LEVERAGING FOURIER DOMAIN FOR UNSUPERVISED REMOVAL OF STRUCTURED NOISE IN FLUORESCENCE MICROSCOPY IMAGING

Haosen Liu *Edmund Y. Lam**

Department of Electrical and Electronic Engineering,
The University of Hong Kong, Pokfulam, Hong Kong, China

2026 IEEE 23rd International Symposium on Biomedical Imaging (ISBI) | 979-8-3315-7763-6/26/$31.00 ©2026 IEEE | DOI: 10.1109/ISBI61048.2026.11515690

ABSTRACT

Due to the practical limitations such as phototoxicity, fluorescence microscopy images are generally with faint signals corrupted by severe and complex noise, necessitating the denoising step to improve image quality for more accurate biological analysis. Recently, unsupervised deep learning techniques have attracted widespread attentions owing to their powerful capability of removing noise without the reliance of collecting clean images. Although existing unsupervised denoising methods such as blind spot networks are effective in suppressing spatially independent noise, developing unsupervised strategies for structured noise is still underexplored. Based on the observation that noise correlated in the spatial domain can exhibit independence in the Fourier domain, this work develops a Fourier-based blind spot strategy to enable effective unsupervised structured noise removal for multi-frame fluorescence microscopy imaging. The evaluation on publicly available real-world microscopy datasets validates its effectiveness in removing structured noise.

***Index Terms*—** Unsupervised denoising, Deep learning, Structured noise, Fluorescence microscopy

1. INTRODUCTION

As a powerful tool for visualizing cellular structures and their dynamic activities, fluorescence microscopy imaging is essential for biological research. However, its signal intensity is severely constrained due to the existence of the limited photon budget imposed by the phototoxicity and the requirements on the high spatiotemporal resolution. Simultaneously, the randomness inherent in the collection, conversion, and processing of photoelectric signals can result in complicated random noise. As such, the fluorescence microscopy image is always with extremely low signal-to-noise ratio (SNR), impeding the effectiveness and accuracy of subsequent biological analysis. This necessitates the development of effective denoising algorithms to improve the image quality so that downstream processing tasks can be benefited.

In the past decades, numerous efforts have been made to tackle the image denoising task. Conventional methods generally focus on explicitly modeling image properties [1] and noise statistics [2] with hand-crafted models. These methods can be powerful when their models are accurate, while inferior for challenging scenarios such as various microscopy imaging, where exact models are almost unavailable. To avoid the reliance of explicit model, deep learning techniques, which can implicitly learn image distributions from the training data, have been applied to the image denoising tasks [3]. Owing to their strong capacity, these methods have exhibited significant performance advantages over conventional hand-crafted techniques. However, such an advantage is not without cost. Collecting tremendous high-SNR and low-SNR image pairs has replaced the explicit modeling, becoming a new burden for image denoising tasks. This is especially challenging for fluorescence microscopy imaging, since many biological samples are vulnerable to high light intensities and their dynamic behaviors can lead to blurry effects if the exposure time is too long.

To release the data collection burden and exploit the potential of deep neural networks, unsupervised denoising techniques that can learn an effective denoising mapping directly from noisy images have attracted widespread attention. The key to enabling unsupervised denoising is the division of every single noisy image into paired data that have independent noise components. For instance, the blind spot strategies [4] treat the center pixel to be predicted as the training target and its neighboring pixels as the paired input. Some other methods sub-sample the original images to multiple sub-images acting as paired training data [5].

Following similar principles, these single image denoising techniques can be extended for multi-frame image denoising. For example, the DeepInterpolation [6] trains a frame interpolation network to recover an unseen noisy frame from its neighboring frames. However, the loss of center frame can have a substantial negative impact on the image recovery results. To overcome this challenge, some unsupervised video denoising methods estimate the optical flow and warp an adjacent frame to pad the missing frame [7]. Some others propose to restrict the blind spot region to pixels at the same spatial position [8] or even only a single pixel [9], instead of making the entire center frame blind to the network. However, these methods can be inferior or even invalid when noise is cor-

*Corresponding author: elam@eee.hku.hk

related, limiting their application to the microscopy imaging where structured noise is a common case.

Inspired by the observation that Fourier coefficients of spatially correlated noise could be independent with each other [10], we develop an unsupervised multi-frame image denoising framework to tackle the structured noise by applying the blind spot strategy to the Fourier domain. More specifically, we train a denoising network by randomly masking parts of Fourier coefficients of the center frame and predicting them with the remaining coefficients. This method can effectively remove noise that is independent in the frequency domain, even if it exhibits long-range spatial correlations. Furthermore, we propose a scheme to combine the blind spots in both the Fourier domain and spatial domain, enabling our method to handle a wider range of noise types and enhancing its potential for applications in fluorescence microscopy imaging.

2. METHOD

This section will first provide a brief analysis of noise in fluorescence microscopy images. Then, we will introduce our developed Fourier-based blind spot strategy and its combination with the blind spot network to suppress both the pixel-wise noise and structured noise.

2.1. Noise Analysis

Noise in practical microscopy images is generally a mixture of different noise sources. This work mainly considers two noise types. The first type is the additive stationary noise n that can be modeled by

$$n = \sum_{i=1}^{I} k_i * \eta_i, \tag{1}$$

where $*$ denotes the convolution operator, k_i is the noise correlation kernel, and η_i denotes $i.i.d.$ noise with zero mean. As illustrated in [11], this noise model can accurately represent structured noise in various microscopes. By setting a reasonable correlation kernel k_i and the distribution of η_i, this model can accurately depict the mixture of $i.i.d.$ pixel-wise noise and spatially correlated structured noise. The second type is the noise holding correlations in the Fourier domain. Some examples include the spatially independent multiplicative noise and the salt-and-pepper noise. The coexistence of these two types of noise challenges the development of unsupervised denoising algorithm for microscopy imaging. To overcome this challenge, we construct a Fourier-based blind spot strategy to remove the additive stationary structured noise. By combining it with the existing spatial-domain blind spot method, we develop an unsupervised denoising method effective for suppressing mixed noise in the microscopy images.

2.2. Fourier-based blind spot strategy

For the additive stationary noise depicted in (1), an important observation reported in [10] is that the Fourier coefficients of such noise are independent with each other. This property ensures that the Fourier coefficients of noisy images can be easily divided into two parts, of which the noise components are independent. As such, when a network is learned to predict one part of coefficients with the remaining coefficients, the best prediction of the noise component contained in this part is its mean value, which is precisely the denoising process. Utilizing this effect, we develop a Fourier-based multi-frame unsupervised denoising technique.

Given multiple noisy frames $\mathbf{x} = [x_{-c} \ldots, x_c]$, our goal is to remove noise n_0 in the center frame x_0 to predict the latent clean image y_0. We assume that noise $\mathbf{n} = [n_{-c} \ldots, n_c]$ contained in these frames are with zero mean and there is no inter-frame noise correlations, *i.e.*, there are $\mathbb{E}(n_i) = 0$ for each i and $\mathbb{E}(n_i n_j) = 0$ for $i \neq j$. To construct the Fourier-based blind spot region, we mask the Fourier coefficients of the center frame as

$$\tilde{x}_0 = \mathscr{F}^{-1}\left(\mathscr{F}\left(x_0\right) \cdot (1-m) + \mathscr{F}\left(x_{\text{adj}}\right) \cdot m\right), \tag{2}$$

where $\mathscr{F}\left(\cdot\right)$ and $\mathscr{F}^{-1}\left(\cdot\right)$ denote the Fourier transform and the inverse Fourier transform, m is a binary mask following the Bernoulli distribution with the probability of p, x_{adj} denotes a randomly selected adjacent frame of x_0. Replacing the original center frame x_0 with $\tilde{x}_0$ to obtain $\tilde{\mathbf{x}}$, we can train a denoising network with the loss function as

$$L\left(f_\theta(\tilde{\mathbf{x}}), x_0\right) = \left\|\left(\mathscr{F}\left(f_\theta(\tilde{\mathbf{x}})\right) - \mathscr{F}\left(x_0\right)\right) \cdot m\right\|_2^2, \tag{3}$$

where f_θ denotes the network parameterized with θ. When noise n_0 is independent in the Fourier domain, we can have

$$\mathbb{E}(\mathscr{F}\left(f_\theta(\tilde{\mathbf{x}})\right)\left(\mathscr{F}\left(x_0\right) \cdot m\right)) = \mathscr{F}\left(f_\theta(\tilde{\mathbf{x}})\right)\left(\mathscr{F}\left(y_0\right) \cdot p\right), \tag{4}$$

which further leads to

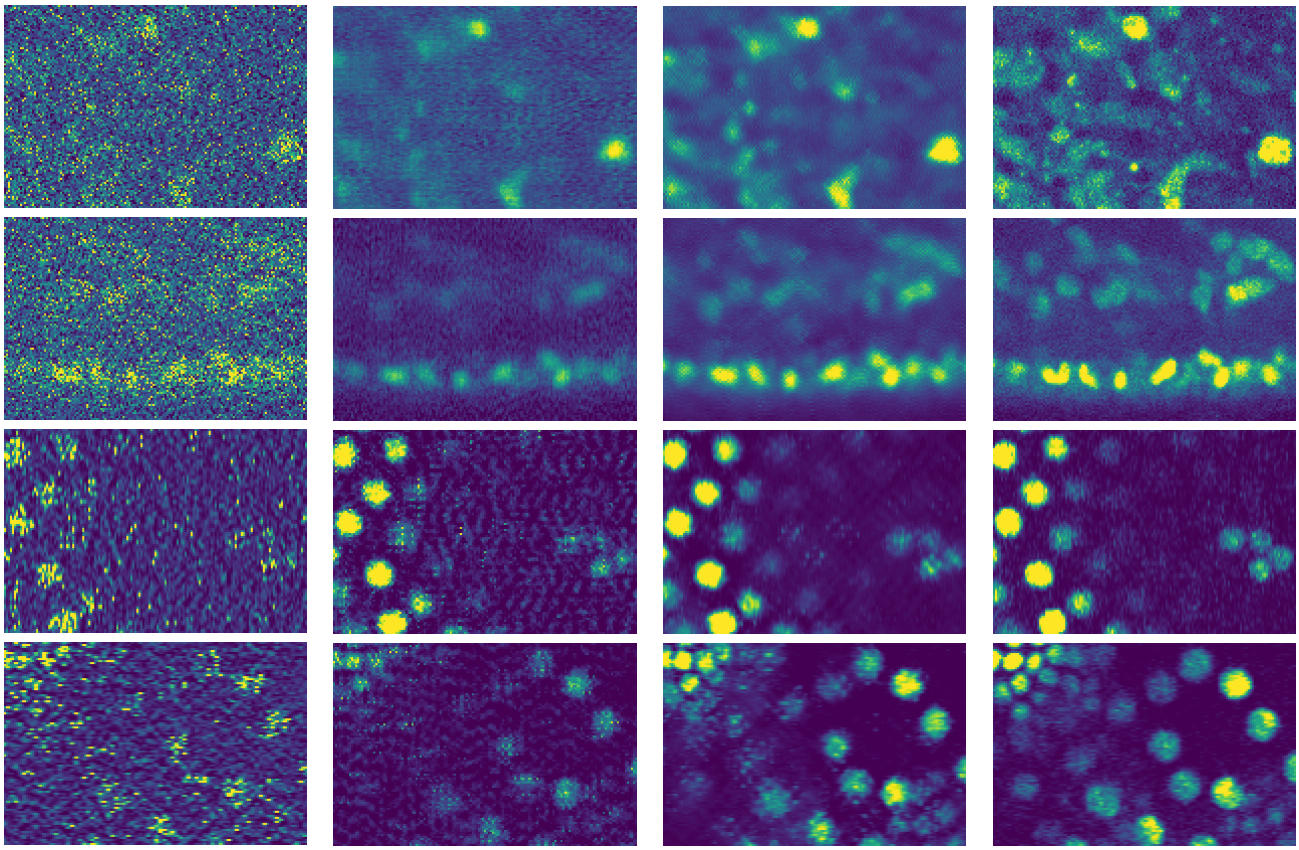
$$\begin{aligned} &\mathbb{E}\Big(\left\|\left(\mathscr{F}\left(f_\theta(\tilde{\mathbf{x}})\right) - \mathscr{F}\left(x_0\right)\right) \cdot m\right\|_2^2\Big) \\ =&p\left(\mathscr{F}\left(f_\theta(\tilde{\mathbf{x}})\right) - 2\mathscr{F}\left(f_\theta(\tilde{\mathbf{x}})\right)\mathscr{F}\left(y_0\right)\right) + const, \\ =&p\left\|\mathscr{F}\left(f_\theta(\tilde{\mathbf{x}})\right) - \mathscr{F}\left(y_0\right)\right\|_2^2 + const. \end{aligned} \tag{5}$$

This equation proves that the learning process utilizing the unsupervised loss function in (3) is statistically equivalent to that of supervised learning.

2.3. Combination with spatial-domain blind spot method

In addition to leveraging Fourier-based blind spot strategy, to further enhance the applicability of our method to various types of noise, we also incorporate the spatial-domain blind spot technique. More specifically, we incorporate the blind spot network f_θ^{BSN} into the loss function in (3) as

$$L\left(f_\theta^{\text{BSN}}(\tilde{\mathbf{x}}), x_0\right) = \left\|\left(\mathscr{F}\left(f_\theta^{\text{BSN}}(\tilde{\mathbf{x}})\right) - \mathscr{F}\left(x_0\right)\right) \cdot m\right\|_2^2 \tag{6}$$



(a) Noisy (b) AT-BSN (c) FourierBSN (d) GT

Fig. 1. Visual evaluation of denoising results on real-world microscopy imaging datasets. From top to bottom, the images are sourced from Planaria C1, Planaria C2, Tribolium C1, and Tribolium C2.

Test set	AT-BSN		FourierBSN	
	PSNR	SSIM	PSNR	SSIM
Planaria C1	23.74	0.3593	27.86	0.5733
Planaria C2	21.49	0.3258	25.22	0.5008
Tribolium C1	27.53	0.5574	30.31	0.8391
Tribolium C2	23.39	0.3109	27.83	0.7659

Table 1. Average PSNR (dB) and SSIM on real-world microscopy imaging datasets.

to collaboratively utilize the blind spot strategies in both the Fourier domain and spatial domain. This combination, which we call FourierBSN, can effectively process practical noise encountered in microscopy imaging.

3. EXPERIMENTS

Our method are validated on two real-world fluorescence microscopy datasets, *i.e.*, Tribolium and Planaria, released by CARE [12]. These datasets have images with different SNR for the same scene. We use the one with the highest SNR as the ground truth reference, and others with two different noise levels, dubbed C1 and C2, as the noisy image to be processed. Our training data $\mathbf{x}$, each of which contains 11 (*i.e.*, $c = 5$) consecutive frames, is extracted from the test sets of Planaria and Tribolium, which contain 20 and 6 image stacks with tens of frames, respectively. The trained network evaluated is conducted on the same data, which is reasonable for the unsupervised technique. The blind spot network we employed is the AT-BSN [13], which also serves as a comparison method in the evaluation. It has an adjustable blind regions size, which we set 1 for Planaria and 5 for Tribolium. The Adam optimizer with a learning rate as 0.0001 is adopted and we set the training step to 10000. To provide quantitative evaluation, the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) of the denoising results are reported in Table 1. Besides, representative results are also provided in Figure 1 to support visual evaluation.

From Figure 1, it can be observed that noise corrupting real-world fluorescence microscopy images is complicated, consisting of both correlated noise and isolated pixels hold-

ing high intensity. For the AT-BSN, it is inferior in eliminating correlated noise, leaving obvious structured artifacts in the denoised results. In contrast, our method can effectively remove mixed noise contained in the real-world fluorescence microscopy images, significantly improving their image qualities. As Table 1 shows, our developed FourierBSN outperforms AT-BSN by a large margin in terms of PSNR and SSIM, which is consistent with the visual results.

4. CONCLUSION

In this study, we have developed a Fourier-based unsupervised multi-frame image denoising technique. By constructing the unsupervised loss function via the blind spots in the Fourier domain, this method excels in removing structured noise. Additionally, the combination of spatial-domain blind spot strategy (*i.e.*, the blind spot network AT-BSN) makes our method able to effectively remove complicated noise encountered in real-world fluorescence microscopy images.

5. COMPLIANCE WITH ETHICAL STANDARDS

This research was conducted retrospectively using fluorescence microscopy data made available in open access. Ethical approval was not required as confirmed by the license attached with the open-access data

6. ACKNOWLEDGMENTS

This work is supported in part by the Research Grants Council of Hong Kong (GRF 17201822) and the Theme-based Research Scheme (T45-701/22-R).

7. REFERENCES

[1] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[2] Ymir Mäkinen, Lucio Azzari, and Alessandro Foi, "Collaborative filtering of correlated noise: Exact transform-domain variance for improved shrinkage and patch matching," *IEEE Transactions on Image Processing*, vol. 29, pp. 8339–8354, 2020.

[3] Xiaojiao Mao, Chunhua Shen, and Yu-Bin Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," *Advances in Neural Information Processing Systems*, vol. 29, 2016.

[4] Alexander Krull, Tim-Oliver Buchholz, and Florian Jug, "Noise2Void-learning denoising from single noisy images," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2129–2137.

[5] Tao Huang, Songjiang Li, Xu Jia, Huchuan Lu, and Jianzhuang Liu, "Neighbor2neighbor: Self-supervised denoising from single noisy images," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 14781–14790.

[6] Jérôme Lecoq, Michael Oliver, Joshua H Siegle, Natalia Orlova, Peter Ledochowitsch, and Christof Koch, "Removing independent noise in systems neuroscience data using DeepInterpolation," *Nature methods*, vol. 18, no. 11, pp. 1401–1408, 2021.

[7] Thibaud Ehret, Axel Davy, Jean-Michel Morel, Gabriele Facciolo, and Pablo Arias, "Model-blind video denoising via frame-to-frame training," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 11369–11378.

[8] Dev Yashpal Sheth, Sreyas Mohan, Joshua L Vincent, Ramon Manzorro, Peter A Crozier, Mitesh M Khapra, Eero P Simoncelli, and Carlos Fernandez-Granda, "Unsupervised deep video denoising," in *IEEE/CVF International Conference on Computer Vision*, 2021, pp. 1759–1768.

[9] Zichun Wang, Yulun Zhang, Debing Zhang, and Ying Fu, "Recurrent self-supervised video denoising with denser receptive field," in *ACM International Conference on Multimedia*, 2023, pp. 7363–7372.

[10] Haosen Liu, Jiahao Liu, Shan Tan, and Edmund Y Lam, "Image restoration learning via noisy supervision in the fourier domain," *arXiv preprint arXiv:2506.00564*, 2025.

[11] Jérôme Fehrenbach, Pierre Weiss, and Corinne Lorenzo, "Variational algorithms to remove stationary noise: applications to microscopy imaging," *IEEE Transactions on Image Processing*, vol. 21, no. 10, pp. 4420–4430, 2012.

[12] Martin Weigert, Uwe Schmidt, Tobias Boothe, Andreas Müller, Alexandr Dibrov, Akanksha Jain, Benjamin Wilhelm, Deborah Schmidt, Coleman Broaddus, Siân Culley, et al., "Content-aware image restoration: pushing the limits of fluorescence microscopy," *Nature methods*, vol. 15, no. 12, pp. 1090–1097, 2018.

[13] Shiyan Chen, Jiyuan Zhang, Zhaofei Yu, and Tiejun Huang, "Exploring efficient asymmetric blind-spots for self-supervised denoising in real-world scenarios," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2024, pp. 2814–2823.