Fusion-Based Enhancement of Multi-Exposure Fourier ptychographic microscopy: supplemental document

This document provides supplementary information to "Fusion-Based Enhancement of Multi-Exposure Fourier ptychographic microscopy". We further elaborated on the structure and parameter details of the multi-exposure image fusion framework, the significance of the feature-domain reconstruction algorithm, its integration with the multi-exposure image fusion framework, and presented numerical simulation results as well as results from traditional Fourier ptychographic microscopy (FPM) reconstruction algorithms. This more comprehensive demonstration highlights the functionality and significance of MEIF, followed by additional discussions.

1. MEIF STRUCTURE DETAILS

In Fig. 1, we introduce the MEIF network, comprising four convolutional layers, a feature fusion module, and input-output components. The first two of these convolutional layers are responsible for feature extraction, while the latter two are tasked with performing a weighted average to generate the output image. This arrangement effectively embodies the dual roles of convolutional layers in image processing: feature extraction and weighted averaging for output [**?**].

A. Feature extraction module

The feature extraction module comprises two pivotal convolutional layers, CONV 1 and CONV 2, essential in image fusion algorithms. However, achieving effective feature extraction using conventional convolutional kernels presents a challenge due to their randomness. Therefore, we leverage pre-trained results from another dataset (ImageNet) to initialize the initial convolutional layer, referred to as CONV 1 in MEIF. This layer comprises 64 convolutional kernels of size 7×7, sufficient for extracting a substantial number of features. Moreover, to enhance effectiveness, the parameters of CONV 1 remain fixed during training. To further facilitate the recognition of features extracted by CONV 1, we introduce CONV 2, which fine-tunes the convolutional features of CONV 1 and enhances their adaptability in feature fusion.

B. Feature fusion module

The purpose of the feature fusion module is to combine features obtained from multiple sets of images into a single set for eventual reconstruction. Here, we select an element-wise fusion approach, which can be mathematically described by Eq. (S1). This method has been proven to be highly effective for fully convolutional image reconstruction networks. Through meticulous experimentation and validation, and considering the emphasis on rich detail in computational microscopy imaging, element-wise maximum fusion has demonstrated the best performance. Therefore, all subsequent work in our research follows the element-wise maximum fusion approach.

$$\widehat{f}^{j}(x,y) = \underset{i}{\mathrm{fuse}}\left(f^{j}_{i,C_2}(x,y)\right), 1 \leq i \leq N \tag{S1}$$

C. Feature reconstruction module

The image reconstruction module consists of an additional two convolutional layers, which utilize CONV 3 and CONV 4 for the weighted combination of features to produce the fused reconstruction result. CONV 3 further refines fused features and CONV 4 then reconstructs these features into a 3-channel output through elementwise averaging. The parameters of these two convolutional layers are determined through training, and their dimensions need to match those of CONV 2 and the final expected single-channel output.

D. Parameter setting of MEIF network

In these four convolutional layers, CONV 1 remains pre-trained, with its parameters unchanged during training, while the parameters of the other three convolutional layers are optimized based on the training data. The purpose of CONV 2 is to further extract features, thus its parameters need to match those of CONV 1; CONV 3 further fine-tunes the convolutional features, hence sharing parameters with CONV 2; CONV 4 reconstructs all features into a single output. The dimensional parameters of each convolutional layer should adhere to these principles. Additionally, for better training, both CONV 2 and CONV 3 are equipped with ReLU activation layers [**?**] and batch normalization layers to achieve improved output performance.

Based on these principles, the final parameters of the network are determined as follows: CONV 1 comprises 64 kernels of size 7 × 7, with a stride and padding set to 1 and 3 respectively; CONV 2 and CONV 3 both consist of 64 kernels with a size of 3 × 3, a stride of 1, and padding of 1; CONV 4 has a kernel size of 1 × 1, aimed at compressing the 64-channel feature maps into the target image.

In conclusion, the structure and parameters of the MEIF network are established through systematic parameter selection and integration of a pre-trained network. This ensures stable implementation of the image fusion functionality as a preprocessing framework, with guaranteed generalizability post pretraining. Although CNN networks have matured in the field of computer vision [**? ?**], making their debut in the realm of microscopy imaging. The training of the network relies entirely on openly available datasets commonly used in computer vision, such as ImageNet and datasets publicly. Given that many datasets are in RGB three-channel format, in the specific process of computing microscopy imaging, we need to actively enhance grayscale images and then directly utilize MEIF for preprocessing to obtain results suitable for reconstruction.

2. EXPERIMENTAL SETUP AND PARAMETERS

Table S1. **Equipment parameters**

	LED Array	LED Spacing	Illumination Distance	Objective
Equipment 1	17*17	4mm	70.0mm	4×/0.1NA,Nikon
Equipment 2	19*19	4mm	65.0mm	4×/0.1NA,Nikon

To comprehensively validate the generality and effectiveness of our proposed MEIF preprocessing method, experiments were conducted using various samples and experimental setups with different parameters (see Table S1). Specifically, raw data for the USAF and onion epidermis samples were acquired using Equipment 1, while data for the connective tissue samples were collected using Equipment 2. According to calculations, the numerical apertures were increased by 6.4 times and 7.2 times for the two systems, respectively. Theoretically, this extends the resolution to 3.1600 micrometers and 2.2850 micrometers, respectively.

Moreover, to underscore the versatility, different exposure times were utilized for Equipment 1 and Equipment 2. Equipment 1 employed red light with a wavelength of 632 nanometers, whereas Equipment 2 utilized blue light with a wavelength of 457 nanometers for illumination (denoted by asterisks in Table S2). Owing to the distinct wavelengths and sample characteristics, the exposure times required to achieve the same exposure level varied. Detailed experimental parameters are provided in Table S2.

Table S2. **Experimental parameters**

Sample	EV-1	EV0	EV+1	EV+2	EV+4	Equipment
USAF resolution calculator	2ms	4ms	8ms	16ms	64ms	Equipment 1
Onion epidermis	8ms	16ms	32ms	64ms	256ms	Equipment 1
Connective tissue*	1ms	2ms	4ms	8ms	32ms	Equipment 2

For all samples, we consistently employed a standardized set of relative exposure parameters. The EV 0 exposure, defined as the normal exposure, was determined using the camera's automatic settings. Given the broad range of selected exposure values, minor deviations are unlikely to impact the results significantly, allowing us to confidently rely on either the automatic exposure or subjective judgment. Once the exposure time for EV 0 is established, we calculate the corresponding exposure times for EV -1, EV +1, EV +2, and EV +4, and proceed with the multi-exposure data collection. The chosen range, spanning from EV -1 to EV +4, is sufficiently broad to capture fine details from both the dark regions in bright-field imaging and the majority of useful information in dark-field imaging. This range ensures comprehensive data acquisition across various conditions.

In terms of exposure selection, the lower EV values (e.g., EV -1, EV 0, EV +1) are primarily used for conventional HDR imaging, while the higher EV values (e.g., EV +2, EV +4) are focused on capturing richer details in dark-field imaging—information that traditional HDR methods often fail to utilize sucessfully (see Sec. 4. INTENSITY INVERSION AND MISALIGNMENT) .

Through validation across multiple samples, this exposure range has proven highly suitable, consistently providing rich detail without introducing excessive noise. As a result, we recommend adopting the same exposure parameters for further experimentation and refinement. Although expanding the exposure range is possible, experimental results indicate that doing so does not significantly enhance image quality. Thus, the selection of EV -1 to EV +4 strikes an ideal balance between optimal imaging performance and efficient information capture.

3. MULTI-EXPOSURE IMAGE FUSION AND HIGH DYNAMIC RANGE

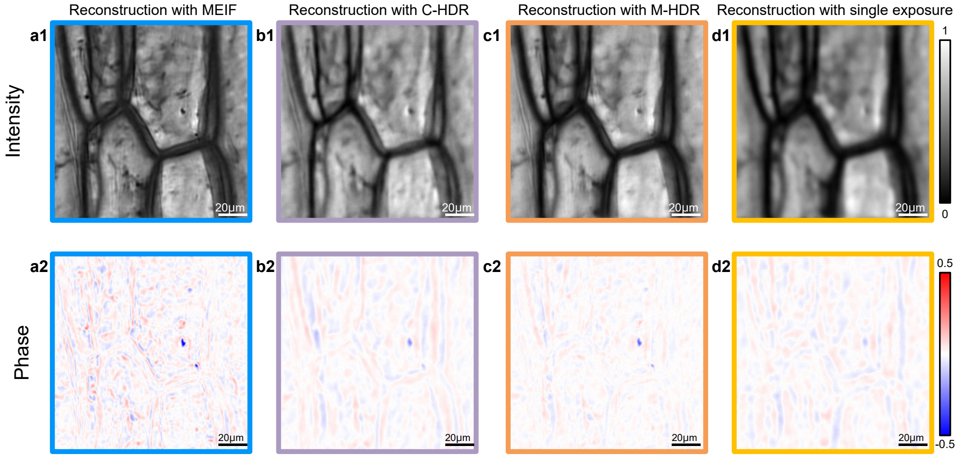


Fig. S1. Reconstruction results of a local region in onion epidermis samples using MEIF, HDR, and single-exposure images. The numbers 1 and represent the intensity and phase of the reconstructed results, respectively. (a) Reconstruction with MEIF. (b) Reconstruction with C-HDR, involving a process that truncates overexposed and underexposed regions before weighted averaging. (c) Reconstruction with M-HDR, directly averaging all exposure results. (d) Reconstruction with single-exposure raw data.

HDR was introduced to FPM in 2013 [**?**] but has not received adequate attention over the years. Essentially, MEIF utilizes a dataset same to HDR and possesses the capability to extend the dynamic range. MEIF leverages CNNs for feature extraction and combination, resulting in a fused output. Traditional HDR methods typically adopt two main approaches: one involves directly linearly combining all multi-exposure images, yielding an average result known as mean HDR (M-HDR). The other method restores each image's exposure gain to the same dynamic range, removes overexposed and underexposed portions, and calculates a weighted average. This approach is consistently referred to as C-HDR.

MEIF and both HDR methods can effectively enhance the information content of the original data (see Supplementary 1). However, the difference lies in the fact that the information enhancement provided by MEIF can be more efficiently utilized by reconstruction algorithms, resulting in better reconstruction outcomes.

In Fig. S1, a small section of onion epidermis is reconstructed using MEIF, different HDR algorithms, and a single normal exposure. The results presented in Fig. S1(b1-c2), obtained through the two HDR methods, demonstrate a noticeable improvement compared to the single-exposure reconstruction. The most impressive improvement is still shown by the meif effect in

Fig. S1(a1-a2) , where all the details become sharp and we can even clearly see some structures in the cell. However, there is no significant difference between the two HDR methods, suggesting that C-HDR and M-HDR show little distinction in FPM. The comparable enhancement effects demonstrate the imperfect utilization of dark-field information by C-HDR methods. From another perspective, this also underscores the immense potential of dark-field information. Therefore, many researchers persist in advocating for the universal use of HDR in any scenario. MEIF may present itself as a better choice for them in the future.

4. INTENSITY INVERSION AND MISALIGNMENT

In multi-exposure imaging, when the exposure range is too large, images with excessively high exposure are inevitably included. Under such conditions, due to the limitations of the sensor's maximum signal capacity, truncation and incomplete absorption are bound to occur, leading to inevitable signal distortion. Although HDR can somewhat alleviate this phenomenon, its reliance on linear combination still leaves it vulnerable to signal distortion in certain exposure groups, resulting in issues when utilizing large-range multi-exposure images. Below, we will provide a detailed demonstration and comparison of this phenomenon.

A. Reconstruction Results with Single Exposure

A.1. Simulation Results

During our simulation experiments, we first saved the generated data as PNG images before reading them, simulating the compression of raw data by camera sensors and processors.

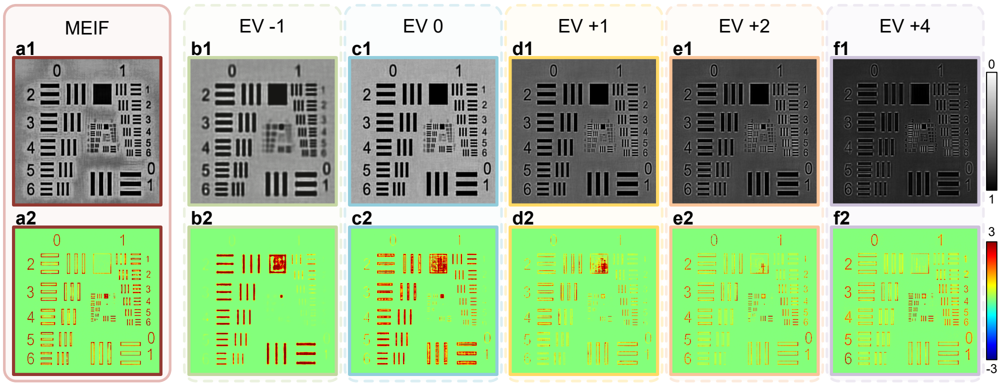


Fig. S2. Reconstruction results of FPM using MEIF based on simulated data. The recovery algorithm employed is FD-FPM. Images (a1) and (a2) depict the phase map and intensity map after MEIF, respectively. Images (b1), (c1), (d1), (e1), and (f1) show the intensity of the recovery results under exposures of EV -1, EV 0, EV +1, EV +2, and EV +4, respectively. Images (b2), (c2), (d2), (e2), and (f2) illustrate the phase of the recovery results.

Fig. S2 displays the recovery results with MEIF and the results under different single-exposure conditions without MEIF. In the single-exposure results, we initially observed that underexposed conditions exhibited a significant lack of high-frequency details (see Fig. S2(b1-b2)), whereas overexposed conditions revealed mismatches in the physical model due to sensor saturation, resulting in abnormal image intensities (see Fig. S2(d1-f1)). This abnormality is the intensity inversion we mentioned, which becomes more pronounced in practical experiments. However, these conditions also provide more phase information, noticeably increasing the resolution of phase data (see Fig. S2(d2-f2)), without experiencing obvious phase overexposure and detail loss as seen in experiments. When focusing on the results after MEIF preprocessing, we observed consistent and optimal resolution with rich phase information. This is reflected not only in sharp edges but also in high phase resolution. Essentially, MEIF intelligently fuses meaningful data from all exposure levels, significantly enhancing the information content of the FPM pipeline.

A.2. Experimental Results

In the experiments, we observed more detailed results of intensity inversion.

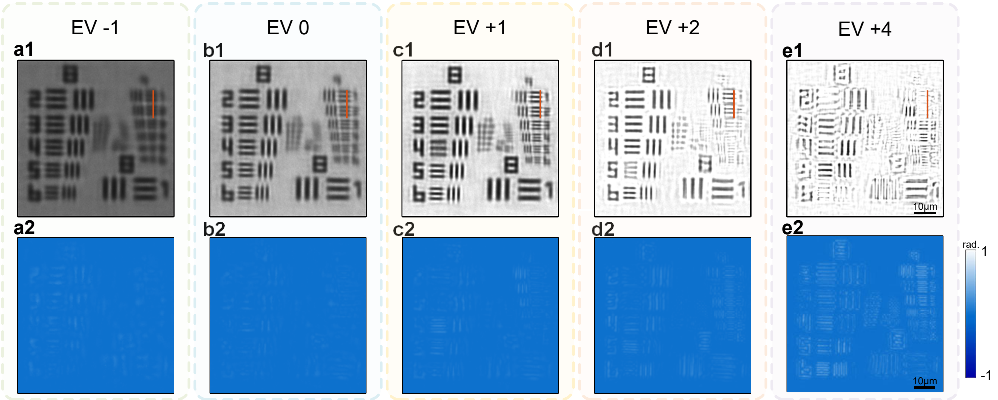


Fig. S3. Reconstruction results of FPM using FD-FPM based on experimental data. (a1-e2) Intensity and phase reconstruction results for EV -1, EV 0, EV +1, EV +2, and EV +4, respectively.

Reconstruction results are optimal under normal and slightly overexposed conditions. This observation is consistent with conventional experimental data and suggests that additional dark-field information enhances reconstruction. It is also evident that intensity inversion becomes more pronounced, and severe misalignment occurs under high exposure conditions. This misalignment will be discussed in more detail in the following section.

B. Inversion and Misalignment

As previously discussed, inversion and misalignment are key issues. Color inversion, in particular, clearly indicates misalignment of detailed regions. This phenomenon is more distinctly observed through the quantitative comparison of the highlighted areas in Fig. S3.

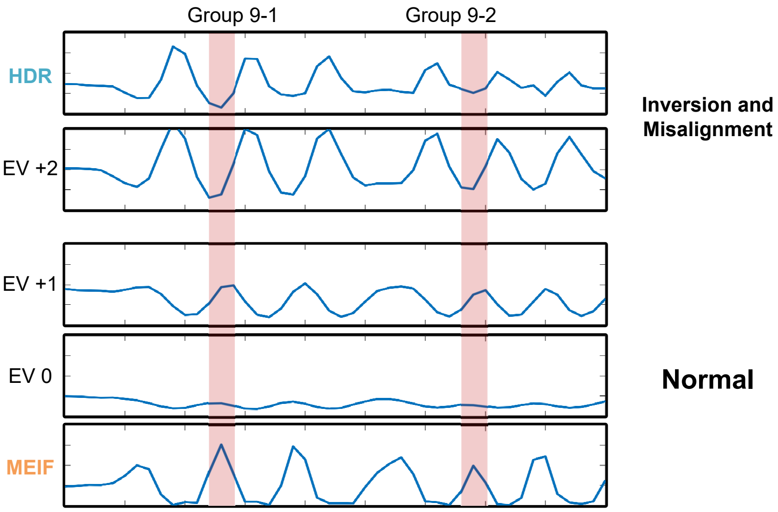


Fig. S4. Quantitative comparison of the highlighted regions based on Fig. S3. (For easier comparison, the maximum value is truncated at 2, rather than the commonly used 1. The minimum value is set to 0.)

From the figure, it is evident that with more appropriate exposure, the intensity contrast becomes more pronounced. However, when the brightness reaches EV +2, intensity inversion occurs, causing the line to curve upwards rather than downwards, which deviates from the expected pattern of white on black. This represents the inversion phenomenon. Additionally, in Fig. S4, we observe that the peaks and valleys in the HDR and EV +2 curves are opposite to those in other curves. The red-highlighted area shows this discrepancy more clearly. The presence of inversion results in reversed peaks and valleys, effectively creating a misalignment effect that impacts observation. Notably, MEIF avoids this issue while maintaining high resolution.

C. Resolution Enhancement

As observed in Fig. S3, many single-exposure images exhibit blurring due to overexposure. However, thanks to MEIF's robust information utilization capabilities, it can effectively reconstruct high-resolution images from these misaligned and blurred data. For instance, with MEIF reconstruction results, the data for Group 9-6 can be resolved, whereas single-exposure images typically only allow resolution up to Group 9-3. This demonstrates that MEIF, through advanced image fusion techniques, can achieve theoretical resolution even with imperfect data, significantly broadening the practical applications of FPM.

5. RAW DATA WITH MEIF AND HDR

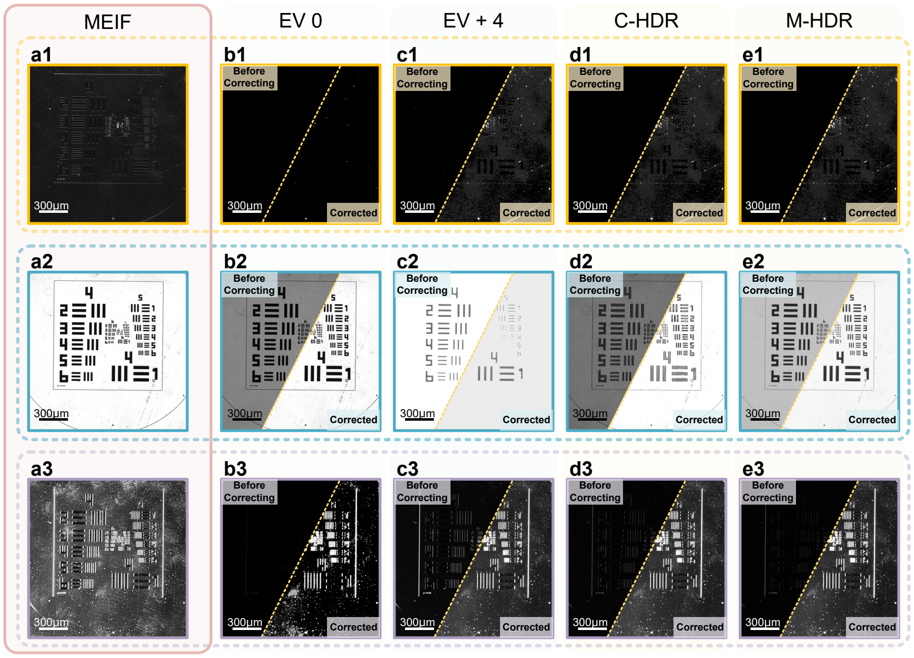


Fig. S5. Comparison of Partial Raw Data and Their Preprocessed Versions with HDR and MEIF. (a1-a3) Raw data after MEIF preprocessing, (b1-b3) Raw data with normal exposure, (c1-c3) Raw data under EV4, (d1-d3) Raw data after c-HDR preprocessing, (e1-d3) Raw data after mHDR preprocessing. The color and number index represents the relative position and illumination angle of the images, corresponding to the color scheme in Fig. 3.

Different preprocessing methods' effects on raw data are demonstrated in Fig. S5. Clearly, MEIF exhibits the best performance. To better illustrate the enhancement of information by MEIF, we conducted brightness matching on the right side of each image, ensuring that this part of the image has the same brightness as the MEIF result, facilitating direct comparison of information content. As shown in Fig. S5(b1-c3), the results of single exposure exhibit a lack of information, with a significant amount of noise and data loss after brightness matching. Traditional HDR methods also show similar phenomena, but it can be seen that c-HDR has better information extraction capability, although this information cannot be perfectly utilized by existing image reconstruction algorithms, as evident from the comparison in Fig. S5(d3,e3). In comparison, the advantages of MEIF are greatly demonstrated. Although it may not be as evident during low-angle illumination, its superiority is clear during high-angle dark-field illumination, as shown in Fig. S5(a1, d1, e1). MEIF extracts more valuable information from the feature domain while discarding the undesirable information from the original overexposed data, which represents a significant advancement over traditional HDR algorithms.

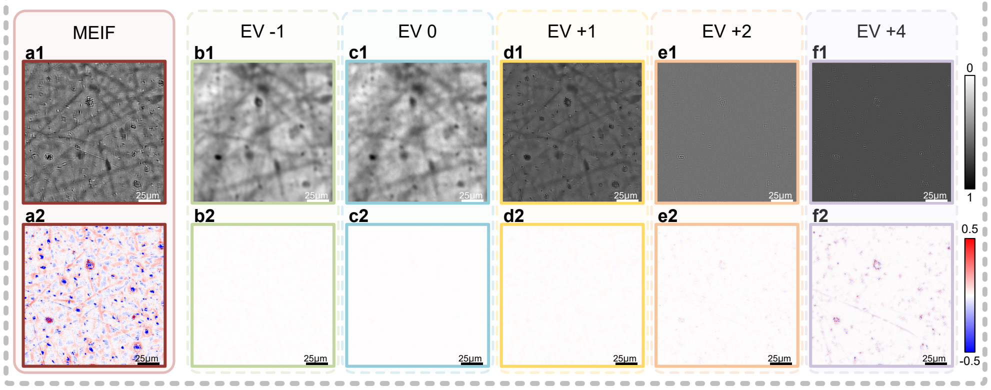
6. CONNECTIVE TISSUE

We chose connective tissue as the subject for further experiments, which is an animal sample with abundant details. Here, we compared the results obtained from traditional FPM, FD-FPM, and MEIF pre-processing.

In principle, the abundance of dark information provided by MEIF makes it possible to achieve better reconstruction results even with traditional FPM. Therefore, it is meaningful to examine the results of the original data from this set of samples under the traditional FPM reconstruction algorithm after MEIF.

The findings indicate that MEIF exhibits significantly better phase resolution and more reliable intensity recovery results. Regardless of the restoration algorithm used, the surplus information provided by MEIF is efficiently utilized, demonstrating the reliability and effectiveness of MEIF.

Box 1 | **Feature-domain FPM reconstruction**



Box 2 | **Traditional FPM reconstruction**

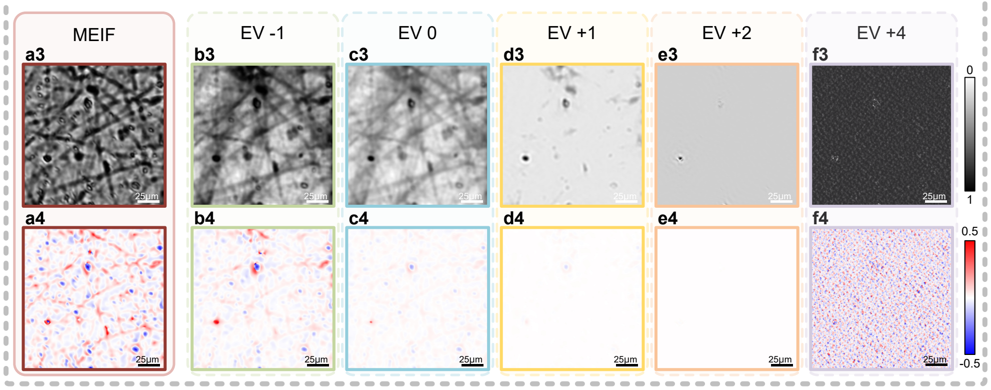


Fig. S6. Recovery results of a region of connective tissue under different FPM reconstruction algorithms. (a1-a4) depict the reconstructions after MEIF preprocessing, while the rest are reconstructions from single-exposure FPM. The images are divided into two boxes: Box 1 shows results of FD-FPM reconstruction, and Box 2 presents results of traditional FPM reconstruction. Within Box 1, (a1-a2) represent the MEIF reconstructions, where (a1) displays intensity and (a2) displays phase. Similarly, (b1-b2), (c1-c2), (d1-d2), (e1-e2), and (f1-f2) showcase intensity and phase reconstructions for EV -1, EV 0, EV +1, EV +2, and EV +4, respectively. The same applies to Box 2, illustrating reconstructions with traditional FPM for the same dataset in a similar sequence: (a3-a4), (b3-b4), (c3-c4), (d3-d4), (e3-e4), and (f3-f4) represent MEIF-preprocessed, EV -1, EV 0, EV +1, EV +2, and EV +4, with intensity on the upper half and phase on the lower half.

First, focusing on the results of FD-FPM in Box 1 of Fig. S6, it is evident that the most informative recovery results are achieved after MEIF preprocessing [see Fig. S6 (a1-a2)]. In terms of intensity, it exhibits the most detailed features, such as textures and grooves within the tissue, as previously demonstrated. Similarly, regarding intensity, overexposure leads to intensity reversals, resulting in reduced details and a cluttered background, which is unacceptable in biological tissues [see Fig. S6 (d1-f1)]. Given the complexity of the structure, a cluttered background implies a significant loss of information. MEIF effectively extracts valuable information, resulting in excellent intensity resolution. Particularly noteworthy are the results in the phase domain, where MEIF provides exceptionally rich phase information. We can clearly discern structures, fibers, and even some finer details, which are mere highlights under single-exposure conditions, representing a significant advancement.

Next, examining the reconstruction results of traditional algorithms, displayed in Box 2 of Fig. S6. Similarly, exhibiting trends akin to traditional FPM, the intensity images from overexposed conditions are entirely unusable [see Fig. S6 (d3-f3)]. However, due to the less advanced algorithms, they fail to deliver reconstructions comparable to FD-FPM. The phase images from single-exposure conditions further highlight the phase noise that increases with exposure time, resulting in a loss of information or clutter [see Fig. S6 (d4-f4)]. Nonetheless, MEIF produces surprising results [see Fig. S6 (a3, a4)], offering excellent contrast and rich information in intensity, along with significant detail in phase, compared to reconstructions from traditional single-exposure algorithms. However, when compared to FD-FPM after MEIF preprocessing showed in Fig. S6 (a1, a2), it falls slightly short in noise performance and detail richness. This further underscores the combination of FD-FPM with MEIF as an exceptionally promising choice for future FPM applications.