Physics-informed deep neural network for image denoising

Emmanouil Xypakis (emmxypakis@gmail.com)
Center for Life Nano- and Neuro-Science, Istituto Italiano di Tecnologia, Viale Regina Elena 291, 00161, Rome, Italy

Valeria de Turris
Center for Life Nano- and Neuro-Science, Istituto Italiano di Tecnologia, Viale Regina Elena 291, 00161, Rome, Italy

Fabrizio Gala
Crestoptics S.p.A. Italy

Giancarlo Ruocco
Center for Life Nano- and Neuro-Science, Istituto Italiano di Tecnologia, Viale Regina Elena 291, 00161, Rome, Italy

Marco Leonetti
Soft and Living Matter Laboratory, Institute of Nanotechnology, Consiglio Nazionale delle Ricerche, 00185, Rome, Italy

Research Article

Keywords: machine learning, microscopy, denoising, Poisson distribution

Posted Date: September 30th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2049014/v2

Physics-informed deep neural network for image denoising

Emmanouil Xypakis [1,2], Valeria de Turris[1], Fabrizio Gala[3], Giancarlo Ruocco[1], Marco Leonetti[1,2,4]

We developed a physics-informed deep neural network architecture able to achieve signal-to-noise ratio improvements starting from low-exposure noisy data. Our model is based on the nature of the photon detection process characterized by a Poisson probability distribution, an information which we included in the training loss function. Our approach surpasses previous algorithm performance for microscopy images; moreover, the generality of the physical concepts employed here, makes it readily exportable to any imaging context.

In any optical and non-optical imaging technology, measurement comes with a noise addition producing a signal that follows a Poisson probability distribution (PPD). Signal enhancement algorithms increase the amount of information by increasing the signal to noise ratio (SNR), making them useful for modeling and visualizing biological data including microscopy images, medical imaging, computer tomography, positron emission tomography and other in-vivo imaging technologies.

Deep neural network (DNNs)[1–7] based algorithms achieve the best signal enhancement results. The performance and the ability to train DNNs, however, depend both on the chosen loss function – a quantity comparing predictions and ground truth (GT) that DNN minimizes to learn its internal parameters – and on the normalization of the network inputs and targets[8]. Two commonly used loss functions for denoising and other image enhancement tasks are the L1-norm and the L2-norm (MSE) [2,4,9,10], where the data are arbitrarily normalized. On the other hand, when the desired output comes from a known probability distribution such as in semantic segmentation (U-net[11]) or other classification tasks, an entropic loss function, and a probabilistic normalization of data are of great significance. Although accounting for the physics of the camera detection process is known to significantly improve imaging efficiency [12], little research has been done in applying these physical properties in DNNs.

Physics-informed machine learning is a new trend in artificial intelligence [13]. Here, we report a physically informed DNN that builds on the PPD of signal detection. Our approach aims to provide a general and exportable approach to deal with Poisson distributed signals: I) we use a non-arbitrary and physics-based normalization process, II) we employ a physically informed loss function, and III) we design DNN architecture which takes advantage of the previous features. First, we remove any arbitrariness on the normalization just working with images in which each pixel count represents the photon number. Then, we design a loss function that considers the distance between probability distributions instead of the distance between count numbers. In particular we propose a symmetrized the Kullback-Leibler divergence (KL[14]; see also methods) which enables the algorithm to work with the same efficiency in all dynamic range windows. We employ a custom DNN architecture capable of classifying each pixel on a predicted photon number, thus preserving the photon number encoding and meaning for the output images. Our custom DNN employs structures from RCAN employed previously for denoising and U-net, which is employed for semantic segmentation.

We first illustrate how the PPD information can be encoded in the DNN architecture: in Fig. 1a we review the detection process of a PPD signal. The sample emits photons from volume V in all directions, and a fraction of them, N_{ph}, arrive at the detector. It is of great importance to characterize the signal as the photon numbers (see Methods and [12]), by calibrating the detector with an efficiency g plus an offset. The detection of photons follows a PPD of average μ which depends on the detector exposure τ as $\mu = \tau V \rho$, while the variance is $\sigma = \sqrt{\mu}$. Thus, the SNR for each voxel scales as $1/\sqrt{\mu}$ thus producing a different accuracy in the detection (see Fig. 1b) and the reconstruction of the density (see Fig. 1c). In Fig. 1d plots two different DNN loss functions that quantify the DNN output W with the GT mean value μ as a function of the PPD relative error $(\mu - W)/\sigma$: the KL divergence (see Methods), and the MSE (other state-of-the-art DNN). The difference between the two losses is that MSE loss strongly depends on the GT mean value while KL is only affected at large relative errors; the curves do not diverge for different μ. In this way, the KL loss gives the same relative loss at the full dynamic range and does not penalize the absolute difference between the GT and the output of the DNN. Here, we identify the classes of the entropic-like KL loss as the number of photons. In Fig. 1e, we show different DNN architectures. A state-of-the-art architecture that is used for classification is U-net, which uses a fine-tuned contracted and expansive path between convolutional blocks, while RCAN employs residual convolutional blocks to perform denoising. We employ RESUNET, which differs from U-net in the fact that instead of having convolutional blocks, it has residual blocks and uses the KL loss modified for photon classification. We train both the RCAN, and RESUNET on the same dataset (DS1) and evaluate them on a different dataset (DS2). DS1 and DS2 are both generated by experimental data for confocal fluorescence microscopy but differ in a) the camera; b) the fluorophores; c) the objective; and d) the biological samples (see Methods). In Fig. 1f, we show that RESUNET produces artifact-free reconstructions when exported to a different imaging condition.

In Fig. 2, we further quantify the performance of RESUNET and compare it with RCAN and the non-local-mean NLM classical denoising algorithm. In Fig. 2a we show the RESUNET, RCAN (trained in DS1 and evaluated in DS2) and

1. Center for Life Nano- and Neuro-Science, Istituto Italiano di Tecnologia, Viale Regina Elena 291, 00161, Rome, Italy
2. D-TAILS srl, 00161, Rome, Italy
3. Crestoptics S.p.A., Italy
4. Soft and Living Matter Laboratory, Institute of Nanotechnology, Consiglio Nazionale delle Ricerche, 00185, Rome, Italy

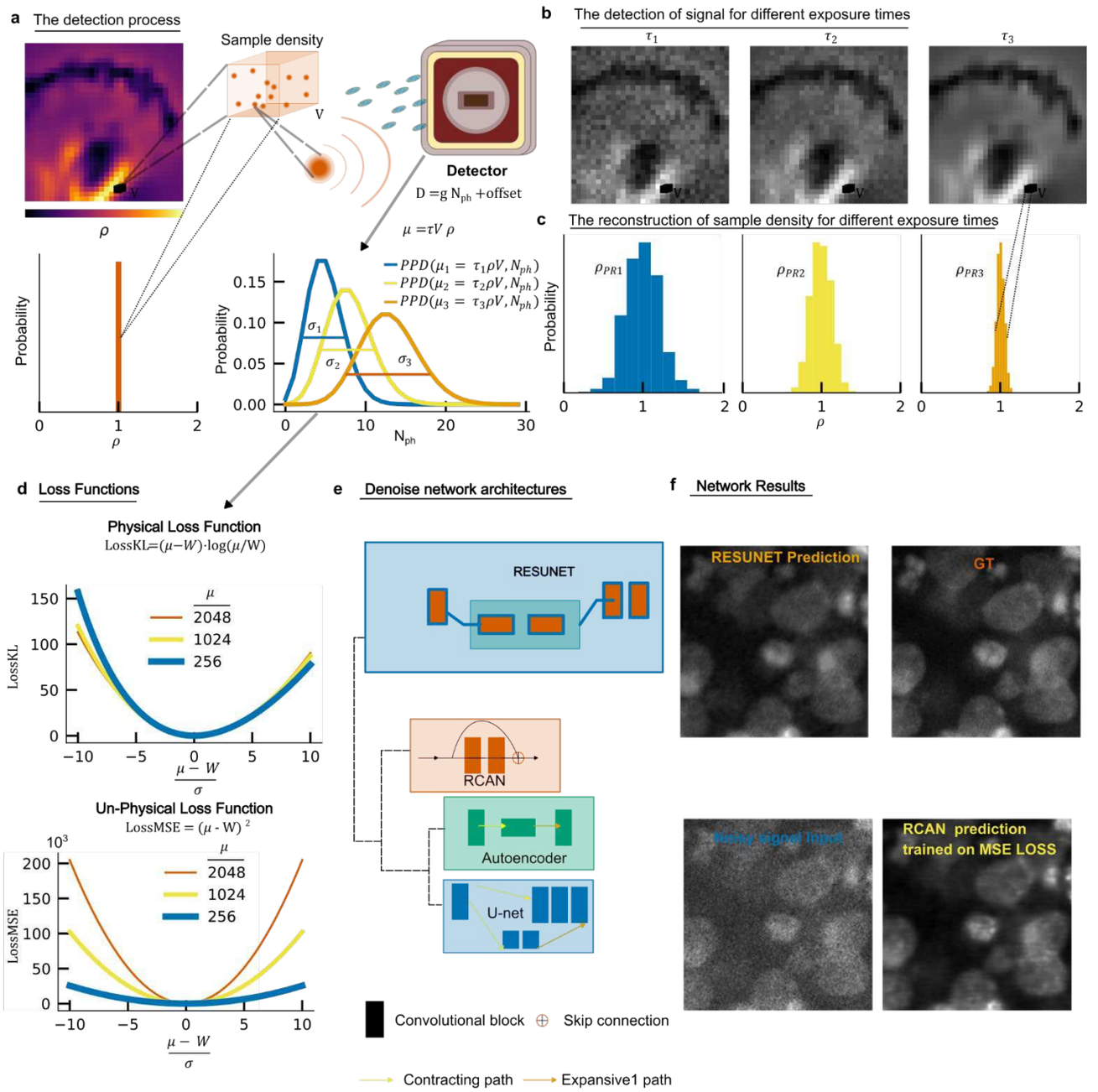

Figure1: a, from a small volume V of the density ρ of the sample N_{ph} photons are absorbed in the detector and transformed into electrical signal with efficiency g plus an offset. The longer the exposure τ the higher the average signal μ originating from the same density value ρ. On the left lower corner the real probability of the density value inside the volume V. On the right lower corner, the detector absorbs N_{ph} for different exposure times $\tau = \tau_1 < \tau_2 < \tau_3$ producing a PPD with average value μ and variance $\sigma = \sqrt{\mu}$. **b**, Three different measurements of the sample density for the exposure times of panel a. **c**, The reconstruction of the density inside V for the exposure times of panel a. **d**, The physical loss function LossKL and the unphysical LossMSE for a single voxel of the sample for three different GT average values μ as a function of the relative error $\frac{\mu-W}{\sigma}$ between the predicted DNN output W for three different average values μ. **e**, A schematic representation of different DNN architectures: RESUNET, RCAN, autoencoder, Unet. RESUNET is a combination of the other three and uses the physical Loss function F) Different predictions between the RCAN and the RESUNET (both trained in dataset DS1, and evaluated in dataset DS2) for a Noisy signal Input and a comparison with the GT.

NLM denoised images for the noisy input (10 msec exposure time image slice taken from dataset DS2) and the GT image (500 msec exposure time, same laser power). Both RESUNET and RCAN have been trained on the same dataset (DS1) for the same number of epochs. RESUNET produces a high-contrast, artifact-free smooth image. RCAN produces artifacts that originate from DS1 (horizontal lines in the lower part of the image) and a grainy image. NLM produces a high contrast image but tends to highlight the high intensity values more and loses some information on the structure of the image (legs on the right lower corner). In Fig. 2b, we show the line profiles of the line in Fig. 2a. RESUNET, in comparison with the other algorithms, produces a higher contrast between the deeps and the peaks along the lines. In Fig. 2c, we show the cumulative probability (Cum. Prob.) of the intensity histogram for the images shown in Fig. 2a. The vertical line represents the Kolmogorov distance. Both RCAN and RESUNET do not surpass GT Cum. Prob, with RESUNET approaching the GT at approximately 20 % of the total counts and RCAN at 50 %, while NLM surpasses the Cum. Prob. At 10 % of the total counts while approaching the GT Cum. Prob. at 50 %. In other words, approximately 40 % of the NLM denoised values are over-estimated, 50 % of the RCAN are under-estimated, while only 20 % of RESUNET are underestimated. In Fig. 2d, we compare different scores of the three algorithms for the same

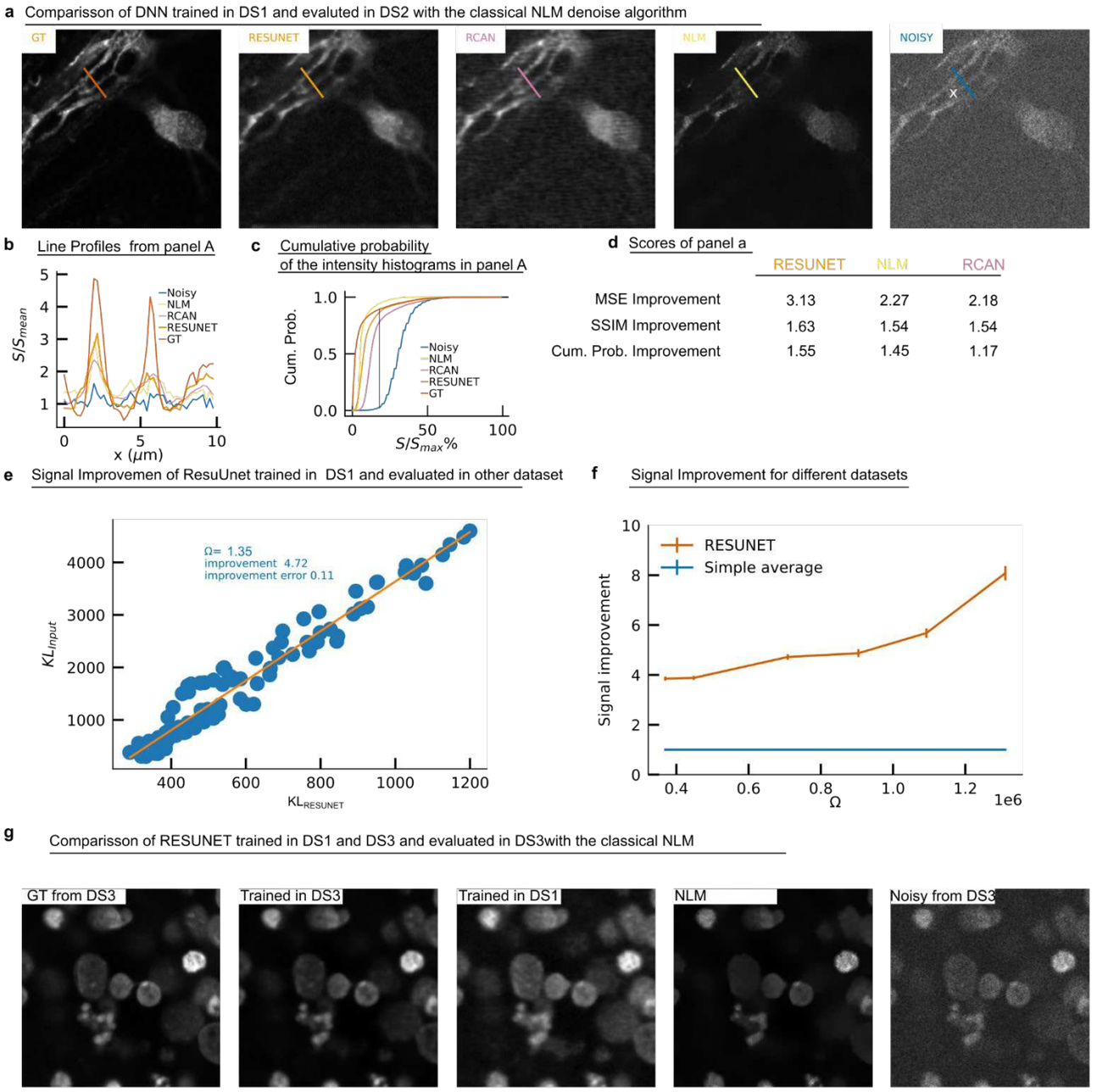

Figure2: **a**, The comparison of the RESUNET with RCAN and NLM algorithm reconstruction of a Noisy input 10msec exposure (Noisy) for a Noci cell image slice from DS2 with the corresponding GT (500msec exposure). Both RESUNET and RCAN have been trained in a different dataset 1. Characterized by different optical systems, fluorophores, and cameras. **b**, the line profile of the panel **a** for the line shown above **c**, The Cumulative probably of the intensity which is characterized by the Kolmogorov distance (vertical line) for panel a. **d**, The MSE, SSIM, Cumulative probability improvements for panel **a**. **e**, We define the signal Improvement as the slope between the KL_{Input} and the $KL_{RESUNET}$ scatter plot linear fit. which we find for each different dataset here shown just for one characterized by the parameter **f**, The Signal improvements for different datasets with different $\Omega = \frac{V_{input}}{V_{psf}}$. **g**, Comparison of RESUNET for the same DS trained and different and comparison with the NLM. The signal improvement is 54.2 (trained in DS3) and 4.4 (trained in DS2)

image, showing that RESUNET achieves better scores. In Fig. 2e, we quantify the signal improvement of RESUNET (trained in DS1) for a different dataset by performing a linear fit on the scatter plot of the KL divergence between the GT and Noisy input and the KL between the GT and the output of RESUNET. The slope of the linear fit is the signal improvement. The parameter Ω (see Methods) identifies the information contained in the image volume. In Fig. 2f, we plot the signal improvement for different datasets with different Ω, showing that signal improvement increases as Ω increases, while a simple average of the input image brings no advantages in the signal improvement. Finally, in Fig. 2g, we show that when trained on the same dataset, under the same imaging conditions, the performance of RESUNET increases even further.

In summary, we showed that a physics-informed loss function employed in a novel DNN architecture outperforms the state-of-the-art denoising algorithms for fluorescence microscopy. We overcome the arbitrariness of the data normalization in denoising DNNs by using a photon model that does not produce instabilities, thus showing portability when used in different imaging conditions. Our loss function is based on the statistics of the Poisson distribution and could be used in DNNs performing tasks other than microscopy, such as optical or non-optical imaging technologies that have Poissonian signals.

Acknowledgment: This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 713694. GR acknowledges the support by European

Research Council Synergy grant ASTRA (No. 855923). ML thanks Project LOCALSCENT, Grant PROT. A0375-2020-36549, Call POR-FESR "Gruppi di Ricerca 2020".

Competing interests: The authors declare no competing interests.

Bibliography

1. Belthangady, C. & Royer, L. A. learning for fluorescence image reconstruction. *Nat. Methods* (2009) doi:10.1038/s41592-019-0458-z.

2. Weigert, M. *et al.* Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090–1097 (2018).

3. Laine, R. F., Arganda-Carreras, I., Henriques, R. & Jacquemet, G. Avoiding a replication crisis in deep-learning-based bioimage analysis. *Nat. Methods* **18**, 1136–1144 (2021).

4. Chen, J. *et al.* Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes. *Nat. Methods* **18**, (2021).

5. Gurrola-ramos, J., Dalmau, O. & Alarcón, T. E. A Residual Dense U-Net Neural Network for Image Denoising. **9**, (2021).

6. Byun, J., Cha, S. & Moon, T. FBI-Denoiser: Fast Blind Image Denoiser for Poisson-Gaussian Noise. (2021).

7. Mayorov, A. S. *et al.* Interaction-Driven Spectrum Reconstruction in Bilayer Graphene. *Science (80-.).* **333**, 860–863 (2011).

8. Shanker, M., Hu, M. Y. & Hung, M. S. Effect of data standardization on neural network training. *Omega* **24**, 385–397 (1996).

9. Qiao, C. *et al.* Evaluation and development of deep neural networks for image super-resolution in optical microscopy. *Nat. Methods* (2021) doi:10.1038/s41592-020-01048-5.

10. Xypakis, E. *et al.* Deep learning for blind structured illumination microscopy. *Sci. Rep.* **12**, 8623 (2022).

11. Falk, T. *et al.* U-Net: deep learning for cell counting, detection, and morphometry. *Nat. Methods* **16**, 67–70 (2019).

12. Mandracchia, B. *et al.* Fast and accurate sCMOS noise correction for fluorescence microscopy. *Nat. Commun.* 1–12 doi:s41467-019-13841-8.

13. Karniadakis, G. E. *et al.* Physics-informed machine learning. *Nat. Rev. Phys.* **3**, 422–440 (2021).

14. Kullback, S. & Leibler, R. A. On Information and Sufficiency. *Ann. Math. Stat.* **22**, 79–86 (1951).

15. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).

16. Demidenko, E. Kolmogorov-smirnov test for image comparison. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **3046 LNCS**, 933–939 (2004).

Methods

Normalization and camera calibration.

One known issue in the training process is the normalization of the data. Improper or arbitrary normalization leads to a failure 4. in the training[8]. To overcome this issue of denoising DNNs for microscopy, an (arbitrary) percentile normalization of the images has been proposed; however, on the evaluation of the algorithms, the efficiency scores that are reported[2,4] not only depend on the normalization choice, but an extra normalization to the ground truth image (GT) has to be performed to produce the best scores. This ambiguity in the results becomes important, especially when a GT is not provided or unknown. Overall, a robust image normalization scheme is missing from the literature. To overcome these issues, we first calibrate the camera as described in [12] and use the photon numbers as the input and target of RESUNET so that each pixel/voxel represents the number of photons. Both the GT and the output of RESUNET represent the average value of the PPD.

Physical loss function.

The Kullback-Leibler divergence is a measure of comparing two probability distribution functions $p(n), q(n)$, for the same sample space n. It is defined as

$$KL(p,q) = \sum_n p(n) * \log\left(\frac{p(n)}{q(n)}\right)$$

For two PPDs with mean values μ, W, $p(n,\mu) = \mu^n\, e^{-\mu}/n!$ $q(n,W) = W^n\, e^{-W}/n!$, it can be shown that $KL(p,q)$ depends only on the mean values μ, W

$$KL(\mu,W) = \mu \log\left(\frac{W}{\mu}\right) - (W-\mu)$$

In this paper we used the symmetrized version of $KL(\mu,W)$ as the loss function.

$$Loss_{KL} = \frac{1}{2}(\,KL(\mu,W) + KL(W,\mu)) = (\mu - W)\log\left(\frac{\mu}{W}\right)$$

Since the signal comes from a probabilistic distribution, we treat the photon number for each pixel of the input and the target of RESUNET as the mean values of the PPD.

Res-U-net

Regarding the architectures, a promising approach is to use skip connections between network layers to bypass residual contents (RCAN[4]) and an encoder-decoder-like architecture efficiently solves semantic segmentation/classification tasks (U-net[11]). In our case, the classes are the number of photons, so we use the efficiency of U-net to classify photons and the power of residues to perform denoising. To do that, we substitute the convolutional layers of U-net with residual blocks.

Evaluation datasets and reproducibility.

To ensure reproducibility[3], we evaluated and compared our architecture with the state of the art DNN architecture and a non DNN classical denoising algorithm. Dataset DS1 is the dataset that RCAN was originally trained for Tubulins and can be found in[4]. The evaluation was performed on confocal spinning disk microscopy experimental data acquired in our lab where we systematically changed the exposure time (1,2,5,10,20,50,100 and 500 msecs) and thus the SNR. We differentiate each dataset by the parameter $\Omega = \frac{V_{input}}{V_{psf}}$ as the sampling rate between the volume of the input and the psf by using three different fluorophores (DAPI, Alexa488, Alexa647) emission wavelengths (420 nm, 525 nm, 665 nm). and three different objectives with different depths of field (UPLSAPO 30XS, UPLSAPO 20X, LUCPLFLN40XPH 40X).

Evaluation metrics.

To provide the best comparison in addition to the widely used structural similarity index (SSIM [15]) and mean square error (MSE) we also use the Kolmogorov-Smirnov index [16] and the Kullback-Leibler divergence. In this way we provide the maximum information for image comparison.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Xmodel.png