

# Self-Supervised Low-Light Quantum RGB Image Demosaicing

Diego Garcia-Barajas, Kebin Contreras, Brayan Monroy and Jorge Bacca

Department of Computer Science, Universidad Industrial de Santander, Colombia

[jbacquin@uis.edu.co](mailto:jbacquin@uis.edu.co)

**Abstract**—Image acquisition in low-light environments is fundamentally challenging due to the photon-limited nature of the scene, which results in severe noise and incomplete color information. Imaging sensors operating under such conditions require robust post-processing to recover visually coherent, full-color images. In these conditions, the photon arrival process can be modeled as a Poisson distribution, which introduces noise that complicates image reconstruction. Furthermore, the use of a color filter array leads to missing color information at each pixel, which exacerbates the challenge. As a result, denoising and demosaicing become ill-posed and interdependent tasks. We propose a self-supervised method that jointly addresses denoising and demosaicing under low-light conditions without requiring clean reference images. Our approach achieves a PSNR higher by 2.0 dB compared to best state-of-the-art methods at gain of 20 and is close to the supervised method.

**Index Terms**—Self-Supervised Learning, Image Denoising, Image Demosaicing, Low-Light, Poisson Noise

## I. INTRODUCTION

In digital photography, low-light environments pose a significant challenge due to the limited number of photons reaching the sensor. This results in image degradation characterized by high levels of noise as many pixels do not accumulate enough light to produce a reliable signal. The random nature of photon arrival under these conditions is commonly modeled as a Poisson process. As a result, an essential post-processing step known as denoising [1] is required to suppress noise and recover a visually coherent and perceptually accurate image.

Recent advancements in sensor design have introduced Quanta Image Sensors (QIS) as a promising alternative for image acquisition in extremely low-light environments, where traditional Complementary Metal-Oxide-Semiconductor (CMOS) sensors often fail. Unlike CMOS sensors, which integrate charge over time and amplify analog signals, QIS operates at the single-photon level using jots, ultrasmall pixels that produce a binary output indicating the presence or absence of photon during a brief exposure [2, 3]. This design reduces readout noise and dark current, enhancing the signal-to-noise ratio in low-light conditions [2, 4].

In addition to the challenges introduced by low-light conditions, QIS sensors, employ a Color Filter Array (CFA), typically arranged in a Bayer pattern [5]. This pattern assigns a specific color filter (red, green, or blue) to each jot in a mosaic layout, resulting in an incomplete color image where two-thirds of the

This work was supported by MINCIENCIAS AMSUD 2023, CTO No. 112721-110-2024 under project "Computational ultrasound imaging: from theory to applications".

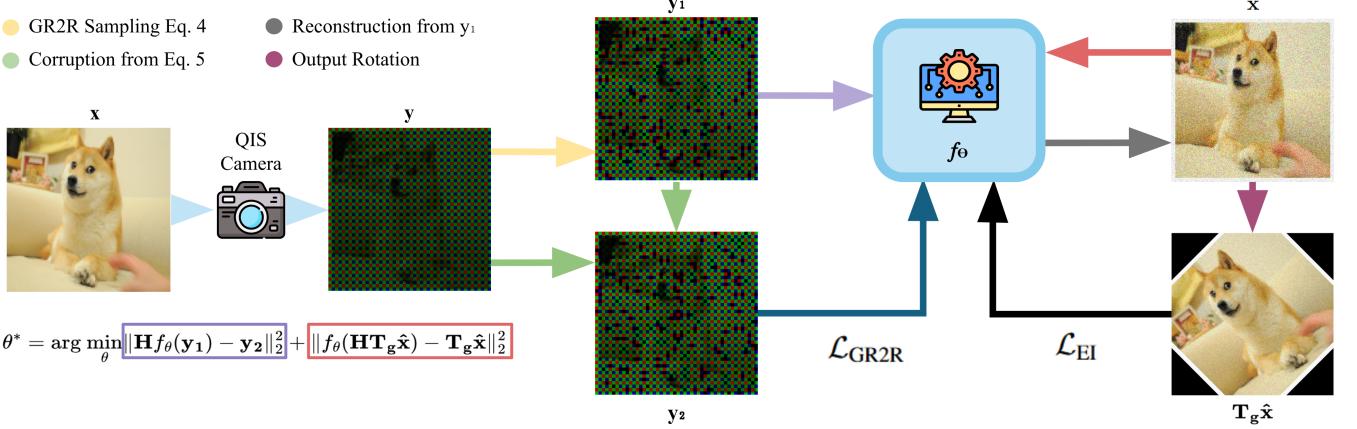
color information is missing. Consequently, reconstructing a full-color image requires an additional processing step known as demosaicing [5, 6].

Denoising and demosaicing are fundamental to reconstructing the captured images and can be solved with optimization algorithms and deep learning techniques consisting of training deep neural networks to solve complex tasks [7]. The most popular of these methods are supervised learning, which consists in using real images to learn a deep model. However, obtaining a ground truth RGB image under low-light conditions is extremely challenging, since in real scenes it is very difficult to acquire a noise-free raw image with accurate color information [8]. In this case, we can use self-supervised learning techniques, consisting of solving the task without labeled data entries [9] or ground truth.

Denoising has been widely studied within the context of self-supervised learning techniques, e.g., Noise2Noise, which trains models on pairs of noisy images [10]; Noise2Void, which predicts missing pixels from their surroundings [11]; Recorrupted2Recorrupted, which leverages multiple corruptions of the same image to train denoising models without clean targets [12]; Generalized Recorrupted2Recorrupted (GR2R), which extends this idea to a broader range of noise models and tasks [13]; and SURE-based methods, which estimate denoising risk directly from noisy observations [14]. Notably, only the last two methods have addressed a broader range of noise types, including Poisson noise, which is particularly relevant to low-light imaging, and their ability to model signal-dependent noise makes them especially important for our study.

In parallel, solve demosaicing has been approached through a variety of techniques, including classical interpolation-based methods such as bilinear and edge-directed interpolation [15]; Joint Demosaicing and Denoising with Self-Guidance, which leverages the higher sampling rate of the green channel to guide the reconstruction of missing color values in a self-supervised manner [16]; JDD-DoubleDIP, a training-free approach that utilizes double deep image priors to perform joint demosaicing and denoising directly on a single RAW image [17]; and Equivariant Imaging (EI), which exploits inherent symmetries in natural images to learn from compressed measurements without requiring ground truth data [18].

Jointly addressing the challenges of demosaicing and denoising is inherently difficult, as both are ill-posed problems that strongly interact with each other. Noise in the sensor data can significantly degrade the accuracy of color interpolation, while poor interpolation may in turn amplify or misrepresent the noise [19]. Tackling these tasks separately often leads to error accumulation, yet solving them together requires models



**Figure 1: Low-light Image Reconstruction Pipeline:** Illustration of a DnCNN-based model designed for low-light Bayer image reconstruction. A photon-limited scene is captured by a camera with a CFA, producing a noisy Bayer image. From this raw sensor output  $y$ , we generate two noisy observations  $y_1$  (Eq. 4) and  $y_2$  (Eq. 5). Reconstructions are produced from  $y_1$ , while  $y_2$  is used for training the model. Specifically,  $\mathbf{T}_g \hat{\mathbf{x}}$  corresponds to passing  $y_1$  through the network  $f_\theta$  and then applying a rotational transformation  $\mathbf{T}_g$ . Our method jointly removes noise and reconstructs the missing color information, ultimately generating a clean, full-color image.

that can reliably distinguish between true image structures and random perturbations, a particularly demanding task in areas with fine textures or sharp edges [19].

We propose a self-supervised approach that resolves missing pixel information using the EI framework with rotational transformations, which allow recovering lost pixel intensity values by repositioning pixels within the image. Simultaneously, we eliminate noise associated with measurements through the GR2R strategy, where multiple re-corruptions enable the reconstruction of a denoised image. Both processes operate without requiring ground truth data, making the method particularly efficient for low-light photography. The results show that our approach maintains stable reconstruction quality across a range of noise levels and demonstrates a clear advantage in high-noise scenarios, where traditional self-supervised techniques tend to degrade more significantly. This suggests that the combination of rotational transformations and re-corruption mechanisms provides a robust framework for image restoration under low-light conditions.

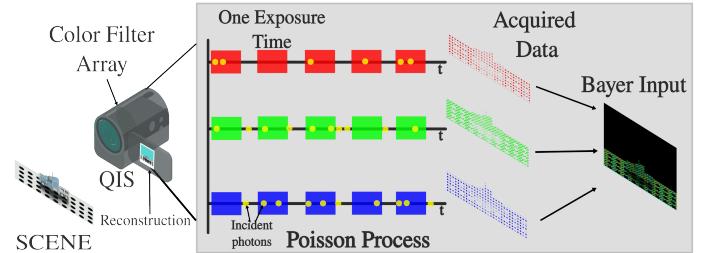
## II. LOW-LIGHT RGB IMAGE MODEL

The data captured by the QIS sensor is governed by two key elements: the use of a Bayer CFA, which assigns a specific color channel (red, green, or blue) to each pixel, and the discrete arrival of photons at each site during the exposure interval. As illustrated in Figure 2, each pixel accumulates photon events according to a Poisson process, depending on its associated color filter. In this way, the sensor generates three separate, sparse color planes, which are then combined into a single raw Bayer image. We model the image formation process under low-light conditions and Bayer CFA sampling as follows:

$$\mathbf{z} \sim \mathcal{P}(\gamma \cdot \mathbf{Hx}), \quad (1)$$

where  $\mathbf{x} \in \mathbb{R}^{3n \times 1}$  is the real scene, represented as a vector containing the true underlying color information of the image before any sensor-related processing takes place,  $\mathbf{H} \in \mathbb{R}^{n \times 3n}$  is

a sampling matrix that models the effect of the color filter array by selecting a single color channel at each pixel location,  $\gamma$  is the gain factor of the sensor, and  $\mathbf{z} \in \mathbb{R}^{n \times 1}$  is our measurement that follows a Poisson process due to the nature of the capture which is represented by  $\mathcal{P}$ .



**Figure 2: QIS Image Acquisition:** Illustration of the image formation process in a QIS with a Bayer CFA. During the exposure time, photons are detected at each jot following a Poisson process. Each jot captures a single color component (red, green, or blue), and the acquired data forms a Bayer-patterned image that serves as input for subsequent demosaicing.

## III. LOW-LIGHT RGB IMAGE DEMOSAICING ALGORITHM

In QIS image restoration, the processes of denoising and demosaicing are addressed jointly to enhance the final image quality. The GR2R approach is used to suppress Poisson noise, which is characteristic of images captured under low-light conditions. This method relies on a self-supervised learning framework that allows the model to be trained using only noisy observations, without requiring clean reference images. The loss function used in GR2R is defined as:

$$\mathcal{L}_{GR2R}(\mathbf{y}; \theta) := \|\mathbf{H}f_{\theta}(\mathbf{y}_1) - \mathbf{y}_2\|_2^2, \quad (2)$$

where  $f_\theta$  denotes the neural network to be trained,  $\mathbf{H}$  is the acquisition operator, and  $\mathbf{y}_1, \mathbf{y}_2$  are two noisy observations

generated from the same base image  $\mathbf{z}$ , using a statistical noise model.

In our case, since the images are captured under low-light conditions, a Poisson noise model is adopted to reflect the nature of the degradation. According to the model in Eq. 1, the observed image can be expressed as

$$\mathbf{y} = \mathbf{z}/\gamma, \quad (3)$$

The division by  $\gamma$  compensates for the variations in photon arrival rates at each pixel, normalizing the observed image in relation to the true image. From this measurement  $\mathbf{y}$ , the two noisy observations  $\mathbf{y}_1$  and  $\mathbf{y}_2$  required for self-supervised training are generated as follows:

$$\mathbf{y}_1 = \frac{\mathbf{y} - \omega/\gamma}{1 - \alpha}, \quad \omega \sim \text{Bin}(\mathbf{z}, \alpha), \quad (4)$$

where  $\text{Bin}(\mathbf{z}, \alpha)$  denotes a sampling of a binomial distribution with probability  $\alpha$  on measurements without normalization.

$$\mathbf{y}_2 = \frac{1}{\alpha}\mathbf{y} - \frac{1 - \alpha}{\alpha}\mathbf{y}_1. \quad (5)$$

These expressions generate two noisy versions of the same normalized measurement  $\mathbf{y}$ , each containing independent realizations of Poisson noise. This enables the model to learn how to predict one noisy observation from another, by leveraging their shared underlying structure. During training, the model is optimized so that the prediction  $f_\theta(\mathbf{y}_1)$ , once passed through the acquisition operator  $\mathbf{H}$ , approximates  $\mathbf{y}_2$ , enabling image restoration to be learned directly from noisy data without external supervision.

The EI method is used to reconstruct missing chromatic information through a self-supervised approach that exploits symmetries present in image data, such as rotations or translations. The key idea behind EI is that natural images exhibit structural invariances that should be preserved during the restoration process. The loss function used in this method is defined as:

$$\mathcal{L}_{\text{EI}}(\mathbf{y}; \theta) := \|f_\theta(\mathbf{HT}_g\hat{\mathbf{x}}) - \mathbf{T}_g\hat{\mathbf{x}}\|_2^2 \quad (6)$$

where  $\hat{\mathbf{x}} = f_\theta(\mathbf{y}_1)$ , and  $\mathbf{T}_g$  is a transformation from a symmetry group (e.g., a rotation or translation) applied to both the reconstructed and target images. This formulation encourages the model to preserve the inherent symmetries of natural images during reconstruction, ultimately improving the accuracy and consistency of the demosaicing process.

The combination of these two techniques is expressed by a composite loss function, which includes the loss associated with noise removal and the loss related to demosaicing. The model parameters  $\theta$  are estimated by solving an optimization problem that minimizes the GR2R loss ( $\mathcal{L}_{\text{GR2R}}$ ) together with the EI loss ( $\mathcal{L}_{\text{EI}}$ ) as:

$$\hat{\theta} \in \arg \min_{\theta} \mathcal{L}_{\text{GR2R}}(\mathbf{y}; \theta) + \lambda \mathcal{L}_{\text{EI}}(\mathbf{y}; \theta). \quad (7)$$

---

**Algorithm 1 Image Restoration Pseudo-code**


---

```
# Input: clean image dataset  $\mathbf{x}$ , Bayer mask  $\mathbf{H}$ , network  $f_\theta$ 
# transformation group  $\mathbf{T}$ , optimizer update

for  $\mathbf{x}$  in dataset:          # iterate over clean images
     $\mathbf{y} \leftarrow \text{Poisson}(\gamma \cdot \mathbf{Hx})$  # simulate noisy measurement
     $\mathbf{y}_1 \leftarrow \text{GR2Rsampling}(\mathbf{y}, \alpha, \gamma)$            # Eq. 4
     $\mathbf{y}_2 \leftarrow \text{Corrupt}(\mathbf{y}, \mathbf{y}_1, \alpha)$            # Eq. 5
     $\hat{\mathbf{x}} \leftarrow f_\theta(\mathbf{y}_1)$                       # restore image from  $\mathbf{y}_1$ 
     $\mathbf{t} \leftarrow \text{select}(\mathbf{T}_g)$                   # random transformation
     $\text{gr2r\_loss} \leftarrow \text{GR2R}(\mathbf{y}_1, \mathbf{y}_2, \mathbf{H}, f_\theta)$       # Eq. 2
     $\text{ei\_loss} \leftarrow \text{EI}(\mathbf{y}_1, \mathbf{t}, \mathbf{H}, \hat{\mathbf{x}}, f_\theta)$       # Eq. 6
     $\text{loss} \leftarrow \text{gr2r\_loss} + \text{ei\_loss}$            # Eq. 7
     $\text{loss.backward}()$ 
     $\text{update}(\theta)$ 
```

---

We opted for DnCNN because it is a well-established, reliable, and fast architecture that delivers the flexibility necessary to distinctly isolate and assess the influence of the proposed framework. DnCNN has demonstrated efficacy in noise reduction for QIS images due to its capability to learn intricate representations while preserving fine details, making it an ideal candidate for our objectives. The pseudocode outlining our methodology, which involves the application of GR2R for denoising followed by EI for demosaicing, is provided in Algorithm 1.

## IV. SIMULATION AND RESULTS

### A. Data

The DIV2K dataset, consisting of 800 images for training and 100 for validation, was used to simulate QIS imaging. To increase the diversity of the training data, a random crop technique was applied to the 800 training images, resulting in a total of 5000 cropped images, each with a size of  $256 \times 256$  pixels. For testing, the original 100 validation images were retained, with a center crop applied to each, producing a size of  $512 \times 512$  pixels. The simulation of real-world data is achieved through image processing techniques that replicate sensor-generated images under low-light conditions. To mimic the behavior of a sensor in such environments, Poisson noise was added to the images from the dataset, reflecting the photon arrival process characteristic of low-light conditions. A Bayer mosaic was also applied to the images to simulate the CFA in the sensor, closely replicating how a sensor captures color information. Different gains of Poisson noise were introduced, mirroring the gains factors observed in the experimental setup. These simulations aim to closely replicate the noise characteristics and spatial resolution typically seen in real QIS sensors under challenging conditions. Figure 3 illustrates the effect of varying Poisson gains and the reconstruction with different methods on a  $512 \times 512$  image captured with a Bayer CFA, highlighting visual degradation as the gain factor  $\gamma$  increases. The code implementation was developed using the DeepInverse library [21].

Table I: Comparison of the PSNR and SSIM values for different gains on the DIV2K dataset

Gain ( $\gamma$ )	Metric	Supervised	Ours	PURE+EI [14, 18]	MC+EI [18]	MC+TV [20]
100	PSNR	30.347 $\pm$ 2.326	27.844 $\pm$ 3.202	<b>27.997 <math>\pm</math> 2.640</b>	25.488 $\pm$ 2.635	25.288 $\pm$ 2.448
	SSIM	0.87 $\pm$ 0.05	0.79 $\pm$ 0.07	<b>0.80 <math>\pm</math> 0.05</b>	0.67 $\pm$ 0.06	0.66 $\pm$ 0.07
20	PSNR	26.710 $\pm$ 2.252	<b>24.492 <math>\pm</math> 2.799</b>	22.332 $\pm$ 2.434	21.298 $\pm$ 2.001	21.448 $\pm$ 1.612
	SSIM	0.76 $\pm$ 0.07	<b>0.67 <math>\pm</math> 0.09</b>	0.55 $\pm$ 0.05	0.42 $\pm$ 0.07	0.45 $\pm$ 0.09
10	PSNR	25.090 $\pm$ 2.186	<b>23.034 <math>\pm</math> 2.694</b>	19.533 $\pm$ 1.925	19.206 $\pm$ 1.874	19.144 $\pm$ 1.484
	SSIM	0.70 $\pm$ 0.08	<b>0.60 <math>\pm</math> 0.10</b>	0.32 $\pm$ 0.05	0.31 $\pm$ 0.06	0.35 $\pm$ 0.09
2	PSNR	21.457 $\pm$ 1.978	<b>19.384 <math>\pm</math> 2.524</b>	15.425 $\pm$ 1.580	13.968 $\pm$ 1.634	12.876 $\pm$ 1.495
	SSIM	0.52 $\pm$ 0.09	<b>0.41 <math>\pm</math> 0.10</b>	0.15 $\pm$ 0.04	0.13 $\pm$ 0.04	0.16 $\pm$ 0.07

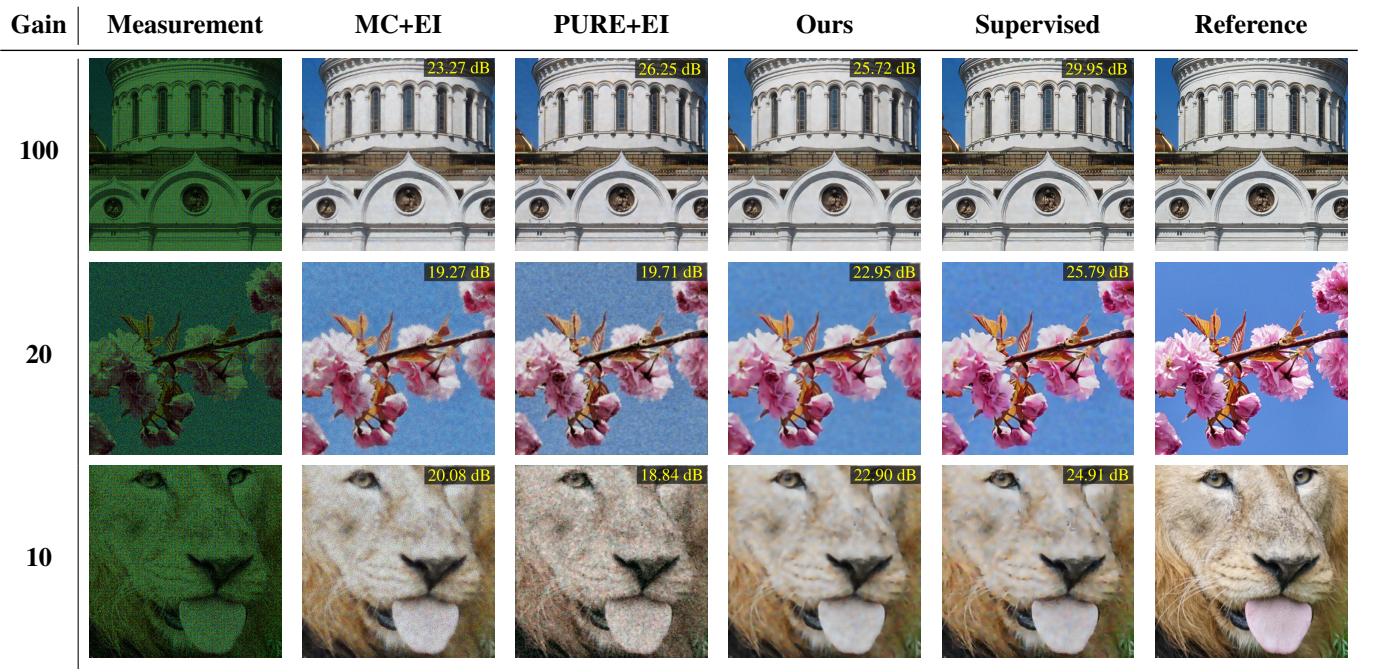


Figure 3: **Visual Results at Varying gain factor:** Qualitative visual comparison of reconstruction methods at different gain. From left to right: input measurement, MC+EI, PURE+EI, our method, supervised baseline, and the ground truth reference. Each row corresponds to a distinct gain setting, illustrating how reconstruction quality varies under increasing noise. The figure highlights the robustness of our approach relative to other self-supervised methods across different noise conditions.

### B. Experiments

The model was trained for 300 epochs with a batch size of 20 and a learning rate of 0.001. A 7-layer DnCNN architecture was used as the backbone, modified to better handle the Poisson noise model by removing the additive residual connection and applying a ReLU activation to the difference between the network output and the input. The training objective included the GR2R loss, estimated using monte carlo estimation with 10 sampling points per batch and configured with  $\alpha = 0.1$ , along with an EI regularization term based on random rotations, weighted by a factor of  $\lambda = 0.01$ . The model was optimized using the Adam algorithm with a weight decay of  $10^{-8}$ , and a step learning rate scheduler was applied after 80% of the total training epochs. Training and evaluation were performed on a GPU, with fixed random seeds to ensure reproducibility. The evaluations were performed with different gain  $\gamma$ , with values at 100, 20, 10, 2. Different approaches were compared, including

a supervised model and recent self-supervised methods such as Measurement Consistency (MC) with Total Variation (TV) [20], MC+EI and PURE method [14] combined to EI, reporting quality metrics such as Peak Signal-to-Noise Ratio (PSNR) [22], which is used to quantify reconstruction fidelity in terms of pixel-wise differences, and Structural Similarity Index (SSIM) [22], which evaluates perceptual similarity based on structural information.

### C. Analysis

Table I shows that our method closely matches the supervised baseline across various gain values, often exceeding other unsupervised methods. Specifically, PURE [14] coupled with EI [18] (PURE+EI) method slightly surpass our approach at a gain of 100, its performance declines notably with increased noise, becoming unstable at lower gain. Similarly, the MC+EI [18]

and MC+TV methods show lower PSNR values throughout the range, with MC+TV generally performing slightly worse than MC+EI. In contrast, our method delivers more consistent and robust results under tougher noise conditions.

Note in Table I that the results obtained with PURE required careful monitoring of the loss behavior and training metrics, since the method tended to diverge when addressing both tasks, indicating instability. For this reason, the best-performing model checkpoints were selected for evaluation. In addition, the EI transformation used in the experiments was based on rotations, which proved highly effective in reconstructing pixel-level information lost due to the CFA in the image acquisition process. This is especially beneficial for the demosaicing, where spatial structure plays a crucial role.

Figure 3 presents a qualitative comparison of reconstruction results across different gain. Each row corresponds to a specific gain value, while each column displays the output of a different method. We can observe that our method achieves better visual results compared to the other approaches, being outperformed only by the supervised method.

For a gain of 100, our method achieves a PSNR of 27.844 dB, only 0.153 dB below PURE+EI, which slightly outperforms us at this gain. However, as the gain decreases, our method advantage becomes more pronounced. At a gain of 20, we outperform PURE+EI by over 2 dB, while also surpassing MC+EI and MC+TV by more than 3 dB. This performance gap widens further at lower gains: at a gain of 10, our method outperforms PURE+EI by more than 3.5 dB and exceeds MC+EI and MC+TV by approximately 4 dB. At the lowest gain of 2, where degradation is most severe, our approach maintains a lead of nearly 4 dB over PURE+EI, more than 5 dB over MC+EI, and over 6 dB above MC+TV. Compared to the supervised baseline, our results remain remarkably close, with PSNR differences consistently under 3 dB across all gain levels, demonstrating robustness even in the most adverse cases. Visually, reconstructions of our method preserve detail and structure, closely resembling supervised results and surpassing other self-supervised approaches.

## V. CONCLUSIONS

This work demonstrates that our proposed method effectively addresses the challenge of reconstructing full-color images from measurements of extremely low light and low photon count, even under high noise conditions where supervised approaches often have difficulties. By operating directly on raw sensor data without requiring clean references, our self-supervised framework maintains stable and accurate reconstruction across a wide range of noise levels. Experimental results confirm that, unlike traditional denoising pipelines that fail under very low photon arrivals, our approach consistently delivers high-quality output both quantitatively (PSNR) and qualitatively, without any additional computational overhead for image recovery. In summary, our method offers a competitive solution for color image reconstruction in extreme low-light scenarios compared to supervised methods, paving the way for reliable imaging in applications such as nocturnal digital photography.

## REFERENCES

- [1] C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo, and C.-W. Lin, “Deep learning on image denoising: An overview,” *Neural Networks*, vol. 131, pp. 251–275, 2020.
- [2] E. R. Fossum, J. Ma, S. Masoodian, L. Anzagira, and R. Zizza, “The quanta image sensor: Every photon counts,” *Sensors*, vol. 16, no. 8, p. 1260, 2016.
- [3] Y. Chi, A. Gnanasambandam, V. Koltun, and S. H. Chan, “Dynamic low-light imaging with quanta image sensors,” in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXI 16*. Springer, 2020, pp. 122–138.
- [4] A. Gnanasambandam and S. H. Chan, “Image classification in the dark using quanta image sensors,” in *European Conference on Computer Vision*. Springer, 2020, pp. 484–501.
- [5] M. Muller, D. Picone, M. Dalla Mura, and M. O. Ulfarsson, “Pattern-invariant unrolling for robust demosaicing,” in *2024 32nd European Signal Processing Conference (EUSIPCO)*. IEEE, 2024, pp. 461–465.
- [6] H. Arguello, S. Pinilla, Y. Peng, H. Ikoma, J. Bacca, and G. Wetzstein, “Shift-variant color-coded diffractive spectral imaging system,” *Optica*, vol. 8, no. 11, pp. 1424–1434, 2021.
- [7] J. Bacca, “Projection-based correction for enhancing deep inverse networks,” *arXiv preprint arXiv:2505.15777*, 2025.
- [8] X. Guo and Q. Hu, “Low-light image enhancement via breaking down the darkness,” *International Journal of Computer Vision*, vol. 131, no. 1, pp. 48–66, 2023.
- [9] R. Balestrieri, M. Ibrahim, V. Sobal, A. Morcos, S. Shekhar, T. Goldstein, F. Bordes, A. Bardes, G. Mialon, Y. Tian *et al.*, “A cookbook of self-supervised learning,” *arXiv preprint arXiv:2304.12210*, 2023.
- [10] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, “Noise2noise: Learning image restoration without clean data,” *arXiv preprint arXiv:1803.04189*, 2018.
- [11] A. Krull, T.-O. Buchholz, and F. Jug, “Noise2void-learning denoising from single noisy images,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 2129–2137.
- [12] T. Pang, H. Zheng, Y. Quan, and H. Ji, “Recorrupted-to-recorrupted: Unsupervised deep learning for image denoising,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 2043–2052.
- [13] B. Monroy, J. Bacca, and J. Tachella, “Generalized recorrupted-to-recorrupted: Self-supervised learning beyond gaussian noise,” in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 28155–28164.
- [14] Y. Le Montagner, E. D. Angelini, and J.-C. Olivo-Marin, “An unbiased risk estimator for image denoising in the presence of mixed poisson-gaussian noise,” *IEEE Transactions on Image processing*, vol. 23, no. 3, pp. 1255–1268, 2014.
- [15] X. Li and M. T. Orchard, “New edge-directed interpolation,” *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1521–1527, 2001.
- [16] L. Liu, X. Jia, J. Liu, and Q. Tian, “Joint demosaicing and denoising with self guidance,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 2240–2249.
- [17] T. Li, A. Lahiri, Y. Dai, and O. Mayer, “Joint demosaicing and denoising with double deep image priors,” *arXiv preprint arXiv:2309.09426*, 2023.
- [18] D. Chen, J. Tachella, and M. E. Davies, “Equivariant imaging: Learning beyond the range space,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 4379–4388.
- [19] Y. Park, S. Lee, B. Jeong, and J. Yoon, “Joint demosaicing and denoising based on a variational deep image prior neural network,” *Sensors*, vol. 20, no. 10, p. 2970, 2020.
- [20] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, T. Pock *et al.*, “An introduction to total variation for image analysis,” *Theoretical foundations and numerical methods for sparse recovery*, vol. 9, no. 263–340, p. 227, 2010.
- [21] J. Tachella, M. Terris, S. Hurault, A. Wang, D. Chen, M.-H. Nguyen, M. Song, T. Davies, L. Davy, J. Dong *et al.*, “Deepinverse: A python package for solving imaging inverse problems with deep learning,” *arXiv preprint arXiv:2505.20160*, 2025.
- [22] A. Hore and D. Ziou, “Image quality metrics: Psnr vs. ssim,” in *2010 20th international conference on pattern recognition*. IEEE, 2010, pp. 2366–2369.