Structured Kernel Estimation for Photon-Limited Deconvolution - Supplementary Document

Yash Sanghvi, Zhiyuan Mao, Stanley H. Chan
School of Electrical and Computer Engineering, Purdue University
{ysanghvi, mao114, stanchan}@purdue.edu

1. Overview of Camera Noise Model

In this section, we present an overview of different camera noise sources followed by a justification of the Poisson noise model used for photon-limited settings. Consider the sensor output of i^{th} pixel of camera, denoted as Y_i. From [10], we model Y_i as the following random variable:

$$Y_i \sim K_d\Big(K_a(\mathcal{P}(I_i) + \eta_a) + \eta_d + \eta_q\Big) \tag{1}$$

Here I_i represents the average number of incident photons during the exposure. K_a, K_d represent the analog and digital gain respectively. η_a represents noise sources before the analog gain (dark current shot noise, flicker noise etc.) and η_d represents noise sources before the digital gain (thermal noise, fixed pattern noise). η_q represent the quantization noise. From Eq. (1), we can see view Y_i is a noisy measurement of the parameter I_i.

Eq. (1) can be simplified in the following way:

$$\tilde{Y_i} \sim \underbrace{\mathcal{P}(I_i)}_{\text{signal-dependent}} \underbrace{+ \eta_a + \frac{1}{K_a}(\eta_d + \eta_q)}_{\text{signal-independent}} \tag{2}$$

where $\tilde{Y_i} \overset{\text{def}}{=} Y_i/(K_aK_d)$ is the normalized the pixel measurement. The noise sources can be easily decoupled into signal dependent Poisson noise and signal independent noise sources. The latter is often approximated as zero mean Gaussian noise in literature [2,3,6] leading to the following Poisson-Gaussian mixture modelling

$$\tilde{Y_i} \sim \mathcal{P}(I_i) + Z_i \qquad Z_i \sim \mathcal{N}(0, \sigma^2) \tag{3}$$

We can further break down average number of incident photon I_i as $I_i \overset{\text{def}}{=} \alpha x_i$ where α is a function of camera parameters such as exposure time, quantum efficiency, sensor area etc. and x_i is the scene radiance corresponding to the i^{th} pixel. This helps us decouple the scene and camera characteristics in the signal I_i. Therefore, we arrive at the following camera model formulation where Y_i is the measurement of the true signal x_i:

$$\tilde{Y_i} \sim \mathcal{P}(\alpha x_i) + \mathcal{N}(0, \sigma^2) \tag{4}$$

1.1. Poisson Noise SNR

For this subsection, we ignore the Gaussian term in Eq. (4) i.e. $\sigma = 0$ to focus on the nature of Poisson noise. An interesting property of the Poisson random variable is that its mean and variance are the same. Thus, the signal-to-noise ratio for $\tilde{Y_i}$, in absence of Gaussian noise, is given as follows:

$$\text{SNR}(\tilde{Y_i}) \overset{\text{def}}{=} \frac{\mathbb{E}[\tilde{Y_i}]}{\sqrt{\text{Var}[\tilde{Y_i}]}} = \sqrt{\alpha x_i} \tag{5}$$

This implies that our measurements get noisier with decreasing number of incident photons. If the scene is not well-illuminated (low x_i) or the exposure is short (small α), the number of incident photons on the sensor is also low leading to noisier images.

1.2. Photon-Limited Scenes

Scenes where the Poisson noise dominates other sources of noise in the measurements are defined as *photon-limited* [4]. For the random variable $\tilde{Y_i}$ defined in (4), this occurs when the variance due to Poisson noise is greater than the variance due to Gaussian noise, i.e. $\alpha x_i \geq \sigma^2$.

However, for the purpose of this paper, not all photon-limited scenes are equally significant. To emphasize this point, we inspect the SNR for Poisson-Gaussian mixture $\tilde{Y_i}$ for different levels of α. The SNR for $\tilde{Y_i}$ in presence of both Poisson and Gaussian noise is given as follows:

$$\text{SNR}[\tilde{Y_i}] \overset{\text{def}}{=} \frac{\mathbb{E}[\tilde{Y_i}]}{\sqrt{\text{Var}[\tilde{Y_i}]}} = \frac{\alpha x_i}{\sqrt{\alpha x_i + \sigma^2}} \tag{6}$$

Consider the case of read noise $\sigma = 1.6$e- [7]. We assume $x_i = 1$ and inspect the random variable Y_i for different α in the photon-limited regime, as shown in Table 1.

From the table we can conclude the following: Images taken in well-illuminated scenes with good exposure can be approximated as noiseless for the purpose of deblurring. However, on the other end of photon-limited regime

Photon level ↓	SNR (in dB)
$\alpha = 1000$	29.98 dB
$\alpha = 40$	15.75 dB
$\alpha = 20$	12.48 dB
$\alpha = \sigma^2$	1.07 dB

Table 1. Signal-to-noise ratio (SNR) for different photon levels in photon-limited regime.

i.e. $\alpha = \sigma^2$, there is too much noise in the image for any meaningful recovery from a single frame. Therefore, for this paper, we explore photon-limited deconvolution for $\alpha \in [10, 40]$ where shot noise dominates read noise but there is still a possibility of extracting the clean image.

2. Initialization Algorithm

In Algorithm 1, we describe the kernel initialization method for the proposed method. This scheme is a minor variation of the kernel estimation method from [1] and used to estimate a rectilinear kernel with parameters $\{\rho, \theta\}$ from the blur-only image $G(\mathbf{y})$. We would like to reiterate to the reader that the scheme in Algorithm 1 is not the kernel estimation process in its entirety and only represents the initialization process. The kernel estimated at the end of this algorithm is further refined in Stage I and II of the main iterative scheme.

Algorithm 1 Initialization for Kernel Estimation

```
Input: Blur-only Image $G(\mathbf{y})$, Number of Key Points $K$
Estimate gradient images $\mathbf{D}_x, \mathbf{D}_y$ from $G(\mathbf{y})$
for $\theta = 1, 2, \cdots, 180$ do
    $\mathbf{D}_\theta \leftarrow \mathbf{D}_x \cos(\theta) - \mathbf{D}_x \sin(\theta)$
    $f_\theta \leftarrow \max(|\mathbf{D}_\theta|)$
end for
$f_{\hat{\theta}}, \hat{\theta} \leftarrow \min f_\theta$
$\rho \leftarrow C_1 \sqrt{\frac{C_0^2}{f_{\hat{\theta}}^2} - \sigma_b^2}$
$(x_0, y_0) \leftarrow (0, 0)$
for $k = 1, 2, \cdots, K - 1$ do
    $(x_k, y_k) \leftarrow \frac{k\rho}{K-1} \cos(\hat{\theta}), \frac{k\rho}{K-1} \sin(\hat{\theta})$
end for
$\mathbf{z}_0 \leftarrow [x_0,\ y_0,\ x_1,\ y_1,\ \ldots,\ x_{K-1},\ y_{K-1}]^T$
return $\mathbf{z}_0$
```

3. Qualitative Comparison

Extended qualitative results comparing end-to-end trained methods to our approach are provided in the supplementary document. Figure 1 and 2 provide qualitative examples from the RealBlur dataset which contains realistic blur. Figure 3 contains examples from the PLDD dataset [9] cotnains real-shot noise corrupted and blurred image sensor data along with the ground-truth kernel, as measured using a point source. Finally, Figure 4 provides reconstruction results on synthetically blurred images using motion kernels from Levin dataset [5].

4. KTN Architecture and Training

The KTN architecture can be summarized as follows: the vectorized control points of dimension $2 \times (K - 1)$ are passed through 3 fully-connected layers followed by reshaping into an image. The reshaped image is then passed through the decoding half of a UNet to give the kernel output. The final output, when used in the iterative scheme, is clipped to zero and then normalized to one. Architecture details of KTN are provide in Figure 5.

5. Implementation Details

Boundary Conditions: While blurring the image synthetically, the boundary conditions are important to take into account. Circular boundary conditions allow the blur operator to be written in terms of FFTs and making it computationally inexpensive, it is not a realistic assumption for natural blur. A more appropriate boundary condition to assume is *symmetric boundary condition*.

This has major implications for our inverse problem scheme. Since PhD-Net assumes circular boundary conditions, we need to pad the image symmetrically, pass through PhD-Net and crop out the relevant portion to deblur the image without any artifacts. Second, when calculating the reblurring loss, $\mathbf{h_z} \circledast F(\mathbf{y}, \mathbf{h_z})$ needs to be calculated using symmetric boundary conditions.

Step Size and Backtracking: For Stage I, we set the initial step size as $\delta = 10^5$ and for Stage II, we set $\delta = 2.0$. For every iteration, we check whether the current choice of step-size decreases the cost-function or not. If it doesn't, then the step-size is reduced by half for rest of the iterative scheme until the next time the cost function increases instead of decreasing. Note that δ is set very large in the first stage compared to the second. This is because the gradients are backpropagated through two networks i.e., $F(.)$ and $T(.)$ instead of one, leading to the vanishing gradient problem and hence justifying the larger step size.

Computational Time: For each stage of the iterative scheme, we limit the number of iterations to 150. The experiments in the main document are performed on a Nvidia TitanX GPU, and take approximately 0.35 seconds per iteration.

References

[1] Mauricio Delbracio, Ignacio Garcia-Dorado, Sungjoon Choi, Damien Kelly, and Peyman Milanfar. Polyblur: Removing

mild blur by polynomial reblurring. *IEEE Transactions on Computational Imaging*, 7:837–848, 2021. 2

[2] Alessandro Foi. Clipped noisy images: Heteroskedastic modeling and practical denoising. *Signal Processing*, 89(12):2609–2629, 2009. 1

[3] Alessandro Foi, Mejdi Trimeche, Vladimir Katkovnik, and Karen Egiazarian. Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data. *IEEE Transactions on Image Processing*, 17(10):1737–1754, 2008. 1

[4] Samuel W Hasinoff. Photon, poisson noise. *Computer Vision, A Reference Guide*, 4:16, 2014. 1

[5] Anat Levin. Blind motion deblurring using image statistics. *Advances in Neural Information Processing Systems (NeurIPS)*, 19, 2006. 2

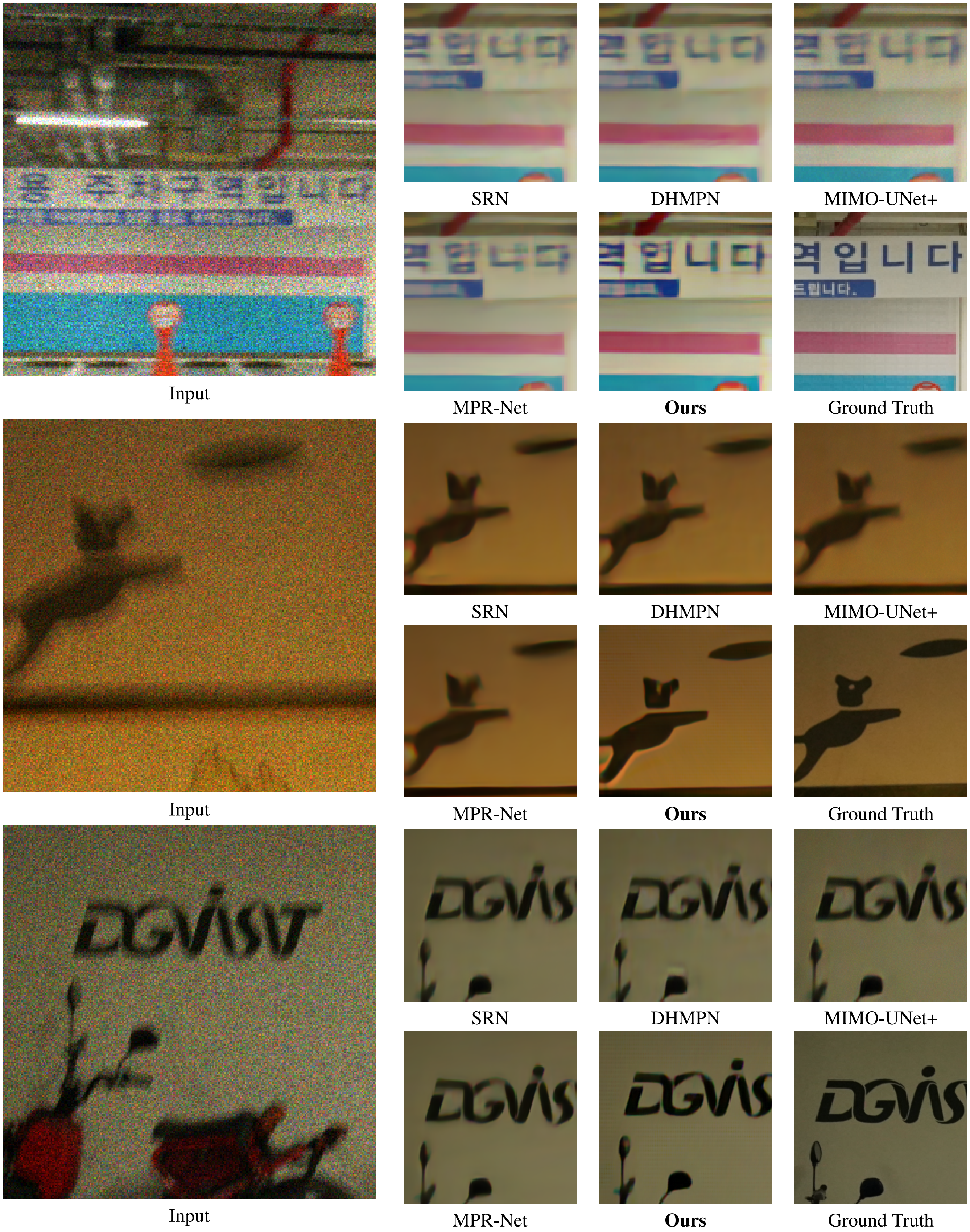
[6] Florian Luisier, Thierry Blu, and Michael Unser. Image denoising in mixed Poisson–Gaussian noise. *IEEE Transactions on Image Processing*, 20(3):696–708, 2010. 1

[7] Jiaju Ma, Dexue Zhang, Dakota Robledo, Leo Anzagira, and Saleh Masoodian. Ultra-high-resolution quanta image sensor with reliable photon-number-resolving and high dynamic range capabilities. *Scientific Reports*, 12(1):1–9, 2022. 1

[8] Yash Sanghvi, Abhiram Gnanasambandam, and Stanley H Chan. Photon limited non-blind deblurring using algorithm unrolling. *IEEE Transactions on Computational Imaging (TCI)*, 8:851–864, 2022. 6

[9] Yash Sanghvi, Abhiram Gnanasambandam, Zhiyuan Mao, and Stanley H. Chan. Photon-limited blind deconvolution using unsupervised iterative kernel estimation. *IEEE Transactions on Computational Imaging*, 8:1051–1062, 2022. 2

[10] Yi Zhang, Hongwei Qin, Xiaogang Wang, and Hongsheng Li. Rethinking noise synthesis and modeling in raw denoising. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4593–4601, 2021. 1



SRN DHMPN MIMO-UNet+

Input MPR-Net **Ours** Ground Truth

SRN DHMPN MIMO-UNet+

Input MPR-Net **Ours** Ground Truth

SRN DHMPN MIMO-UNet+

Input MPR-Net **Ours** Ground Truth

Figure 1. **Qualitative examples on the Real-Blur Dataset**

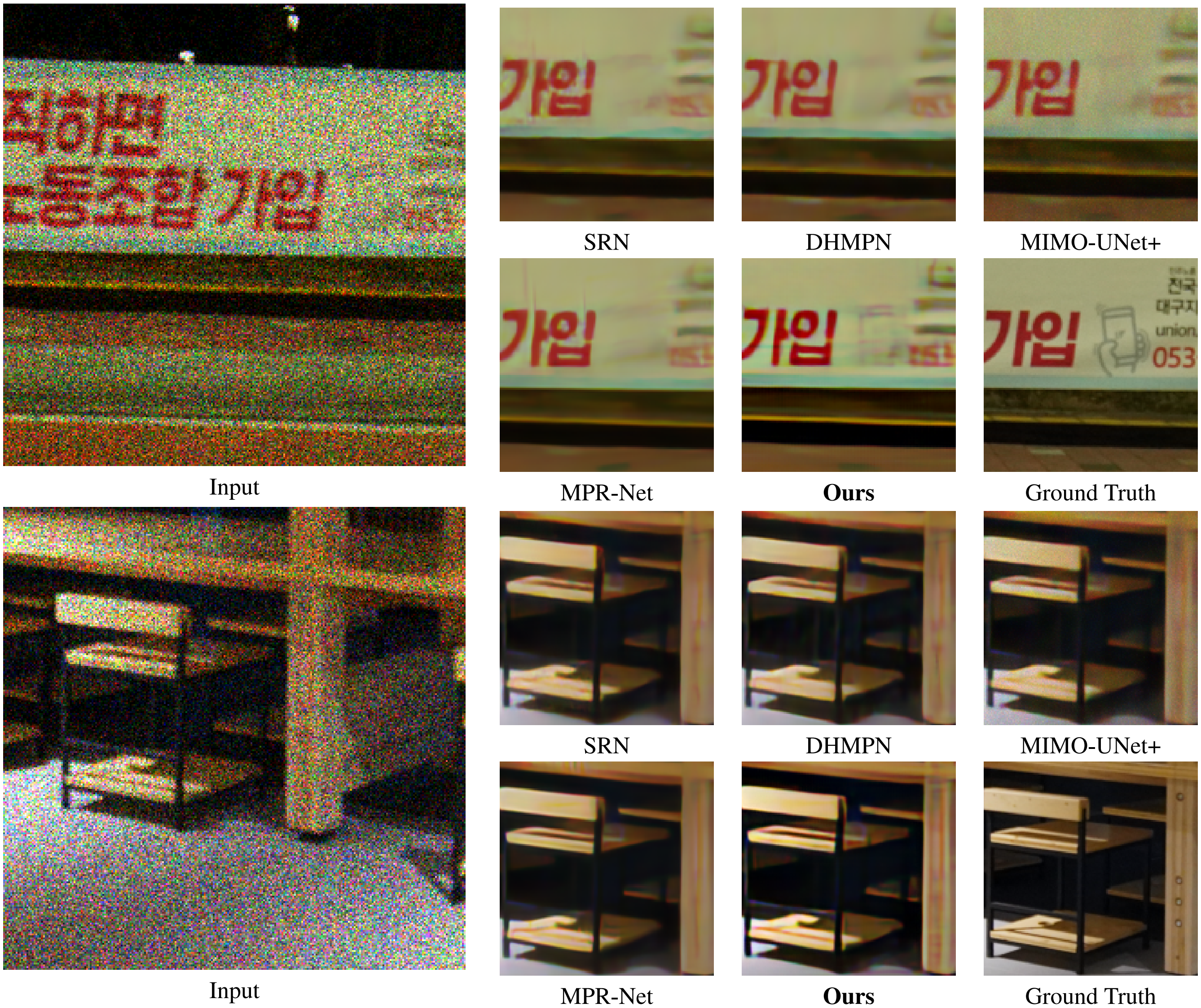


Figure 2. **More Qualitative examples on the Real-Blur Dataset**

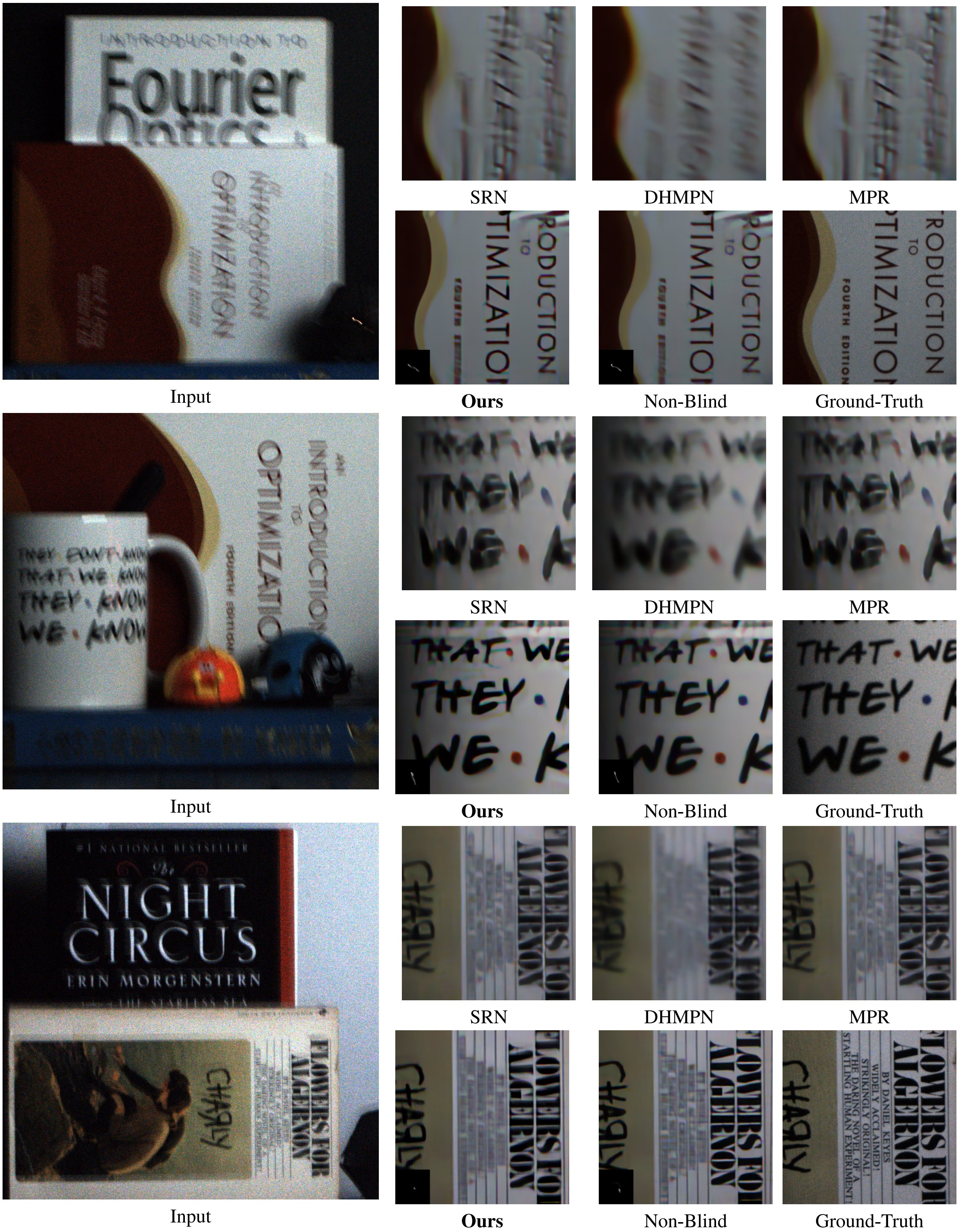


Figure 3. **More comparisons on Photon-Limited Deblurring Dataset** [8].

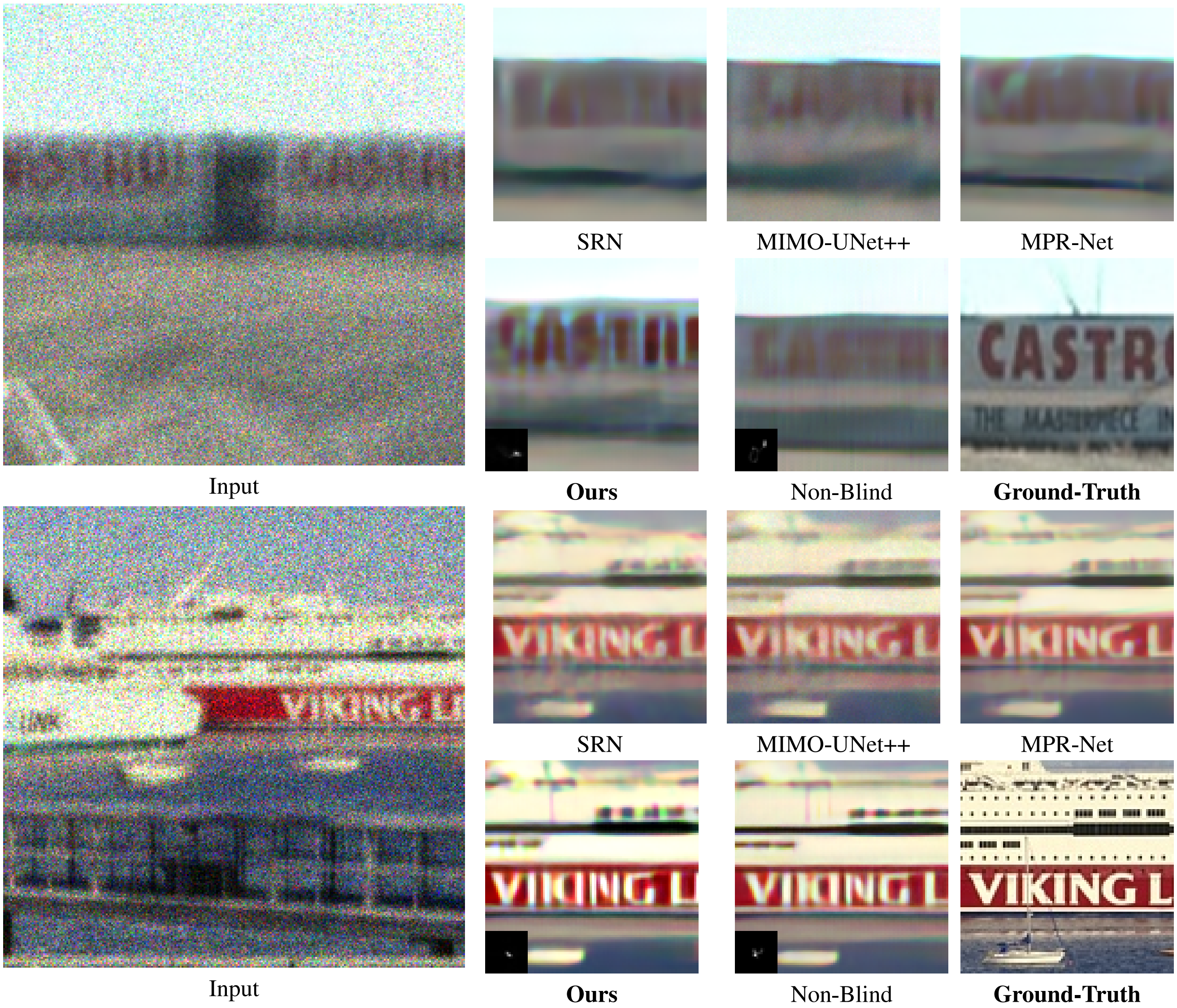


Figure 4. **Qualitative Examples on Synthetic Blur**: ”Non-Blind” is provided for reference and serves as an upper bound on the deconvolution performance. It is obtained through *PhD-Net* with noisy-blurred image and ground truth kernel as inputs. The kernels in inset of ”Ours” and ”Non-Blind” represent the estimated and true blur kernel respectively.

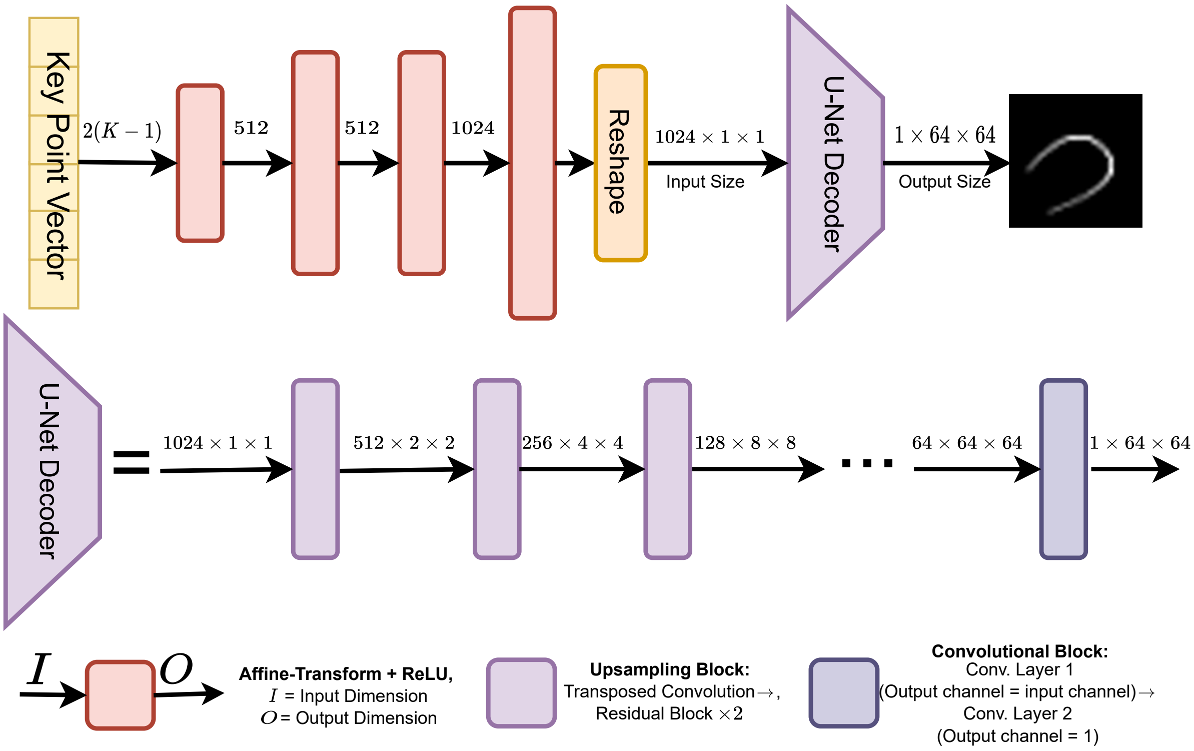


Figure 5. **Kernel Trajectory Network Architecture** 3 fully connected layers followed by a U-Net decoder