Supplementary Information:
Cascaded adaptive aberration-eliminating multimode fiber imaging

Zhong Wen,[a, b, c †] Qilin Deng, [a, b†] Quanzhi Li, [a, b, c] Yizhou Tan,[d, e] Jingshan Zhong,[f] Chiming Zhang, [a, b, c] Jiahe Zhang,[c] Clemens F. Kaminski,[g] Ying Gu, [d, e] Xu Liu, [a, b, c] Qing Yang, [a, b, c *]

[a] Zhejiang University, College of Optical Science and Engineering, State Key Laboratory of Extreme Photonics and Instrumentation, Hangzhou, China.
[b] Zhejiang University, International Research Center for Advanced Photonics, Zhejiang University, Hangzhou, China.
[c] Zhejiang University, ZJU-Hangzhou Global Scientific and Technological Innovation Center, Hangzhou, China.
[d] Chinese PLA General Hospital, The First Medical Center, Department of Laser Medicine, Beijing, China.
[e] Chinese PLA General Hospital, Hainan Hospital, Laser Medicine Center, Sanya, China.
[f] Zhejiang Laboratory, Research Center for Intelligent Manufacturing Computing, Hangzhou, China.
[g] University of Cambridge, Department of Chemical Engineering and Biotechnology, Cambridge, UK.

* Corresponding authors, E-mail: qingyang@zju.edu.cn (Qing Yang)

Supplementary Note 1: Implementation steps of DECOUPLE

The DECOUPLE is divided into four stages. The first and second stages are depicted in Fig S1, while the third and fourth stages are illustrated in Fig S2. To effectively control the wavefront across different pupils at a fiber's output, DECOUPLE must first accurately ascertain the multimode fiber's transmission matrix. This is accomplished by aligning two orthogonally polarized Hadamard measurement basis vectors at the fiber and subsequently measuring the two orthogonal complex amplitude distributions -post-fiber, as described in [18]. In stage 1, a raster scan is implemented firstly using the transmission matrix. The focusing points scattered by the scattering samples are out of point shape after transmission. This distorted point spread function leads to blurred images, as shown in stage 1 in Fig. S1. Then we select the field of view (FOV) by applying a mask. The purpose of stage 1 is to identify the target field of view.

In stage 2, different tilts are added to different sub-pupils so that the focus point of each sub-pupil can align at the same spatial position. Without these tilts, aberrations would deflect these focuses, preventing them from focusing at the same position. As demonstrated in Fig. S1, a Fourier transform is applied to transmission matrix to obtain spatial frequency domain transmission matrix T_f. This output spatial frequency domain is then segmented and engineered accordingly. Combined with full vector modulation [18], we can modulate the wavefront on the pupil at the output of the fiber as follows:

$$\psi_{In}(\lambda, f_x, f_y) = \mathrm{conj}(T_f\ )K\psi_{Out}(\lambda, f'_x, f'_y) \qquad (S1)$$

Here, $\psi_{In}(\lambda, x_f, y_f)$ is the spatial frequency domain of wavelength, phase, and polarization of the input light field, $\psi_{Out}(\lambda, x'_f, y'_f)$ represents the spatial frequency domain distribution of the wavelength, phase, and polarization of the output light field, K is the pupil splitting factor, and f_x, f_y and f'_x, f'_y denote the two-dimensional spatial frequency coordinates at the ends of the optical fiber, respectively. $\mathrm{conj}(T_f\ )$ is the transposed conjugation of T_f .

Using the above formula, construct a position scanning matrix $TM_{position}$ for the MMF output, where only one sub- pupil is active in the spatial frequency domain. Each sub-pupil undergoes an independent scan and imaging to obtain an image, as shown in stage 2 in Figure S1. We can calculate the deflect component in each sub-pupil by comparing the lateral position of the image (based on the cross-correlation operation of the image). Then, these components can be inversely added into the pupil to cancel out the aberration influence.

The next part is identifying the phase shift in each sub-pupil. In stage 3, the relative phase of each sub-pupil is adjusted following the correction of their spatial positions. Open two sub-pupils at a time until the phase-shift correction of all sub-pupils is complete. This process requires turning on two sub-pupils, fixing the phase of one reference pupil, and adding different phases to the other sub-pupil one by one. Maximum phase length interference is determined by the fluorescence/reflected light intensity returned from the sample, as shown in

stage 3 in Figure S1. The maximum value means achieving the constructive interference.

The final stage involves applying the deflection components and the phase shifts identified in stages 2 and 3 to all sub-pupils. As shown in Fig S2, stage 4 corrects the tilt and phase-shift of all sub-pupils, and then we use this corrected pupil for raster scan imaging and get sharp images.

Supplementary Note 2: The tradeoff of the number of pupil divisions

When the optical pupil is divided into multiple sub-pupils, the number of modulated modes for each sub-pupil decreases accordingly. We tested this by constructing the transmission matrix of a 50-micron core diameter, 0.22 NA fibre, and reducing the number of input modes. A linear decrease in PR (the ratio of power stored in the optimized focus to the total power leaving the MMF's output facet) and intensity was observed (as shown in Fig. S3). This indicates that excessive pupil segmentation significantly reduces the progress of measurements for individual sub-pupils, leading to less effective aberration correction compared to cases with fewer divisions.

Supplementary Note 3: The difference from previous work

Compared with STABLE [18], there are essential differences in the core problem solved and techniques. Firstly, our current paper addresses an important problem that diverges from previous work (as shown in Fig. S4). Previous research primarily focused on real-time tracking and compensation of fiber optic state changes during bending, offering high stability and robustness against the movement and deformation of the probe. However, a significant limitation of this approach was that it only provided surface images (left image in Fig. S4). DECOUPLE aims to solve the problem of scattering and to achieve clear imaging through tissues (right image in Fig. S4). This advancement is crucial for non-invasive visualization of deeper structures, which has significant implications for medical applications. Furthermore, the current work enables simultaneous multicolor imaging, which compensates the dispersion, deformation, and aberration dispersion simultaneously. The previous work STABLE primarily compensates for deformation and operates at a single wavelength.

Secondly, DECOUPLE differs from prior methods in its novel techniques (as shown in Fig. S5). Previous work primarily used the STABLE method, which relied on a single-pixel beacon to track fiber state changes and compensating in real time by referencing a pre-recording library of transfer matrices. STABLE doesn't correct aberration. In contrast, this work introduces a novel aberration measurement scheme that does not require wavefront sensors. DECOUPLE accurately capture tissue aberrations over time and compensate for these aberrations by adjusting the incident wavefront, thereby achieving deep imaging. Additionally, DECOUPLE can achieve dual-color imaging while ensuring accurate aberration measurement and compensation even with fiber bending.

Compared to traditional adaptive optical imaging, which is limited to rigid fixed-mode systems and low-order aberrations in objective lenses or GRIN lens, DECOUPLE introduces a novel approach tailored for the complex light transmission of MMF. This marks a significant departure from the rigid structures used in other studies (Table S2). The flexibility of MMF is crucial for adaptive imaging in the dynamic environment of biological tissues. Our method innovatively combines a cascaded transmission matrix with pupil segmentation wavefront sensing to address and correct both fiber dispersion and sample aberrations, overcoming the limitations of traditional methods. Unlike approaches such as single-pixel spatial frequency beacon tracking or frequency multiplexing with wavefront modulating devices, our technique is designed specifically for in vivo imaging. It achieves non-invasive volumetric imaging in two colors, surpassing the single-color capability of some counterparts and avoiding the need for sample movement compensation. In summary, our work presents a unique, advanced solution for endoscopic imaging, offering high-resolution, flexible imaging with comprehensive aberration correction for complex optical systems.

Supplementary Note 4: Reflection matrix measurement method

We developed an algorithm to retrieve the reflection matrix from intensity images measured using a multi-color reflection measuring setup without a reference light. The reflection matrix records the linear relationship between the incident field and the reflected field. This method modulates the incident light field through a series

of phase modulations, then measures the intensity images of the reflected light field to achieve transmission inversion. The intensity images are denoted as $P_n(x, y)$, where n = 1...N, and x, y represent spatial coordinates. Each image contains $N_x * N_y$ pixels. The phase modulation is represented by the vector ψ_{in_n}, with N_k modulation modes for both polarizations. The forward model for intensity measurements is

$$\mathrm{P}_n = \left|\mathrm{RM}(\lambda)\psi_{in\_n}\right|^2 \tag{S2}$$

Here, P_n is the vector corresponding to $P_n(x, y)$, $\mathrm{RM}(\lambda)$ is the reflection matrix, and $|\cdot|^2$ denotes the squared modulus of the complex number. The size of the reflection matrix is $N_x * N_y * N_k$.

$$\min_{\mathrm{rm}^k} f(\mathrm{rm}^k) = \left\|\mathrm{P}^k - |\mathrm{Q}\mathrm{rm}^k|^2\right\|_2^2 \tag{S3}$$

The number of unknowns in the transmission matrix is typically very large, which makes the optimization problem difficult to solve directly. However, it can be decomposed into $N_x * N_y$ smaller optimization problems. Each problem is based on an equation derived from the intensity measurement of a single pixel.

$$\mathrm{P}^k = \begin{bmatrix} \mathrm{P}_1^k \\ \mathrm{P}_2^k \\ \vdots \\ \mathrm{P}_n^k \end{bmatrix}, \mathrm{Q} = \begin{bmatrix} \psi_{in\_1}{}' \\ \psi_{in\_2}{}' \\ \vdots \\ \psi_{in\_n}{}' \end{bmatrix} \tag{S4}$$

The column vector rm^k is the transpose of the k-th row of the reflection matrix and P_n^k is the k-th element of P_n, where $'$ denotes the transpose. The vector P^k contains all measurements at the same pixel indexed by k in the intensity image. The matrix Q is referred to as the measurement matrix; each row of Q corresponds to an incident light field. Each optimization problem recovers one row of the reflection matrix from the measurements at the corresponding pixel. Therefore, by independently solving these smaller optimization problems, the entire reflection matrix can be recovered.

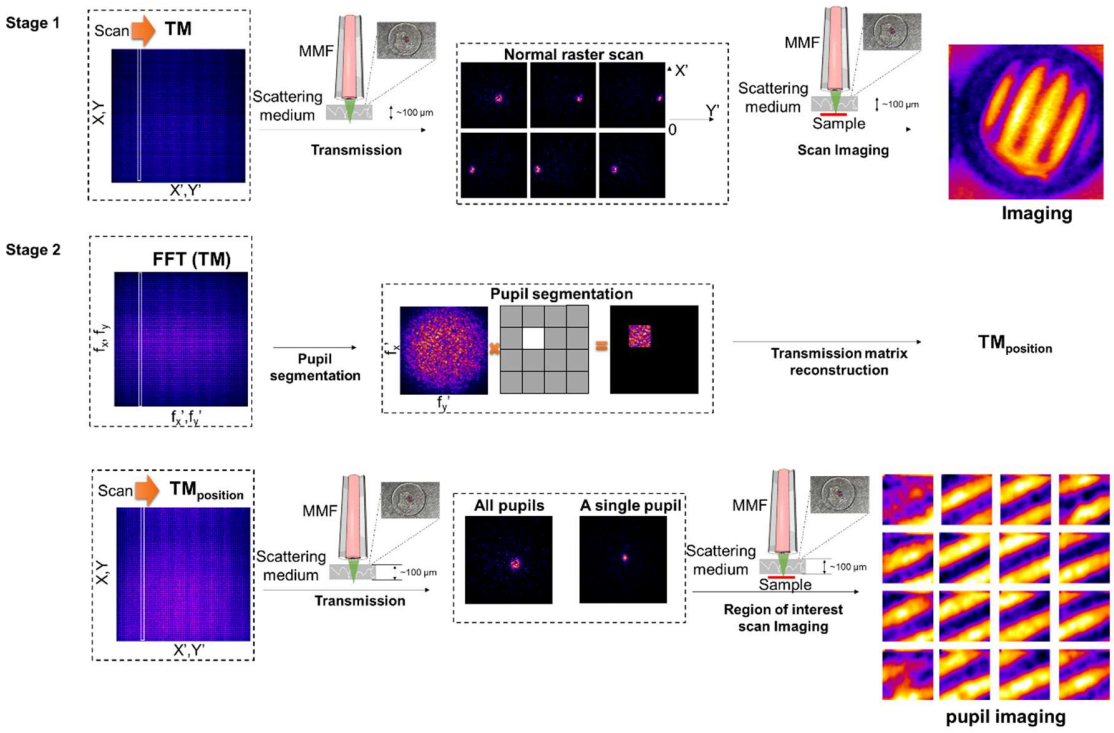

Fig. S1 Stage 1 and stage 2 diagram. Stage 1 is scanning and to identify the target field of view. Stage 2 is to calculate the deflect component in each sub-pupil. The $TM_{position}$ represents position scanning matrix.

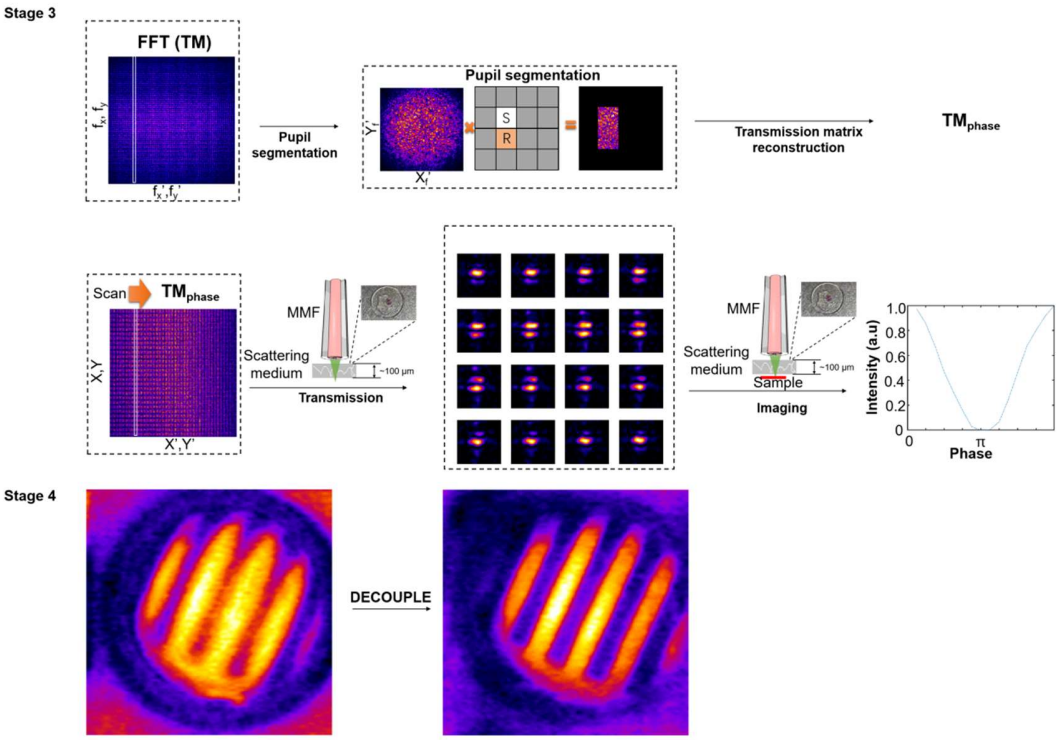

Fig. S2 Stage 3 is to identify the phase shift in each sub-pupil. Stage 4 is applying the deflection components and the phase shifts identified in Stage 2 and 3 to all sub-pupils, then getting clear image. The TM_{phase} represents phase scanning matrix.

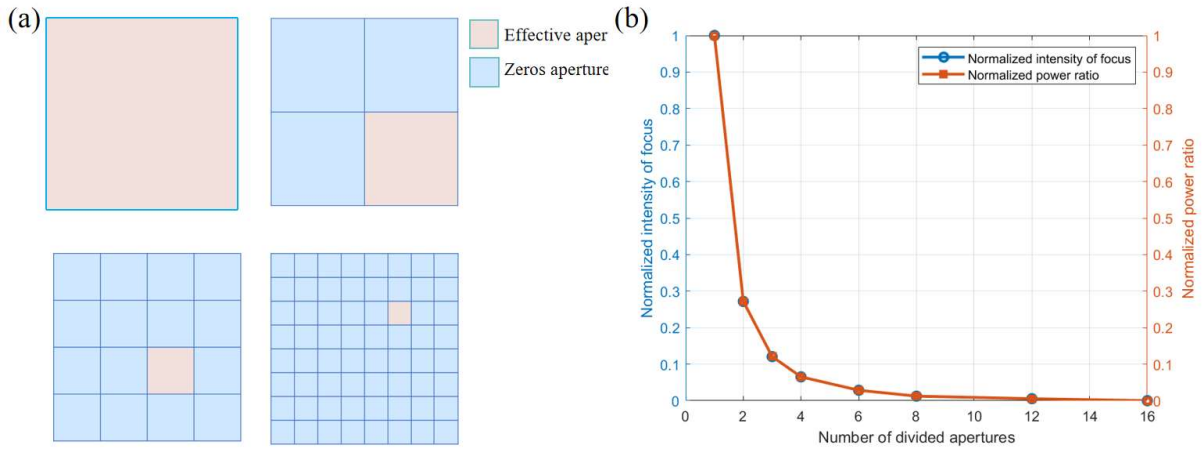

Fig. S3 Schematic of region segmentation versus signal-to-noise ratio (SNR) and focal intensity. (a) Illustrates the segmentation of the modulation region into pupil sizes of 1×1, 2×2, 4×4, and 8×8. (b) Shows the normalized SNR and intensity at the focal point for different pupil segmentation. The SNR represents the ratio of the energy at the focal point to the total energy. The normalized SNR is the SNR value relative to the 1×1 pupil, with other pupil segmentations compared to this maximum value.

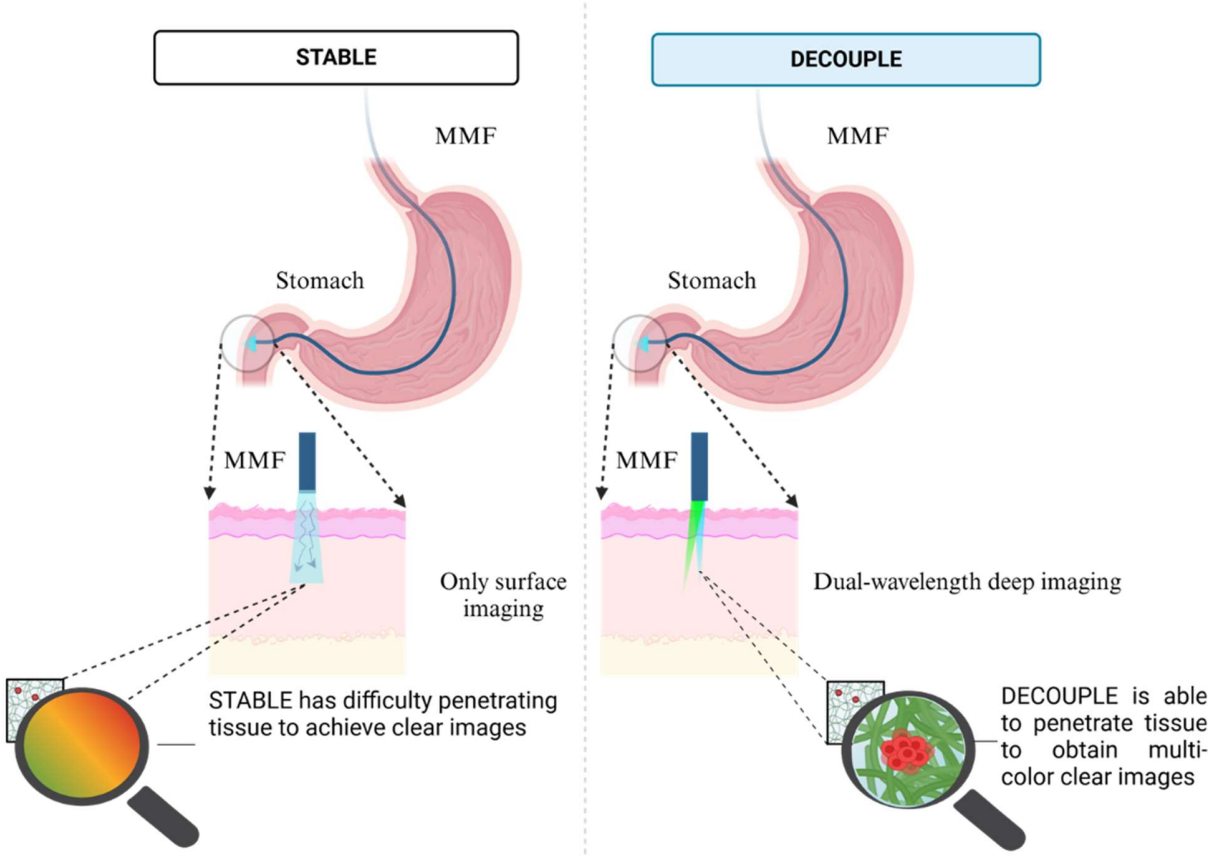

Fig. S4 STABLE and DECOUPLE solve different problems.

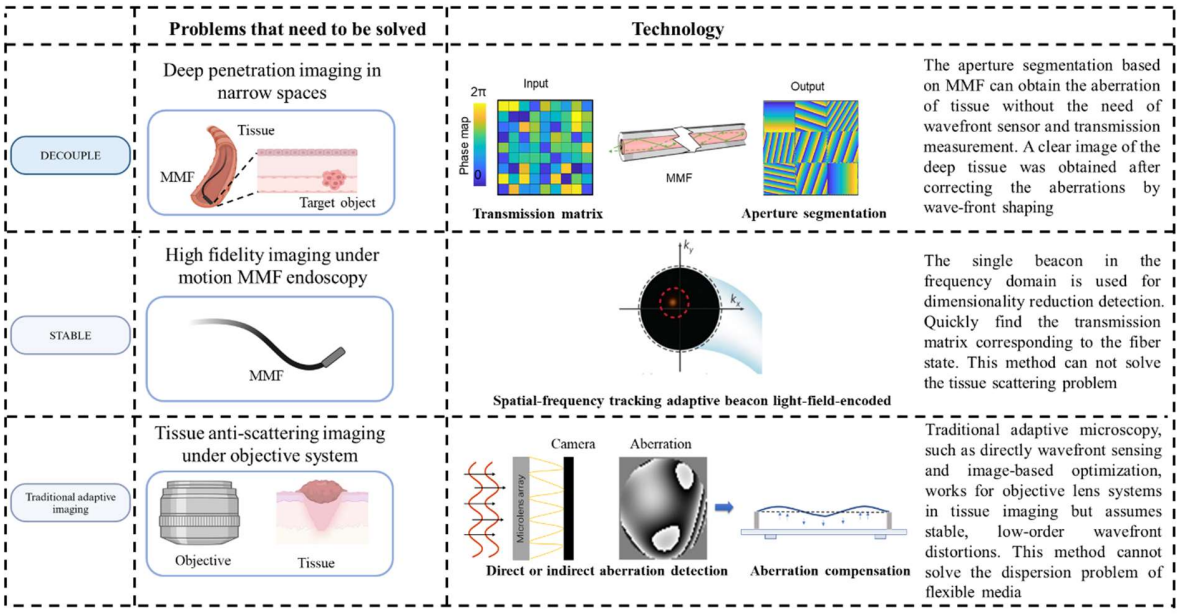

Fig. S5 STABLE and DECOUPLE employ different technical approaches.

Optical components	Model number
Laser 1	YDFL-488-SF-0.5-CW （PRECILASERS）
Laser 2	0561-0501-0500500 （COBOLT）
PBS1-3	MPBS641 (LBTEK)
BS1-3	MBS1455-A (LBTEK)
FOBC	RGB50HF（THORLABS）
M	BDM1-A(LBTEK)
QWP and HWP	Custom (UNION OPTIC)
DMD	FLDISCOVERY HCSLM137D70
OBJ	OPLN20X (OLYMPUS)
OF	MBF10-650-25 (LBTEK)
PMT	H7422P-40 （HAMAMATSU）
DM1	DM05-490SP (LBTEK)
DM2	DM05-550LP (LBTEK)
L1 and L8	MAD303-A (LBTEK)
L2 and L9	MAD419-A (LBTEK)
L3 and L5	AC254-200-A-ML（THORLABS）
L4	AC254-075-A-ML（THORLABS）
L7	MAD303-A (LBTEK)
L10	AC254-075-A-ML（THORLABS）
L11	AC254-200-A-ML（THORLABS）
CMOS	acA720-520um (BASLER)
Iris	SM1DP12-1A （THORLABS）
LP	FLP25-VIS-M (LBTEK)
F1 and F2	A240TM-A（THORLABS）
F3	AC254-100-A-ML（THORLABS）

Table. S1 Model and manufacturers of the optical components

Article/Feature	Suitable for flexible media (such as optical fibers)	Penetrating tissue imaging	Technological Approach
Our Research	**YES**	**YES**	Eliminate the fiber's dispersion and sample aberrations through **cascaded transmission matrix and pupil segmentation wavefront sensing.**
Opt. Lett., 32(16), 2309-2311 (2007) [39]	YES	NO	**Iteratively optimize** the input wavefront to focus the transmitted light onto a target area the size of a single spot
Nat. Photonics, 17, 679-687 (2023) [18]	YES	NO	**Single-pixel spatial frequency beacon tracking** to correct mode transformation
Light Sci. Appl., 12, 270 (2023) [38]	NO	YES	**Machine learning** to process imaging information, with a physics-informed neural network architecture
Sci. Adv., 6, eabc6521 (2020) [40]	NO	YES	Measure the aberrations of the GRIN lens in advance, and correspond them through a **lookup table method** during the experiment. Then, directly measure the tissue aberrations with a **Hartmann wavefront sensor** and compensate with a deformable mirror.
Nat. Methods 18, 1259–1264 (2021) [30]	NO	YES	The method of partitioned indirect wavefront sensing for aberration measurement using **frequency multiplexing with two wavefront modulating devices.**

Table. S2 DECOUPLE A different summary from previous work.