

Generative Quanta Color Imaging

Vishal Purohit, Junjie Luo, Yiheng Chi, Qi Guo, Stanley H. Chan, Qiang Qiu
 Elmore Family School of Electrical and Computer Engineering
 Purdue University, West Lafayette, USA.

{purohitv, luo330, chi14, guo675, stanchan, qqiu}@purdue.edu

Abstract

The astonishing development of single-photon cameras has created an unprecedented opportunity for scientific and industrial imaging. However, the high data throughput generated by these 1-bit sensors creates a significant bottleneck for low-power applications. In this paper, we explore the possibility of generating a color image from a single binary frame of a single-photon camera. We evidently find this problem being particularly difficult to standard colorization approaches due to the substantial degree of exposure variation. The core innovation of our paper is an exposure synthesis model framed under a neural ordinary differential equation (Neural ODE) that allows us to generate a continuum of exposures from a single observation. This innovation ensures consistent exposure in binary images that colorizers take on, resulting in notably enhanced colorization. We demonstrate applications of the method in single-image and burst colorization and show superior generative performance over baselines. Project website can be found at https://vishal-s-p.github.io/projects/2023/generative_quanta_color.html

1. Introduction

Single-photon image sensors are gaining a strong momentum over the past decade due to their impeccable photon-counting capability. This not only opens the door for advanced scientific imaging in low-light environment, but also creates the opportunity for new imaging in high-speed, high dynamic range, and low-bit conditions. Single-photon imaging today are realized through single-photon avalanche diodes (SPAD) and quanta image sensors (QIS). The latest development of QIS has demonstrated a mega-pixel resolution [48] which is compatible with mainstream cell phone cameras, whereas SPAD has achieved hundreds of thousands of (binary) frames per second which has enabled a plethora of industrial and 3D imaging applications [16].

Problem Statement. In the context of 1-bit single-photon imaging, the data volume generated is huge. Taking a 1000×1000 SPAD array as an example, a video of

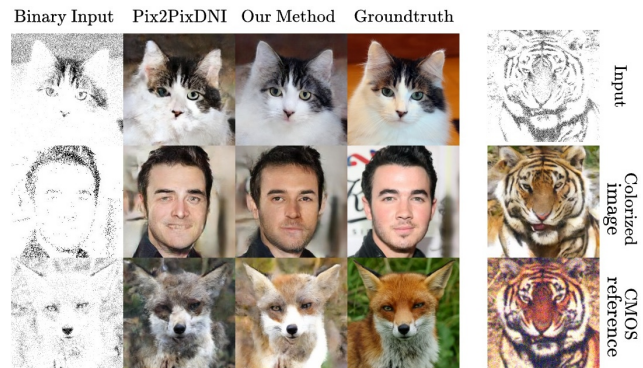


Figure 1. We introduce generative quanta color imaging. Given a binary frame captured by a single-photon camera (quanta image sensor (QIS) in this example), the proposed method generates a continuum of exposures using a neural ordinary differential equation framework. Color are then generated based on these exposures. **Left:** A qualitative comparison between our approach and existing methods is shown. **Right:** Comparison between colorization results and image captured using a RGB CMOS camera under similar conditions.

500 binary frames requires a 500 Mb/s data transfer rate; if the frame rate becomes 150k fps, the data transfer rate increases to 150 Gb/s. Since data throughput scales proportionally with the power consumption, it quickly becomes infeasible to use these sensors continuously without considering power supply, memory, and transmission issues. While solid progress are being made across the sensor and signal processing community [1, 11, 17, 23, 24, 45–47, 49, 50], the data throughput remains an unsolved problem for single-photon sensors to be considered applicable in many applications, including low-power devices, wireless devices, underwater vehicles, spacecrafts, etc., where bandwidth is the bottleneck.

In this paper, instead of compressing data or redesigning the sensor, we ask the question

Is it possible to recover a color image from a *single* 1-bit image?

Significance. There are several reasons why this problem is significant. Firstly, while there is an ocean of literature on image generation (through methods such as Generative Adversarial Networks [26] and diffusion [31]), our problem is unique in the sense none of the existing work addresses it. Besides being binary (and hence colorless), in practice the input data usually also suffers from under- or over-exposures. Therefore, before generating color, it is even more critical to synthesize the unseen exposure. Secondly, any solution offered to our problem will provide a path to generative imaging in the future. From the point of view of compressed acquisition, our method demonstrates the extreme case of compression where the incoming data is a single bit. For future AR/VR applications where power is severely limited, our approach shows the possibility of extreme data compression.

Contributions. The core algorithmic contribution of this paper is two-fold

- We propose a neural ordinary differential equation (neural ODE) based framework to synthesize a continuum of exposures that are not available in the measurement. By controlling the integration interval of neural ODE, we obtain convolutional filters that are responsible for converting input representation to image representation of desired exposure.
- We theoretically and empirically show that a controlled variation of filter atoms leads to controlled exposure changes in the output image.

We verify our proposed method using both synthetic and real single-photon image sensor data. A snapshot of the results is shown in Fig. 1.

2. Related Work

Single-Photon Imaging. The context of this paper is single-photon image sensors, of which we are mostly interested in *one-bit* sensing because of its potential in high-speed and high dynamic range imaging [2]. Today, 1-bit sensing has been realized using quanta image sensors (QIS) [19–21, 50, 67] and single-photon avalanche diodes (SPAD) [1, 16, 17]. From an image reconstruction point of view, since 1-bit data contains so little information, all existing image reconstruction methods need to use multiple frames. For QIS, these methods range from classical optimization [3, 5], transform-denoise [4], neural network [12], color demosaicking [18], frequency demodulation [18], and student-teacher learning [11, 22]. For SPAD, quanta burst photography [51] and its variants [30, 32, 35, 57] first align frames and then apply image fusions [27–29, 59]. The question we ask in this paper is to push the input requirement to the limit: can we recover/generate a color image from a binary input?

Image-to-Image Translation. This paper focuses on using generative models. Generative Adversarial Networks

(GANs) [26] have been a significant breakthrough in the field of image translation. Pix2Pix [33] and CycleGAN [68] demonstrated the effectiveness of GANs in image-to-image translation tasks, with paired and unpaired datasets, respectively. Several subsequent works have focused on improving image fidelity, disentangling style and content, and generating images with varying styles, such as Pix2PixHD [61], StarGAN [13], StyleGAN [37], and StyleGANv2 [14]. Recent advancements in continuous cross-domain translation techniques have further improved the output quality by leveraging images from intermediate domains. These techniques include interpolation of networks [40, 41], interpolation of two latent vectors [14, 37], and exploiting the interpolation path and translation of manifold [9, 42, 55].

Convolutional Filter Decomposition. In the proposed method, one of the key components is the convolutional filter decomposition. Several techniques have been proposed to decompose filters of Convolutional Neural Networks (CNNs) in order to reduce the computation complexity by exploiting the redundancy in convolutional filters [10, 39, 53, 54, 63, 69]. Filter decomposition methods achieve similar performance as compared to the original CNNs while offering speedups in computation. As it is expensive to model the high-dimension space of convolutional filters directly, Qiu *et al.* [56] observed that the convolutional filters can often be approximated using a linear combination small set of basis elements. By decomposing filters into a small set of basis, Wang *et al.* [64, 65] were able to sample the basis and demonstrated theoretical and empirical results on stochastic image generation tasks. Futher, Chen *et al.* [7] showed that filter decomposition can be used to calculate network similarity and large convolutional models can be fine tuned using filter decomposition [8].

3. Proposed Approach

3.1. One-Bit Image Formation Model

We consider a simplified image formation model for 1-bit image sensors. Letting θ be the underlying quanta exposure [20], the observed signal is defined by

$$Y = \text{ADC} \left\{ \text{Poisson}(\theta) + \text{Gaussian}(0, \sigma_r^2) \right\},$$

where σ_r denotes the standard deviation of read noise, and ADC is the analog-to-digital conversion. For 1-bit sensing, the ADC can be modeled as a threshold such that $\text{ADC}(x) = 1$ if $x \geq q$ and $\text{ADC}(x) = 0$ if $x < q$ for some threshold q . Since Y is a binary random variable, it is fully characterized by the probability mass function, which is [2]

$$p_Y(1) = \sum_{k=0}^{\infty} \frac{e^{-\theta} \theta^k}{k!} \Phi \left(\frac{k - q}{\sigma_r} \right),$$

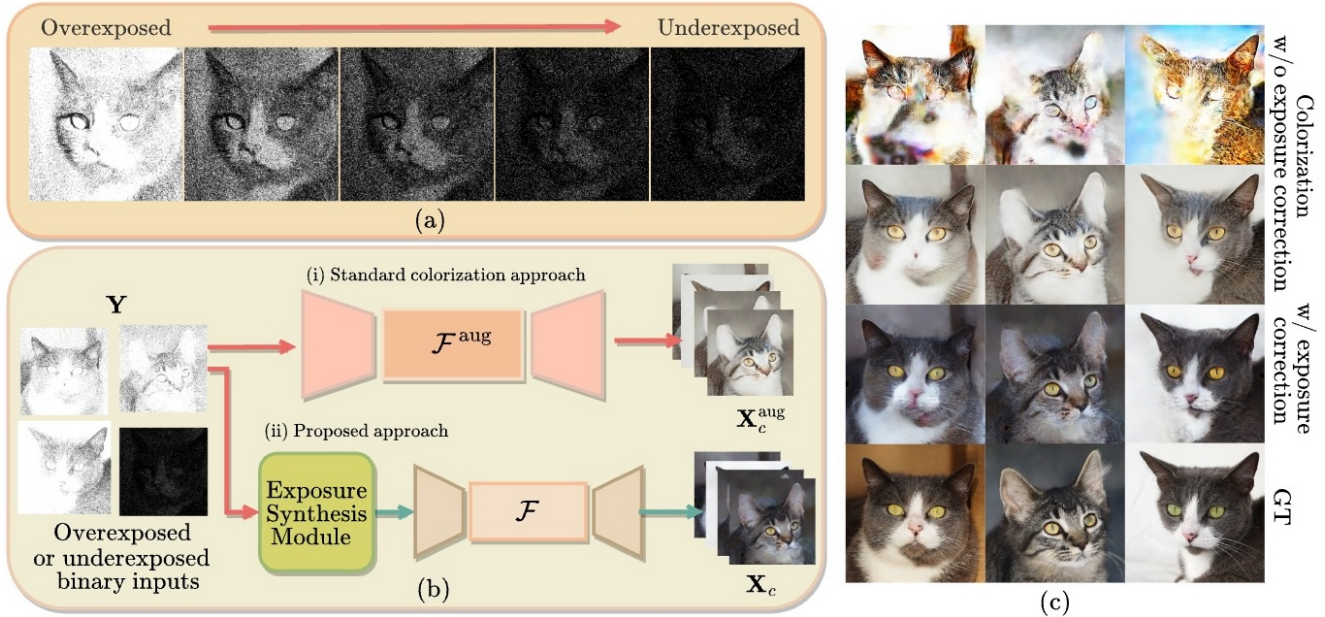


Figure 2. Illustration of exposure correction and colorization of binary images using neural networks. (a) depicts a range of images from overexposed to underexposed, illustrating the degradation of image details due to exposure variation. (b) contrasts the standard colorization workflow and our proposed approach. (i) In standard colorization approaches, a neural network learns to map a binary image Y to the corresponding color image X_c via a neural network, \mathcal{F}^{aug} , where superscript ‘aug’ indicates the colorizer is trained using dataset with augmented exposure images. (ii) In contrast, our approach does not require training colorizer with augmented exposure images. (c) compares the colorization results: the first row is the output of a colorizer trained without augmentation, the second row is the output of colorizer trained with augmented data, the third row corresponds to the results of our method and the last provides the ground truth images for reference.

where $\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} \exp\{-t^2/2\} dt$ is the cumulative distribution of a standard Gaussian. It then follows that the mean (or the bit density) is $\mu = \mathbb{E}[Y] = p_Y(1)$.

Method	Generator Size	Dataset Size	Training Time	FID
Standard colorization	ResNet-9	3x	1x	62.34
+ Increased training time	ResNet-9	3x	1.5x	54.25
+ Increased model size	ResNet-10	3x	1x	74.58
Proposed method	ResNet-9	1x	1x	49.88

Table 1. Performance comparison of the standard colorization approach under various training conditions versus the proposed method.

3.2. Why Exposure Synthesis is Required for Colorization?

Given a binary image $Y \in \mathbb{R}^N$, our goal is to recover a color image $X_c \in \mathbb{R}^{N \times 3}$. However, since the measurement does not contain any color information, our problem fits the classical work of image colorization [38, 58, 60, 66]. However, it is non-trivial to deploy such an extension in real world scenarios, because very likely there is a discrepancy in the exposure levels of the input images compared to those used during training.

An example of binary images of different exposures are shown in Figure 2(a) illustrating the effect of exposure change on the details in the captured binary image. Consider a typical colorization network for example, an Image-to-Image translation model like Pix2Pix [34] trained on well exposed image but tested on overexposed images, the colorization results are notably poorer (shown in Figure 2(c) row 1). This is a realistic concern because the underlying exposure θ can fluctuate dramatically in real life.

Possible solutions include (i) **Exposure Augmentation** : training dataset is enriched with images of varying exposures, at a cost of additional training time. We find, however, this approach partially addresses the issue, as demonstrated in our results shown in Figure 2(b) row 2. In particular, this approach falls short at over- or under-exposure. One can alleviate this issue by applying multiple colorizers, each handling a narrow range of exposures, but this is not ideal as it demands substantially more computational resources. (ii) **Increased Model size** : Another solution is to increase the size of generator used in Pix2Pix [34] and train for longer period of time. In Table 1, we present the colorization performance represented by the FID metric under various training scenarios. Under increased generator size or increased training time the standard colorization ap-

proach performs poorly compared to proposed method detailed in subsequent sections.

Proposed Solution. We address the challenge by a dedicated exposure correction based on neural ordinary differential equations (neural ODEs) to standardize the exposure of binary images. This specialization allows the colorizer to focus on consistent-exposure images, resulting in notably enhanced colorization, as evidenced Figure 2(c) row 3. Additionally, our exposure synthesis can generate multiple exposures from a single binary image, enabling more effective multi-exposure colorization. Some of the added advantages of our approach include: (a) **Modularity.** The colorizer in our design is interchangeable, allowing for upgrades to more advanced architectures without requiring retraining on a large, augmented dataset. (b) **Parameter Efficiency.** Our method achieves its objectives without resorting to multiple networks with large parameter sets, ensuring a more streamlined and resource-efficient process.

3.3. Continuous Exposure Synthesis using Filter Decomposed Neural ODEs

The image formation model defined in Section 3.1 indicates that we can formulate the continuum burst of binary images as $\mathbf{Y}(\tilde{\theta})$, where the binary image \mathbf{Y} is a function of a continuous exposure value $\tilde{\theta} \in (0, 1]$. A small $\tilde{\theta}$ value indicates a higher exposed image \mathbf{Y} . One of the core contributions of this work is a deep neural network architecture that synthesizes a binary image \mathbf{Y}_t at a target exposure $\tilde{\theta}_{\text{target}}$ given an input binary image \mathbf{Y}_i measured at the exposure value $\tilde{\theta}_{\text{input}}$.

As the exposure values are continuous, we may need an exhaustive training set that densely samples every possible pair of input exposure value $\tilde{\theta}_{\text{input}}$ and target exposure value $\tilde{\theta}_{\text{target}}$ for every scene to train our network. To tackle this problem, we designed a convolutional architecture that enforces the parameter of the network to smoothly change according to the exposures.

We use the convolution filter decomposition technique [56, 61, 64] to represent the convolutional filters \mathbf{F} with a low number of parameters, and we leverage neural ODEs [6] to associate parameters with image exposures. With these techniques, we enable the adaptive control of filter parameters \mathbf{F} and thereby the learning of the exposure synthesis operation \mathcal{G} efficiently.

Efficient Modelling of Filter Subspace. Specifically, we formulate the filter parameters \mathbf{F} using the atom-coefficient [56] decomposition, which permits the smallest computational bottleneck in modeling filter subspaces among the candidates. Given a convolutional filter $\mathbf{F} \in \mathbb{R}^{c' \times c \times k \times k}$ where c , c' and k are the output, input channels, and spatial size, respectively. The atom-coefficient decomposition over a set of m filter atoms decomposes \mathbf{F} into coefficients, denoted by $\phi \in \mathbb{R}^{c' \times c \times m}$, and atoms, denoted

by $\Lambda \in \mathbb{R}^{m \times k \times k}$

$$\mathbf{F} = \phi \Lambda.$$

When $m \ll c$, the dimension of the parameters is significantly reduced. Extending the decomposition to a L -layer convolutional network gives a set of L filter atoms $\Lambda = \{\Lambda^l\}_{l=1}^L$ and L coefficients $\phi = \{\phi^l\}_{l=1}^L$.

Enforcing Filter Atom Continuity. We enforce the filter atoms Λ to smoothly change according to the input and target exposure levels $\tilde{\theta}_{\text{input}}$ and $\tilde{\theta}_{\text{target}}$. To do so, we incorporate the continuous-time networks like neural ODEs to describe the relationship of $\Lambda(\tilde{\theta}_{\text{input}}, \tilde{\theta}_{\text{target}})$, as the neural ODEs provide advantages in the number of parameters required, memory consumption, and computational adaptivity [6]. The derivative of the filter atoms $d\Lambda/d\tilde{\theta}$ w.r.t. the exposure value $\tilde{\theta}$ as a neural network $g(\cdot)$

$$d\Lambda/d\tilde{\theta} = g(\tilde{\theta}_{\text{input}}, \tilde{\theta}_{\text{target}}, \Lambda_{\text{init}}; \mathcal{W}),$$

where Λ_{init} is the initial state of filter atoms, and \mathcal{W} is the learnable parameters of $g(\cdot)$. The initial condition Λ_{init} is a learnable parameter and is optimized jointly with \mathcal{W} . Given the input and target exposure values $\tilde{\theta}_{\text{input}}$ and $\tilde{\theta}_{\text{target}}$, the corresponding filter atoms can be calculated as

$$\Lambda(\tilde{\theta}_{\text{target}}, \tilde{\theta}_{\text{input}}) = \Lambda_{\text{init}} + \int_{\tilde{\theta}_{\text{input}}}^{\tilde{\theta}_{\text{target}}} g(\Lambda_{\text{init}}, \mathcal{W}) d\tilde{\theta}.$$

According to Picard-Lindelöf existence theorem [15], for a given initial condition, the solution to the IVP is unique. Theoretically, we prove continuity in filter atoms induces controlled changes in image exposure.

Theorem 3.1 *If Λ_1 and Λ_2 are filter atoms generated with neural ODE using integration intervals $(\tilde{\theta}_0, \tilde{\theta}_1)$ and $(\tilde{\theta}_0, \tilde{\theta}_2)$, respectively such that $\|\Lambda_1 - \Lambda_2\| \leq \epsilon|\tilde{\theta}_1 - \tilde{\theta}_2|$, for $\epsilon > 0$. If $\tilde{\theta}_1$ and $\tilde{\theta}_2$ are the exposures at $\tilde{\theta}_1$ and $\tilde{\theta}_2$, and activation function σ is non-expansive, then change in exposure $\Delta\theta$ is continuous in filter atoms.*

Details of Theorem 3.1 and its proof can be found in the Appendix section A1. An illustration of our parameter efficient approach is shown in Figure 3.

3.4. Exposure Adaptive Colorization

Our proposed exposure adaptive colorization consists of an exposure correction realized by the stated synthesis and exposure level specific colorizer.

Exposure correction. Given a colorization method specific to exposure $\tilde{\theta}_{\text{target}}$ and an input sample $(\mathbf{Y}_i, \tilde{\theta}_{\text{input}})$, the proposed exposure correction generates image $\tilde{\mathbf{Y}}_t$ at the target exposure in two steps. With the operation \mathcal{G} carried out

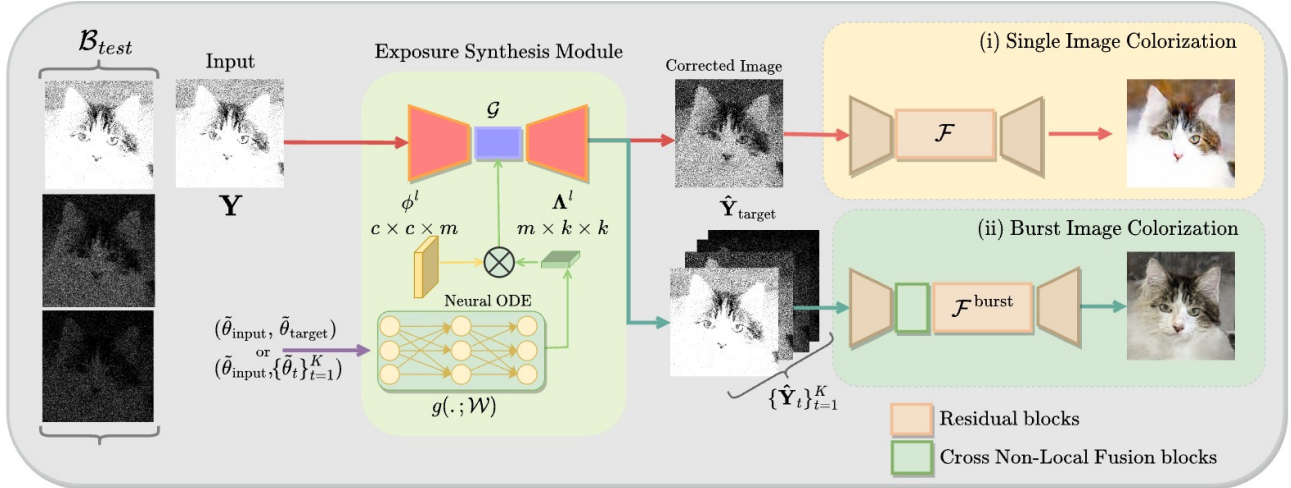


Figure 3. An illustration of our proposed method for exposure adaptive colorization: A binary image, \mathbf{Y} , which can be overexposed or underexposed, is input into the proposed exposure synthesis module. The colorization of binary image can be achieved using Single Image Colorization (SIC) or Burst Image Colorization (BIC). (i) **SIC**: Based on the input and target exposure levels, $\tilde{\theta}_{\text{input}}$ and $\tilde{\theta}_{\text{target}}$, this module adjusts the weights of an exposure synthesis network \mathcal{G} , which then generates an exposure-corrected image. Note that corrected image is not necessarily a binary image. Since the colorization module \mathcal{F} is trained to colorize only image of specific exposure, the exposure synthesis module ensures corrected binary image has similar exposure to the one on which \mathcal{F} is trained. (ii) **BIC**: For BIC we generate images with varying exposures as input to the burst image colorization network, $\mathcal{F}_{\text{burst}}$. The trained network is able to exploit the complementary information across multiple exposures with the help of Cross Non-Local Fusion blocks [44] to synthesize colors in regions of the image that is otherwise not possible by SIC approach.

by some generative network, first, the filter atoms of the network’s convolutional filters are computed using the neural ODEs based on the $(\tilde{\theta}_{\text{input}}, \tilde{\theta}_{\text{target}})$ pair; next, the convolutional filter weights of the network are updated using the filter atoms and their corresponding coefficients, and the network generate the exposure corrected image $\hat{\mathbf{Y}}_t$ based on \mathbf{Y}_i . In the case of burst colorization, multiple images are generated corresponding to a set of target exposures.

Colorization. Both single-image-based colorization and burst-based colorization are considered in our work. The first approach involves a colorization network $\mathcal{F} : \hat{\mathbf{Y}}_t \rightarrow \mathbf{X}_c$ which takes a exposure corrected binary image as input and outputs corresponding RGB image \mathbf{X}_c . The network is trained exclusively at a fixed exposure level, and variations in the exposure of the input image is pre-adjusted by the exposure correction step. The second approach uses a colorization network $\mathcal{F}_{\text{burst}} : \{\hat{\mathbf{Y}}_t\}_{t=1}^K \rightarrow \mathbf{X}_c$ that takes in a burst of K binary images $\{\hat{\mathbf{Y}}_t\}_{t=1}^K$ corresponding to pre-defined varying exposure values $\{\theta_t\}_{t=1}^K$ and outputs one RGB image \mathbf{X}_c .

4. Experiments

4.1. Experimental Setup

We evaluate the performance of both the exposure synthesis module standalone and the entire proposed approach. We conduct exposure burst recovery experiment to evalu-

ate the exposure synthesis and Single Image Colorization (SIC) and Burst Image Colorization (BIC) experiments to evaluate the final colorization results.

Implementation of Exposure Adaptive Colorization.

We implement the neural ODEs using a 6-layer multilayer perceptron. The generative networks we use for exposure synthesis are variants of Pix2Pix [34] and CycleGAN [68], with which the models are named AtomODE-Pix2Pix and AtomODE-CycleGAN. For colorization, we train a Pix2Pix colorizer dedicated to SIC task and another to BIC. For BIC, the colorizer is additionally equipped with a Cross Non-Local Fusion (CNLF) [44] to constructively fuse the representations obtained from the burst. Note that the colorizers are trained directly on pairs of binary and color image data at fixed $\tilde{\theta}$, independent of exposure synthesis. In our colorization experiment, the same trained colorizer is applied on all exposure correction methods for comparison. Details can be found in the Appendix section A2.

Datasets and Synthetic Data Generation. We train and evaluate the proposed method as well as the baselines on synthetic data generated using images from two datasets: AFHQ [14] and CelebA-HQ [36]. AFHQ dataset contains images of three classes, cats, dogs, and wild animals, and CelebA-HQ contains images of male and female human faces.

To synthesize binary images, we use the image formation model described in Section 3.1 and apply *exposure bracket-*

Exposure Synthesis Method	AFHQ (512 x 512)						CelebA-HQ (256 x 256)			
	Cat		Dog		Wild		Male		Female	
	MSE (\downarrow) / RL (\uparrow) / FID (\downarrow)	MSE (\downarrow) / RL (\uparrow) / FID (\downarrow)	MSE (\downarrow) / RL (\uparrow) / FID (\downarrow)	MSE (\downarrow) / RL (\uparrow) / FID (\downarrow)	MSE (\downarrow) / RL (\uparrow) / FID (\downarrow)	MSE (\downarrow) / RL (\uparrow) / FID (\downarrow)	MSE (\downarrow) / RL (\uparrow) / FID (\downarrow)	MSE (\downarrow) / RL (\uparrow) / FID (\downarrow)	MSE (\downarrow) / RL (\uparrow) / FID (\downarrow)	MSE (\downarrow) / RL (\uparrow) / FID (\downarrow)
Pix2Pix-DNI	3.75 / 0.9816 / 61.89	5.91 / 0.9844 / 95.43	8.13 / 0.9780 / 39.29	3.74 / 0.9563 / 56.13	4.78 / 0.9681 / 39.81					
CycleGAN-DNI	9.61 / 0.9816 / 207.82	16.79 / 0.9824 / 220.90	9.62 / 0.9820 / 336.38	9.31 / 0.9824 / 94.78	14.26 / 0.9810 / 250.61					
SAVI2I	41.15 / 46.27 / 178.29	46.27 / 0.9804 / 217.98	38.87 / 0.9019 / 139.47	67.26 / 0.9563 / 111.05	69.21 / 0.9698 / 261.45					
DLOW	2.44 / 0.9654 / 92.30	2.32 / 0.9489 / 173.65	2.75 / 0.9632 / 179.23	1.87 / 0.9655 / 62.57	1.37 / 0.9798 / 63.31					
AtomODE-Pix2Pix (ours)	2.29 / 0.9997 / 57.88	3.17 / 0.9994 / 140.23	2.15 / 0.9998 / 79.76	2.15 / 0.9995 / 53.44	3.20 / 0.9995 / 40.08					
AtomODE-CycleGAN (ours)	1.15 / 0.9948 / 60.54	1.18 / 0.9971 / 87.43	1.23 / 0.9979 / 35.75	1.58 / 0.9986 / 77.82	1.87 / 0.9994 / 62.59					

Table 2. Colorization results on AFHQ and CelebA-HQ datasets. The input image for all these experiments is an overexposed image at index $\tau = 0$ of exposure burst. All the colorizers are trained using an image corresponding to $\tau = 10$ in the exposure burst. At test time we correct the overexposed input to correspond to the one used during training. In the first column of the table, we note the combination of the exposure correction method used and colorizer is same across all the methods.

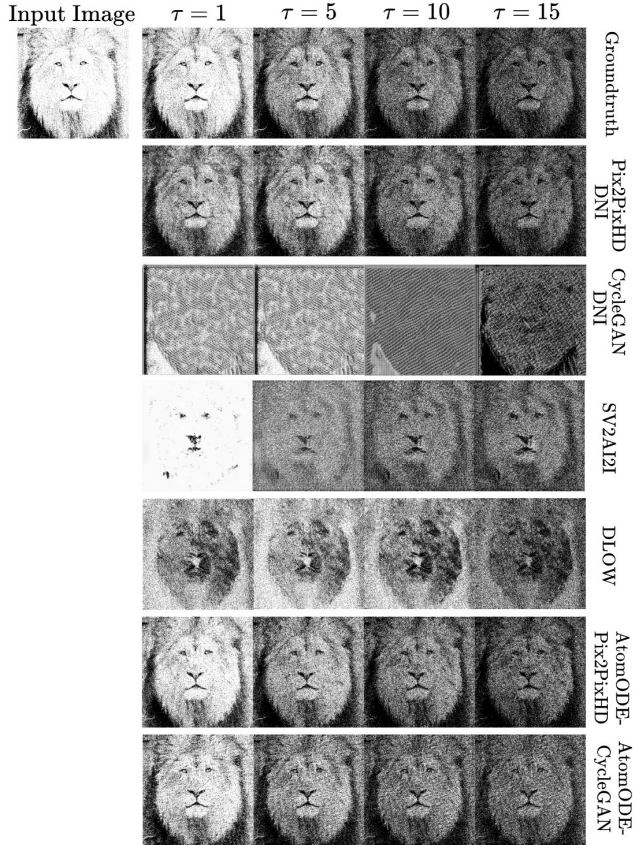


Figure 4. Qualitative results for exposure burst recovery experiments for AFHQ dataset. We compare our methods, AtomODE-Pix2Pix and AtomODE-CycleGAN, with DLOW, SAVI2I and DNI.

ing to generate various exposure. Details of data synthesis can be found in the Appendix section A2. Our compiled test datasets consist of 15-frame bursts of exposure varying from overexposed to underexposed. We index the exposure of these images in a burst by τ , a discretized label of $\tilde{\theta}$.

Baselines. In the exposure burst recovery experiment, we compare our approach against DLOW [25] and SAVI2I [52]. We further consider DNI [62] for its capability of *interpolating* network parameters to generate continuous variations of outputs. We prepare baselines Pix2Pix-DNI and CycleGAN-DNI by coupling DNI with Pix2Pix and Cycle-

GAN exposure generators. In colorization experiments, we apply the independently trained Pix2Pix colorizers to the exposure correction results generated by all methods. For more details on baseline training and hyperparameters, we refer readers to the Appendix section A2.

Metrics. In the burst recovery experiment, we choose to use Mean Squared Error (MSE) and Relative Linearity (RL) [43] to distinguish a smoothly varying exposure burst with consistent appearance changes. A high RL score indicates a better transition in exposure among consecutive samples. In the colorization experiments, we assess the quality of exposure correction by computing the Fréchet Inception Distance (FID) between the colorized image and ground truth.

4.2. Exposure Burst Recovery Results

Figure 4 presents the qualitative comparison of all methods on AFHQ dataset. We find our methods, AtomODE Pix2Pix and AtomODE CycleGAN are able to generate consistent set of smoothly varying exposure images than baseline methods and match ground truth images generated from exposure bracketing. We also observe some baselines struggle with generation of overexposed frames. All methods for comparison are further assessed on both datasets using MSE and RL, with the quantitative results shown in Table 2. Results suggest that the proposed AtomODE provides consistently better performance in all testing conditions across multiple datasets. Additional qualitative results can be found in Appendix section A4.

4.3. Single and Burst Image Colorization Results

The SIC colorization results of all methods on AFHQ dataset are shown in Figure 5, where the exposure of the test samples is different from the that of the colorizer’s training samples. Corrupted colorized images are generated by most of the baselines, while the proposed method produces visually appealing outputs. Figure 7 shows the SIC colorization results on CelebA-HQ dataset. Additionally, we also present the qualitative results for BIC in Figure 6 where K denotes the number of binary frames input to the colorizer. The quantitative results are shown in Table 2, assessed with the FID metric.

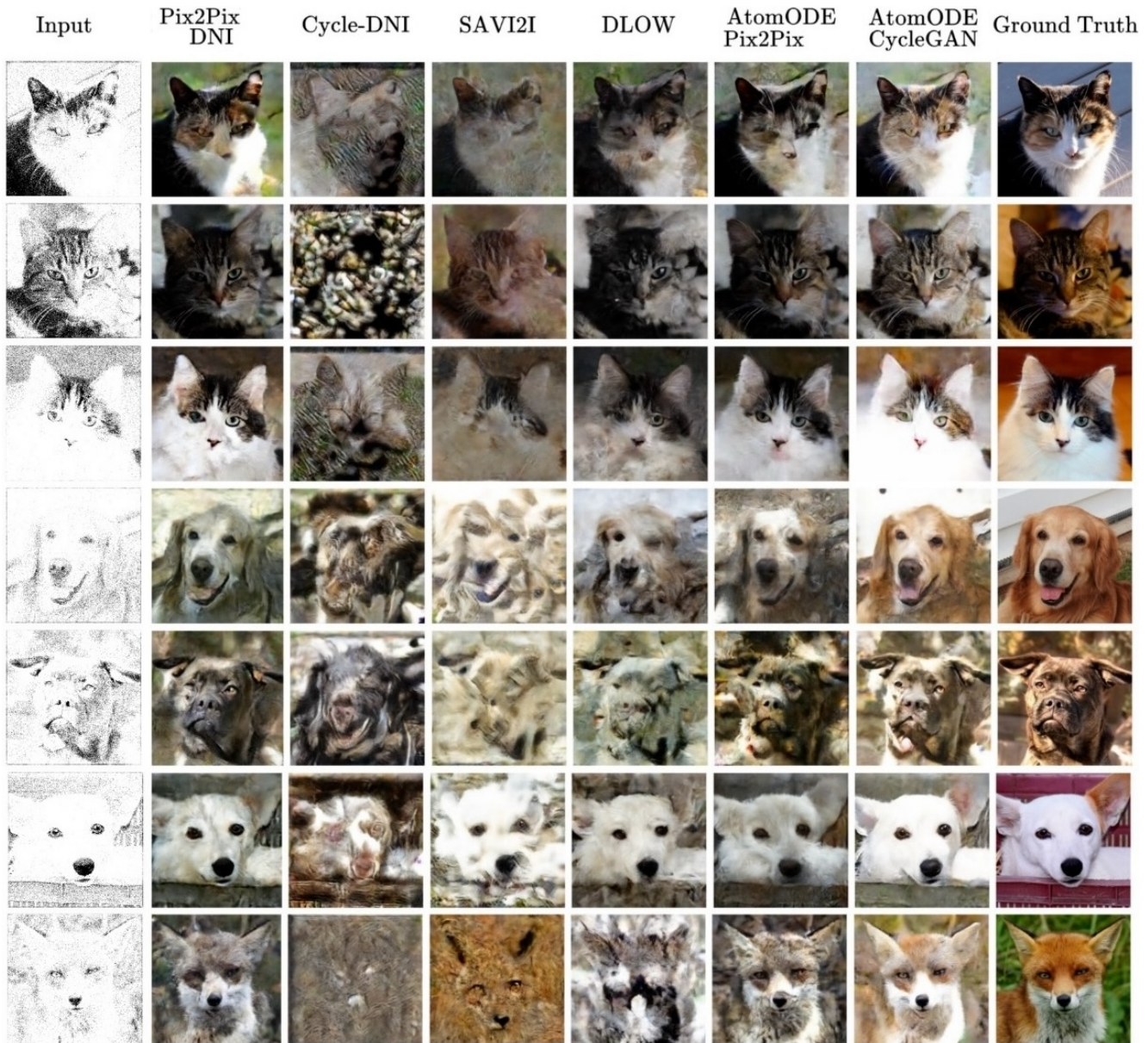


Figure 5. Colorized results for AFHQ dataset using various combinations of exposure correction methods and colorizer. The first column shows the overexposed binary input image for exposure correction. The results of our methods **AtomODE Pix2Pix** and **AtomODE CycleGAN** are shown in columns six and seven.

4.4. Real Data Experiments

As our approach is trained on synthetic data, it is essential to evaluate its performance on real-world data. To this end, we conduct experiments using both a CMOS camera and a prototype QIS camera. We refer the readers to Appendix for further details about the experiment setup. We print 100 images from the training set and 25 from the test set, and we capture pictures of them using both cameras, collecting 125 pairs of grayscale and RGB images. We fine-tune the colorization model with the 100 real data samples, and we evaluate our proposed approach on the set of 25 test data, in-

cluding both binarized CMOS data and QIS data. In Figure 8, we present qualitative results from the evaluation on the real data obtained by the CMOS camera. Despite that the proposed method was not trained with vast real sensor data, it is able to produce visually pleasing colorized images. We observe that fine-tuning the model using a small training dataset significantly improves the colorization quality. We further present results from the evaluation on images captured by the QIS camera in Figure 9. These results demonstrate the potential of the proposed method for real-world single-photon imaging applications.



Figure 6. An illustration of burst image colorization for AFHQ dataset using multiple binary images of varying exposures as input. The colorizer is trained using images of exposure $\tau = \{0, 4, 8, 10, 14\}$. An overexposed input image is used to synthesize multiple images of varying exposures and passed to the colorizer.

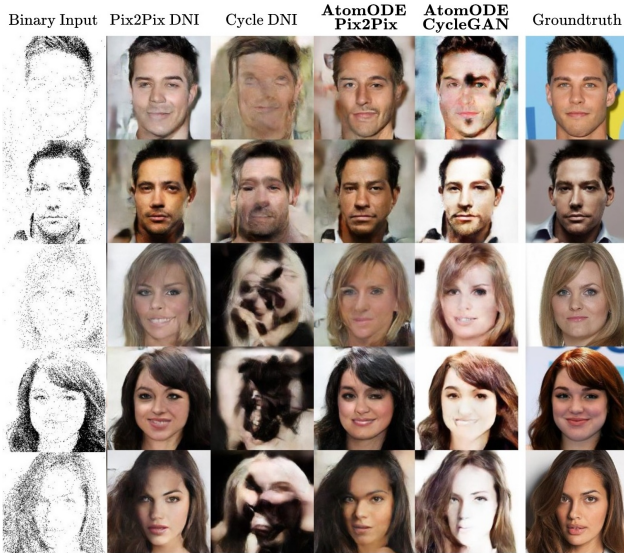


Figure 7. Colorized results for CelebA-HQ dataset using various combinations of exposure correction methods and colorizers. The results of our method are shown in columns four and five.

4.5. Benefits for Practical Applications

Our methodology is specifically designed with modularity and parameter efficiency in mind, features crucial for real-world industry applications. The Pix2PixDNI method, the second-best performing approach, necessitates training multiple networks and either interpolating model weights on-the-fly during inference or precomputing and storing them to generate images with varied exposures. Conversely, our method circumvents the additional storage requirements and the need to train multiple model parameters by incorporating the exposure correction into the neural ODE framework.



Figure 8. Results on real-world images captured from a CMOS camera (images shown along the row with a red line are colorized images, and along the green line are the ground-truth images captured using an RGB CMOS camera).



Figure 9. Colorized output using input from prototype QIS camera.

5. Conclusion

In this work, we proposed a generative exposure correction approach that can synthesize a spectrum of exposure given a single 1-bit image. We modeled the binary signal variation with respect to exposure change as a continuous function by adapting a small set of parameters of a network, known as filter atoms, to desired exposure using neural ODE. Theoretically and empirically, we demonstrated successful recovery of bursts of 1-bit images from a single input image, and we verified its application on single-image-based and burst-image-based 1-bit image colorization. Experiments on real CMOS and QIS data further indicated the potentials of the proposed approach on real-world single-photon imaging applications.

References

- [1] Samuel Burri, Yuki Maruyama, Xavier Michalet, Francesco Regazzoni, Claudio Bruschini, and Edoardo Charbon. Architecture and applications of a high resolution gated SPAD image sensor. *Optics Express*, 22(14):17573–17589, 2014. 1, 2
- [2] Stanley H. Chan. What does a one-bit Quanta Image Sensor offer? *IEEE Transactions on Computational Imaging*, 8: 770–783, 2022. 2
- [3] Stanley H. Chan and Yue M. Lu. Efficient image reconstruction for gigapixel quantum image sensors. In *IEEE Global Conference on Signal and Information Processing*, pages 312–316, 2014. 2
- [4] Stanley H. Chan, Omar A. Elgendy, and Xiran Wang. Images from bits: Non-iterative image reconstruction for Quanta Image Sensors. *MDPI Sensors*, 16(11), 2016. 2
- [5] Stanley H. Chan, Xiran Wang, and Omar A. Elgendy. Plug-and-Play ADMM for image restoration: Fixed-point convergence and applications. *IEEE Transactions on Computational Imaging*, 3:84–98, 2016. 2
- [6] Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David K. Duvenaud. Neural Ordinary Differential Equations. In *NeurIPS*, 2018. 4
- [7] Wei Chen, Zichen Miao, and Qiang Qiu. Inner product-based neural network similarity. In *NeurIPS*, 2023. 2
- [8] Wei Chen, Zichen Miao, and Qiang Qiu. Large convolutional model tuning via filter subspace. Available online : <https://arxiv.org/abs/2403.00269>, 2024. Accessed 03/06/2024. 2
- [9] Ying-Cong Chen, Xiaogang Xu, Zhuotao Tian, and Jiaya Jia. Homomorphic latent space interpolation for unpaired Image-To-Image translation. In *CVPR*, pages 2403–2411, 2019. 2
- [10] Xiuyuan Cheng, Qiang Qiu, Robert Calderbank, and Guillermo Sapiro. RotDCF: Decomposition of convolutional filters for rotation-equivariant deep networks. In *ICLR*, 2019. 2
- [11] Yiheng Chi, Abhiram Gnanasambandam, Vladlen Koltun, and Stanley H. Chan. Dynamic low-light imaging with Quanta Image Sensors. In *ECCV*, pages 122–138, 2020. 1, 2
- [12] Joon Hee Choi, Omar A. Elgendy, and Stanley H. Chan. Image reconstruction for Quanta Image Sensors using deep neural networks. In *IEEE ICASSP*, pages 6543–6547, 2018. 2
- [13] Yunjey Choi, Minje Choi, Munyoung Kim, Jung-Woo Ha, Sunghun Kim, and Jaegul Choo. StarGAN: Unified Generative Adversarial Networks for multi-domain Image-to-Image translation. In *CVPR*, pages 8789–8797, 2018. 2
- [14] Yunjey Choi, Youngjung Uh, Jaejun Yoo, and Jung-Woo Ha. StarGAN v2: Diverse image synthesis for multiple domains. In *CVPR*, pages 8185–8194, 2020. 2, 5
- [15] Earl A. Coddington and Norman Levinson. *Theory of Ordinary Differential Equations*. McGraw-Hill, 1955. 4
- [16] Neale A. W. Dutton, Istvan Gyongy, Luca Parmesan, Salvatore Gneccchi, Neil Calder, Bruce R. Rae, Sara Pellegrini, Lindsay A. Grant, and Robert K. Henderson. A SPAD-based QVGA image sensor for single-photon counting and quanta imaging. *IEEE Transactions on Electron Devices*, 63(1): 189–196, 2016. 1, 2
- [17] Neale A. W. Dutton, Istvan Gyongy, Luca Parmesan, and Robert K. Henderson. Single photon counting performance and noise analysis of CMOS SPAD-based image sensors. *MDPI Sensors*, 16(7), 2016. 1, 2
- [18] Omar A. Elgendy, Abhiram Gnanasambandam, Stanley H. Chan, and Jiaju Ma. Low-light demosaicking and denoising for small pixels using learned frequency selection. *IEEE Transactions on Computational Imaging*, 7:137–150, 2021. 2
- [19] Eric R. Fossum. Some thoughts on future digital still cameras. In *Image Sensors and Signal Processing for Digital Still Cameras*, pages 305–314. Taylor and Francis Group, 2005. 2
- [20] Eric R. Fossum. Modeling the performance of single-bit and multi-bit Quanta Image Sensors. *IEEE Journal of the Electron Devices Society*, 1(9):166–174, 2013. 2
- [21] Eric R. Fossum. Analog read noise and quantizer threshold estimation from Quanta Image Sensor bit density. *IEEE Journal of the Electron Devices Society*, 10:269–274, 2022. 2
- [22] Abhiram Gnanasambandam and Stanley H. Chan. Image classification in the dark using Quanta Image Sensors. In *ECCV*, pages 484–501, 2020. 2
- [23] Abhiram Gnanasambandam and Stanley H. Chan. HDRimaging with Quanta Image Sensors: Theoretical limits and optimal reconstruction. *IEEE Transactions on Computational Imaging*, 6:1571–1585, 2020. 1
- [24] Abhiram Gnanasambandam and Stanley H. Chan. Exposure-referred signal-to-noise ratio for digital image sensors. *IEEE Transactions on Computational Imaging*, 8:561–575, 2022. 1
- [25] Rui Gong, Wen Li, Yuhua Chen, and Luc Van Gool. DLOW: Domain flow for adaptation and generalization. In *CVPR*, pages 2472–2481, 2019. 6
- [26] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Nets. In *NeurIPS*, 2014. 2
- [27] Anant Gupta, Atul Ingle, and Mohit Gupta. Asynchronous single-photon 3D imaging. In *ICCV*, pages 7908–7917, 2019. 2
- [28] Anant Gupta, Atul Ingle, Andreas Velten, and Mohit Gupta. Photon-flooded single-photon 3D cameras. In *CVPR*, pages 6763–6772, 2019.
- [29] Felipe Gutierrez-Barragan, Atul Ingle, Trevor Seets, Mohit Gupta, and Andreas Velten. Compressive single-photon 3D cameras. In *CVPR*, pages 17833–17843, 2022. 2
- [30] Istvan Gyongy, Neale AW Dutton, and Robert K Henderson. Single-photon tracking for high-speed vision. *MDPI Sensors*, 18:323, 2018. 2
- [31] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *NeurIPS*, 2020. 2
- [32] Atul Ingle, Andreas Velten, and Mohit Gupta. High flux passive imaging with single-photon sensors. In *CVPR*, pages 6753–6762, 2021. 2

- [33] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-Image translation with conditional adversarial networks. In *CVPR*, pages 5967–5976, 2017. 2
- [34] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-Image translation with conditional adversarial networks. In *CVPR*, pages 5967–5976, 2017. 3, 5
- [35] Kiyotaka Iwabuchi, Yusuke Kameda, and Takayuki Hamamoto. Image quality improvements based on motion-based deblurring for single-photon imaging. *IEEE Access*, 9:30080–30094, 2021. 2
- [36] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of GANs for improved quality, stability, and variation. In *ICLR*, 2018. 5
- [37] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for Generative Adversarial Networks. *IEEE TPAMI*, 43(12):4217–4228, 2021. 2
- [38] Manoj Kumar, Dirk Weissenborn, and Nal Kalchbrenner. Colorization transformer. In *ICLR*, 2021. 3
- [39] Vadim Lebedev, Yaroslav Ganin, Maksim Rakhuba, Ivan Oseledets, and Victor Lempitsky. Speeding-up convolutional neural networks using fine-tuned CP-decomposition. In *ICLR*, 2015. 2
- [40] Yu-Jing Lin, Po-Wei Wu, Che-Han Chang, Edward Chang, and Shih-Wei Liao. RelGAN: Multi-domain Image-to-Image translation via relative attributes. In *ICCV*, pages 5913–5921, 2019. 2
- [41] Wallace Lira, Johannes Merz, Daniel Ritchie, Daniel Cohen-Or, and Hao Zhang. GANHopper: Multi-hop GAN for unsupervised Image-to-Image translation. In *ECCV*, pages 363–379, 2020. 2
- [42] Yahui Liu, Enver Sangineto, Yajing Chen, Linchao Bao, Haoxian Zhang, Nicu Sebe, Bruno Lepri, Wei Wang, and Marco De Nadai. Smoothing the disentangled latent style space for unsupervised Image-to-Image translation. In *CVPR*, pages 10780–10789, 2021. 2
- [43] Kun Lu, Rongpeng Li, and Honggang Zhang. Contrastive monotonic pixel-level modulation. In *ECCV*, pages 493–510, 2022. 6
- [44] Ziwei Luo, Lei Yu, Xuan Mo, Youwei Li, Lanpeng Jia, Haoqiang Fan, Jian Sun, and Shuaicheng Liu. EBSR: Feature Enhanced Burst Super-Resolution with deformable alignment. In *CVPR*, pages 471–478, 2021. 5
- [45] Jiaju Ma and Eric Fossum. Quanta Image Sensor jot with sub 0.3e- r.m.s. read noise and photon counting capability. *IEEE Electron Device Letters*, 36(9):926–928, 2015. 1
- [46] Jiaju Ma and Eric R. Fossum. A Pump-Gate Jot device with high conversion gain for a Quanta Image Sensor. *IEEE Journal of the Electron Devices Society*, 3:73–77, 2015.
- [47] Jiaju Ma, Dakota Starkey, Arun Rao, Kofi Odame, and Eric R. Fossum. Characterization of Quanta Image Sensor pump-gate jots with deep sub-electron read noise. *IEEE Journal of the Electron Devices Society*, 3(6):472–480, 2015. 1
- [48] Jiaju Ma, Saleh Masoodian, Dakota A Starkey, and Eric R. Fossum. Photon-number-resolving megapixel image sensor at room temperature without avalanche gain. *Optica*, 4(12):1474–1481, 2017. 1
- [49] Jiaju Ma, Dexue Zhang, Omar Elgendy, and Saleh Masoodian. A 0.19e- rms read noise 16.7mpixel stacked Quanta Image Sensor with 1.1 μm -pitch backside illuminated pixels. *IEEE Electron Device Letters*, 42(6):891–894, 2021. 1
- [50] Jiaju Ma, Stanley H. Chan, and Eric R. Fossum. Review of Quanta Image Sensors for ultralow-light imaging. *IEEE Transactions on Electron Devices*, 69(6):2824–2839, 2022. 1, 2
- [51] Sizhuo Ma, Paul Mos, Edoardo Charbon, and Mohit Gupta. Burst vision using single-photon cameras. In *WACV*, pages 5364–5374, 2021. 2
- [52] Qi Mao, Hsin-Ying Lee, Hung-Yu Tseng, Jia-Bin Huang, Siwei Ma, and Ming-Hsuan Yang. Continuous and diverse Image-to-Image translation via signed attribute vectors. *IJCV*, 130(2):517–549, 2020. 6
- [53] Zichen Miao, Ze Wang, Xiuyuan Cheng, and Qiang Qiu. Spatiotemporal joint filter decomposition in 3D convolutional neural networks. In *NeurIPS*, 2021. 2
- [54] Zichen Miao, Ze Wang, Wei Chen, and Qiang Qiu. Continual learning with filter atom swapping. In *ICLR*, 2022. 2
- [55] Fabio Pizzati, Pietro Cerri, and Raoul de Charette. CoMoGAN: Continuous model-guided Image-to-Image translation. In *CVPR*, pages 14283–14293, 2021. 2
- [56] Qiang Qiu, Xiuyuan Cheng, Robert Calderbank, and Guillermo Sapiro. DCFNet: Deep neural network with decomposed convolutional filters. In *ICML*, 2018. 2, 4
- [57] Trevor Seets, Atul Ingle, Martin Laurenzis, and Andreas Velten. Motion adaptive deblurring with single-photon cameras. In *WACV*, pages 1944–1953, 2021. 2
- [58] Jheng-Wei Su, Hung-Kuo Chu, and Jia-Bin Huang. Instance-aware image colorization. In *CVPR*, pages 7965–7974, 2020. 3
- [59] Varun Sundar, Sizhuo Ma, Aswin Sankaranarayanan, and Mohit Gupta. Single-photon structured light. In *CVPR*, pages 17844–17854, 2022. 2
- [60] Patricia Vitoria, Lara Raad, and Coloma Ballester. ChromaGAN: Adversarial picture colorization with semantic class distribution. In *WACV*, pages 2434–2443, 2020. 3
- [61] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. High-resolution image synthesis and semantic manipulation with conditional GANs. In *CVPR*, pages 8798–8807, 2018. 2, 4
- [62] X. Wang, K. Yu, C. Dong, X. Tang, and C. Loy. Deep network interpolation for continuous imagery effect transition. In *CVPR*, pages 1692–1701, 2019. 6
- [63] Ze Wang. *Efficient Adaptation of Deep Vision Models*. PhD thesis, Purdue University, 2023. Available online: https://hammer.purdue.edu/articles/thesis/Efficient_Adaptation_of_Deep_Vision_Models/22696129. 2
- [64] Ze Wang, Xiuyuan Cheng, Guillermo Sapiro, and Qiang Qiu. Stochastic conditional generative networks with basis decomposition. In *ICLR*, 2020. 2, 4
- [65] Ze Wang, Seunghyun Hwang, Zichen Miao, and Qiang Qiu. Image generation using continuous filter atoms. In *NeurIPS*, 2021. 2

- [66] Yanze Wu, Xintao Wang, Yu Li, Honglun Zhang, Xun Zhao, and Ying Shan. Towards vivid and diverse image colorization with generative color prior. In *ICCV*, pages 14357–14366, 2021. [3](#)
- [67] Zhaoyang Yin, Jiaju Ma, Saleh Masoodian, and Eric R Fossum. Threshold uniformity improvement in 1b Quanta Image Sensor readout circuit. *MDPI Sensors*, 22(7):2578, 2022. [2](#)
- [68] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *ICCV*, pages 2242–2251, 2017. [2](#), [5](#)
- [69] Wei Zhu, Qiang Qiu, Robert Calderbank, Guillermo Sapiro, and Xiuyuan Cheng. Scaling-translation-equivariant networks with decomposed convolutional filters. *Journal of Machine Learning Research*, 23(68):1–45, 2022. [2](#)