nature methods

Article

https://doi.org/10.1038/s41592-024-02244-3

Pretraining a foundation model for generalizable fluorescence microscopy-based image restoration

Received: 27 July 2023

Accepted: 13 March 2024

Published online: 12 April 2024

Check for updates

Chenxi Ma[1,2], Weimin Tan[1,2], Ruian He[1] & Bo Yan[1]

Fluorescence microscopy-based image restoration has received widespread attention in the life sciences and has led to significant progress, benefiting from deep learning technology. However, most current task-specific methods have limited generalizability to different fluorescence microscopy-based image restoration problems. Here, we seek to improve generalizability and explore the potential of applying a pretrained foundation model to fluorescence microscopy-based image restoration. We provide a universal fluorescence microscopy-based image restoration (UniFMIR) model to address different restoration problems, and show that UniFMIR offers higher image restoration precision, better generalization and increased versatility. Demonstrations on five tasks and 14 datasets covering a wide range of microscopy imaging modalities and biological samples demonstrate that the pretrained UniFMIR can effectively transfer knowledge to a specific situation via fine-tuning, uncover clear nanoscale biomolecular structures and facilitate high-quality imaging. This work has the potential to inspire and trigger new research highlights for fluorescence microscopy-based image restoration.

Fluorescence microscopy image restoration (FMIR), which aims to provide images with high signal-to-noise ratios (SNRs) from low-SNR images, has received significant attention from the research community, as it helps reveal important nanoscale imaging information for the accurate observation and scientific analysis of biological structures and processes[1–3]. Currently, benefiting from the rapid development of deep learning, the literature is experiencing a large influx of contributions in this direction. Much deep learning-based fluorescence microscopy-based image restoration works[4–15] (Supplementary Note 1) have pushed the physical limits of fluorescence microscopy through computations and have achieved significant improvements over the classic deconvolution algorithms[1,16].

Although significant progress has been achieved, these deep learning-based fluorescence microscopy-based image restoration methods are still affected by several weaknesses, limiting the further development of biological processes. First, the prevailing models address specific fluorescence microscopy-based image restoration problems, such as denoising, super-resolution (SR) and isotropic reconstruction, by training a specific deep model (for example, U-Net-inspired models[4,7–9,14] and RCAN-inspired models[5,6,10,13]) with limited parameters (no more than a few million) on a specific dataset from scratch (Supplementary Table 1b). In addition, these models have poor generalization, as significant performance degradations can be observed when facing large domain gaps between different datasets and different fluorescence microscopy-based image restoration problems. Achieving promising results across different imaging modalities, biological samples and image restoration tasks requires training multiple specific models. Last, the common data dependence problem in the deep learning field also affects most fluorescence microscopy-based image restoration models, the performance of which highly depends on the quality and quantity of the training data due to the data-driven characteristics of deep learning-based methods. Consequently,

[1]School of Computer Science, Shanghai Key Laboratory of Intelligent Information Processing, Fudan University, Shanghai, China. [2]These authors contributed equally: Chenxi Ma, Weimin Tan. e-mail: byan@fudan.edu.cn

the realistic difficulty of experimentally acquiring low-quality and high-quality training image pairs makes the practical application of deep learning-based fluorescence microscopy-based image restoration methods complicated[4]. Therefore, the main purpose of this paper is to overcome the above weaknesses and explore the performance upper limit of deep models while inspiring and fostering subsequent research.

As the latest generation of artificial intelligence models, foundation models[17,18], which can be applied to a diverse set of downstream tasks by training them on massive and diverse datasets, have profoundly advanced the development of deep learning and have exhibited remarkable domain transfer capabilities, particularly in the fields of natural language processing[19,20], computer vision[21,22] and multimodal learning[23–25]. Recently, the concept of foundation models has been utilized in diverse life science applications, and foundation models have also demonstrated their impressive capabilities for clinical cases[26–30], biotechnology[31,32] and so on (Supplementary Note 1).

As stated by Moor et al.[30], foundation models will offer amazing abilities through dataset size and model size increases, in addition to model architecture advances, in agreement with observations[32,33] that large-scale pretraining with larger and more diverse data consistently improved the model's predictive potential. We observed a similar phenomenon in the fluorescence microscopy-based image restoration field, which has never been studied before (Supplementary Note 2). Pretraining also enables better generalization ability and efficient training on new datasets with limited training data, by transferring knowledge in a pretrained model to a specific task or data modality. This claim can be supported by Zamir et al.[34], who explored the idea of transfer learning between many visual learning tasks and showed that the amount of training data required for solving multiple tasks together can be greatly reduced compared to the amount of training data required for independent training. In addition, assembling multiple image restoration processes in a foundation model is a more practical and convenient strategy, as it is difficult to directly determine which type of image restoration operation is needed for the realistic fluorescence microscopy images at hand.

However, task-specific or modality-specific deep models are still the main deep learning-based approaches for fluorescence microscopy-based image restoration. Although individual models can now achieve state-of-the-art (SOTA) performance, foundation models have the merit of versatility. Instead of training a new model from scratch for each task, the above approaches have demonstrated that a foundation model can democratize the fundamental knowledge, learned during the pretraining phase, in general datasets and can transfer this knowledge to a multitude of tasks through fine-tuning[32]. The enormous progress made by pretrained large-scale models brings new momentum to the development of fluorescence microscopy-based image restoration approaches.

Here, we first presented a UniFMIR solution to handle diverse image degradations and imaging modalities simultaneously. We took inspiration from the existing foundation models, where large pretrained models can be flexibly transferred to solve diverse tasks and achieve significant performance improvements via efficient fine-tuning. Specifically, we constructed the UniFMIR model, which adopted a multihead and multitail network structure (Fig. 1a and Extended Data Figs. 1 and 2). Specifically, UniFMIR consists of a multihead module, a feature enhancement module and a multitail module, where the multihead module and the multitail module adopt different branches to extract task-specific shallow features and yield accurate results for different image restoration problems, respectively. The feature enhancement module uses an advanced Swin transformer structure[35] to enhance the feature representations and to reconstruct general and effective features for high-quality fluorescence microscopy-based image restoration. Different fluorescence microscopy-based image restoration operations cover different head and tail branches, but share the same feature enhancement module. We collected a large training dataset (~30 GB) from 14 public datasets, covering a wide range of imaging modalities, biological samples and image restoration tasks. Furthermore, the UniFMIR is pretrained on this large-scale dataset and fine-tuned on different subdatasets, covering various degradation conditions, imaging modalities and biological samples. We also created a baseline by training models, which have the same architecture as UniFMIR, from scratch on specific datasets for different tasks to enable better comparison. We showcased that efficient fine-tuning can feasibly transfer the prior knowledge learned during pretraining to handle different problems (Supplementary Figs. 15–17). We demonstrated the effectiveness of the UniFMIR model on a set of high-impact applications and compared its performance with that of SOTA methods for solving specific problems.

Results

SR

The lack of high-resolution (HR) microscopy images has impeded the further exploration of the life science phenomena of related structures or cellular tissues. To overcome the theoretical spatial resolution limitation in live-cell imaging, SR, aiming to enhance the resolution of scientific microscopy images, has been widely studied in the field of fluorescence microscopy imaging. Deep learning-based SR models have greatly promoted the development of conventional SR microscopy approaches by reconstructing HR microscopy images from their low-resolution (LR) versions.

We first determined the potential of our UniFMIR approach to deal with the SR problem (×2 upscaling) involving images with increasing structural complexity levels from the BioSR dataset[5] obtained via multimodal structured illumination microscopy (SIM) system, including clathrin-coated pits (CCPs), endoplasmic reticula (ERs), microtubules (MTs) and F-actin filaments. Our UniFMIR successfully inferred SR SIM images from wide-field (WF) images at a diffraction-limited scale with a high fluorescence level and revealed clear structural details. Compared with two deep learning-based fluorescence microscopy SR models (XTC[15] and DFCAN[5]) and a single-image super-resolution model (ENLCN[36]) for macroscale photographs, UniFMIR could correctly reconstruct most MTs without losing or merging them, even if the MTs were densely distributed and were close to each other. The baselines for different datasets were obtained by training a model of the same network structure as UniFMIR from scratch on the specific training dataset. For diverse subcellular structures, UniFMIR also restored hollow, ring-shaped CCPs and crisscrossing F-actin with high fidelity (Fig. 1b).

We also quantified the attained SR accuracy with the peak SNR (PSNR), structural similarity index measure (SSIM), normalized root mean square error (NRMSE), resolution estimate of a decorrelation analysis[37], Fourier ring correlation (FRC)[38], SQUIRREL analysis[39] and segmentation metrics (Fig. 1c, Supplementary Figs. 2–4 and 10). Higher PSNR/SSIM values and lower NRMSE values denote better SR when assessing SR SIM images in terms of their fluorescence intensities and structures, which signify images that are closer to the ground-truth (GT) SIM images.

Isotropic reconstruction

Volumetric fluorescence microscopy methods, such as three-dimensional (3D) SIM, are generally limited by the anisotropic spatial resolution, where the axial resolution of 300 nm is inferior to the lateral resolution. Such an anisotropy, which is caused by the inherent optical point spread function of microscopy or a low axial sampling rate, compromises the imaging quality of the volumes of interest. Therefore, isotropic reconstruction is also a common problem encountered in fluorescence microscopy when restoring isotropic image resolutions.

We applied our UniFMIR approach on anisotropic raw data (with up to tenfold lower axial resolutions) from volumetric mouse liver imaging[4] to predict isotropic axial slices and compared it with two

https://doi.org/10.1038/s41592-024-02244-3

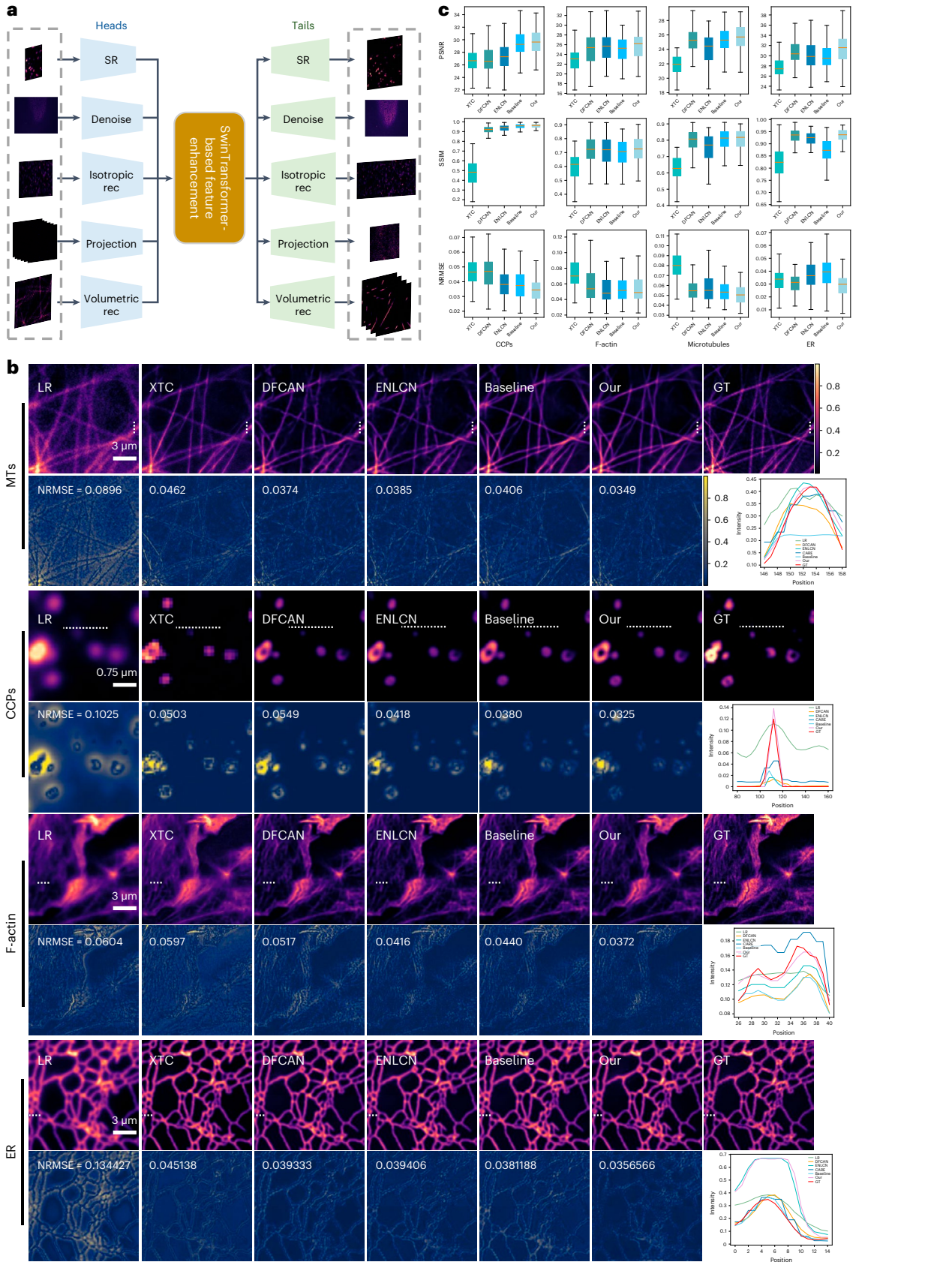


Fig. 1 | Applying the proposed UniFMIR approach to reconstruct SR SIM images from diffraction-limited WF images. a, Architecture of UniFMIR, which comprises multihead, multitail and Swin transformer-based feature enhancement modules. **b**, Shown are the LR inputs; the SR results obtained by the SOTA methods (XTC[15], DFCAN[5] and ENLCN[36]), baseline (same network structure as UniFMIR trained from scratch) and our fine-tuned UniFMIR approach; and the GT SIM images. The NRMSE↓ (lower is better) values are shown on the residual images under the SR results. **c**, PSNR↑/SSIM↑ (higher is better)/NRMSE↓ comparisons for ×2 upscaling on different datasets (MTs, CCPs, F-actin and ERs), $n = 100$. Scale bar, 0.75 μm for CCPs; 3 μm for other specimens.

https://doi.org/10.1038/s41592-024-02244-3

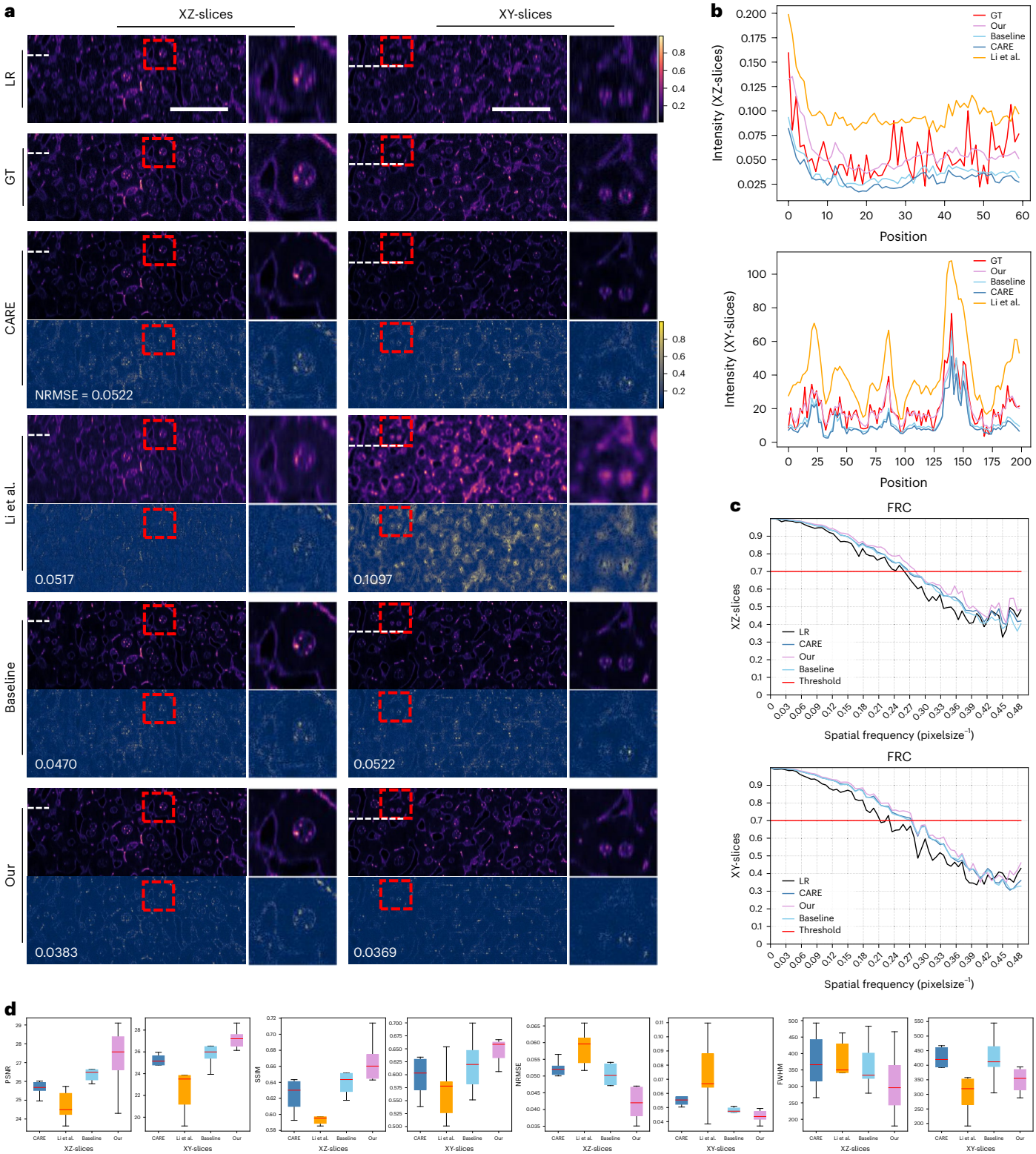


Fig. 2 | Applying UniFMIR to the isotropic reconstruction of 3D volumes. **a**, XZ and XY slices of anisotropic raw LR data (with a subsampling rate of 10); the GTs; and the reconstruction results of CARE[4], Li et al.[40], baseline (same network structure as UniFMIR trained from scratch) and our fine-tuned UniFMIR model. Magnified images of the regions of interest (red boxes) are displayed to the right of the corresponding images. The NRMSE is shown on each residual image. **b**, The line plots show the pixel intensities along the dashed lines for the images in **a**. **c**, FRC curves[38] for resolution estimation. **d**, Statistical comparison on the liver dataset in terms of PSNR↑/SSIM↑/NRMSE↓/FWHM↓ across $n = 7$ slices. Scale bar, 50 μm.

deep learning-based isotropic reconstruction models, CARE[4] and the 3D U-Net model proposed by Li et al.[40]. The proposed UniFMIR method could offer near-isotropic imaging by enhancing the axial resolution, facilitating the subsequent quantification of the shapes and volumes of biological samples (Fig. 2a and Supplementary Fig. 5). Our UniFMIR method yielded isotropic reconstruction results with more accurate

https://doi.org/10.1038/s41592-024-02244-3

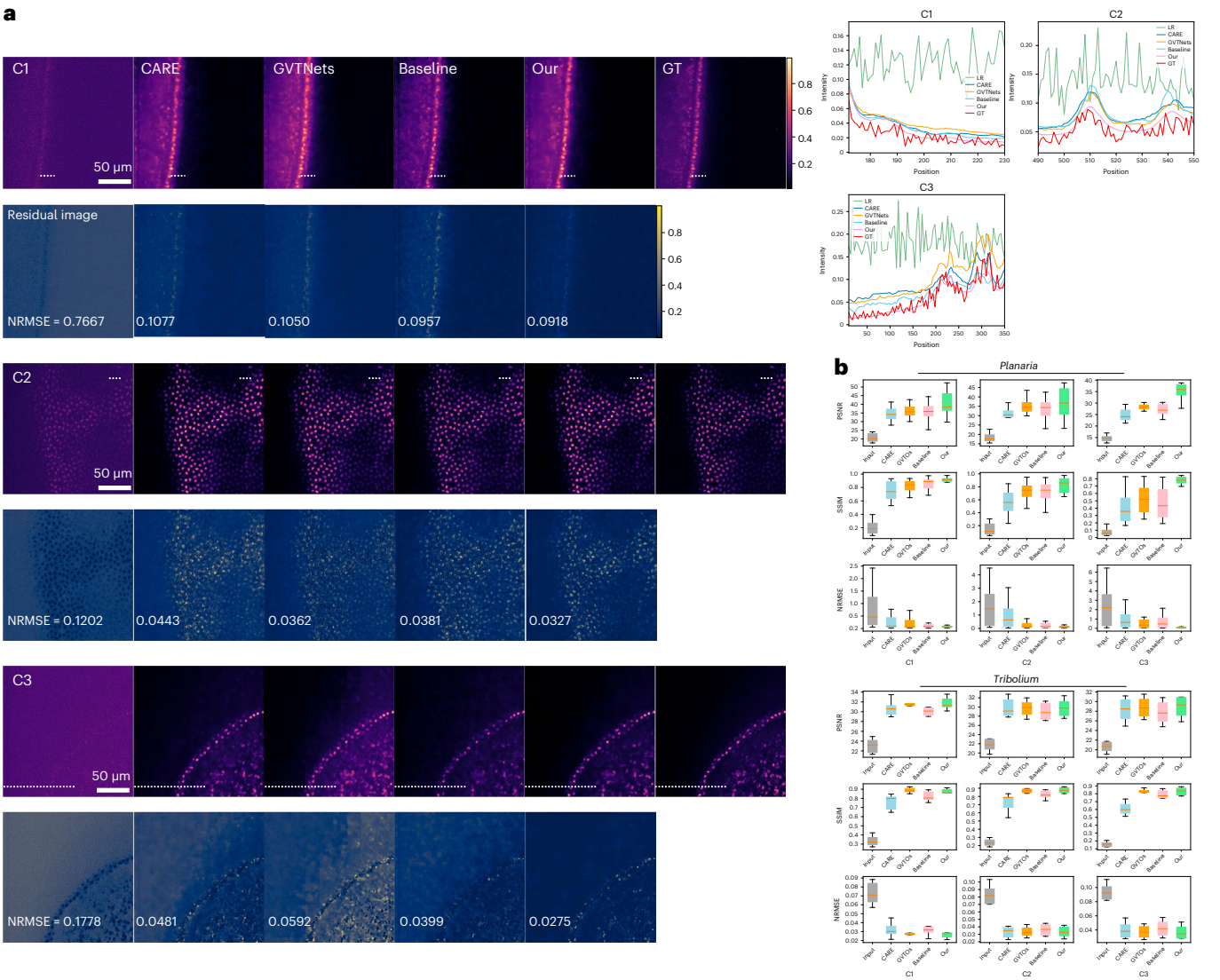


Fig. 3 | Applying UniFMIR to content-aware 3D image denoising. a, Visual results of a 3D image denoising task conducted on flatworm (*S. mediterranea*). Comparison among CARE[4], GVTNets[7], baseline (same network structure as UniFMIR trained from scratch) and our fine-tuned UniFMIR model. The line plots show the pixel intensities along the dashed lines for the images. **b**, Box plots of the PSNR↑/SSIM↑/NRMSE↓ results obtained on the *Planaria* ($n = 20$) and *Tribolium* ($n = 6$) datasets[4] under three imaging conditions (C1–C3). Scale bar, 50 μm.

pixel distributions, and the pixel intensities along the columns of our axial outputs were closer to those of the GTs (Fig. 2b). The same conclusion could also be drawn from the average PSNR/SSIM/NRMSE results obtained on data from the liver dataset (Fig. 2c).

3D image denoising

High-SNR fluorescence microscopy imaging always requires high laser power or long exposure times, which are accompanied by bleaching, phototoxicity and other side effects that are detrimental to the sample. Deep learning-based denoising methods[4,6,7] can computationally restore acquired low-SNR fluorescence microscopy images by leveraging the available knowledge about the data at hand.

We further benchmarked the performance of our UniFMIR approach in a live-cell image denoising task conducted on the *Planaria* and *Tribolium* datasets[4], each of which contains well-registered high-/low-SNR 3D images captured by a spinning-disk confocal microscope and a multiphoton laser-scanning microscope, for training and testing. Compared with two U-Net-based denoising models, CARE[4] and GVTNets[7], our UniFMIR model considerably suppressed the noise of the low-SNR fluorescence microscopy images under different laser powers/exposure time (C1–C3) and clearly depicted the planarian *Schmidtea mediterranea* and *Tribolium castaneum* volumes with labeled nuclei, helping to observe embryonic development (Fig. 3a and Supplementary Figs. 6 and 7). UniFMIR resulted in higher statistical accuracy (PSNR/SSIM/NRMSE values) between the denoised images and the well-registered high-SNR images (GTs; Fig. 3b).

Surface projection

To better analyze and study the cell behavior in developing epithelia of the *Drosophila melanogaster* fruit fly, surface projection helps project a 3D volume into a two-dimensional (2D) surface image. The current deep learning models (CARE[4] and GVTNets[7]) formulate this image restoration problem as two subproblems, 3D-to-2D surface projection and 2D image denoising, and use two task-specific networks, following the same encoder–decoder framework as that of U-Net, to solve them.

We further examined UniFMIR in a more complex composite fluorescence microscopy-based image restoration task and adopted the public Flywing dataset[4], which contains 3D–2D image pairs for training

https://doi.org/10.1038/s41592-024-02244-3

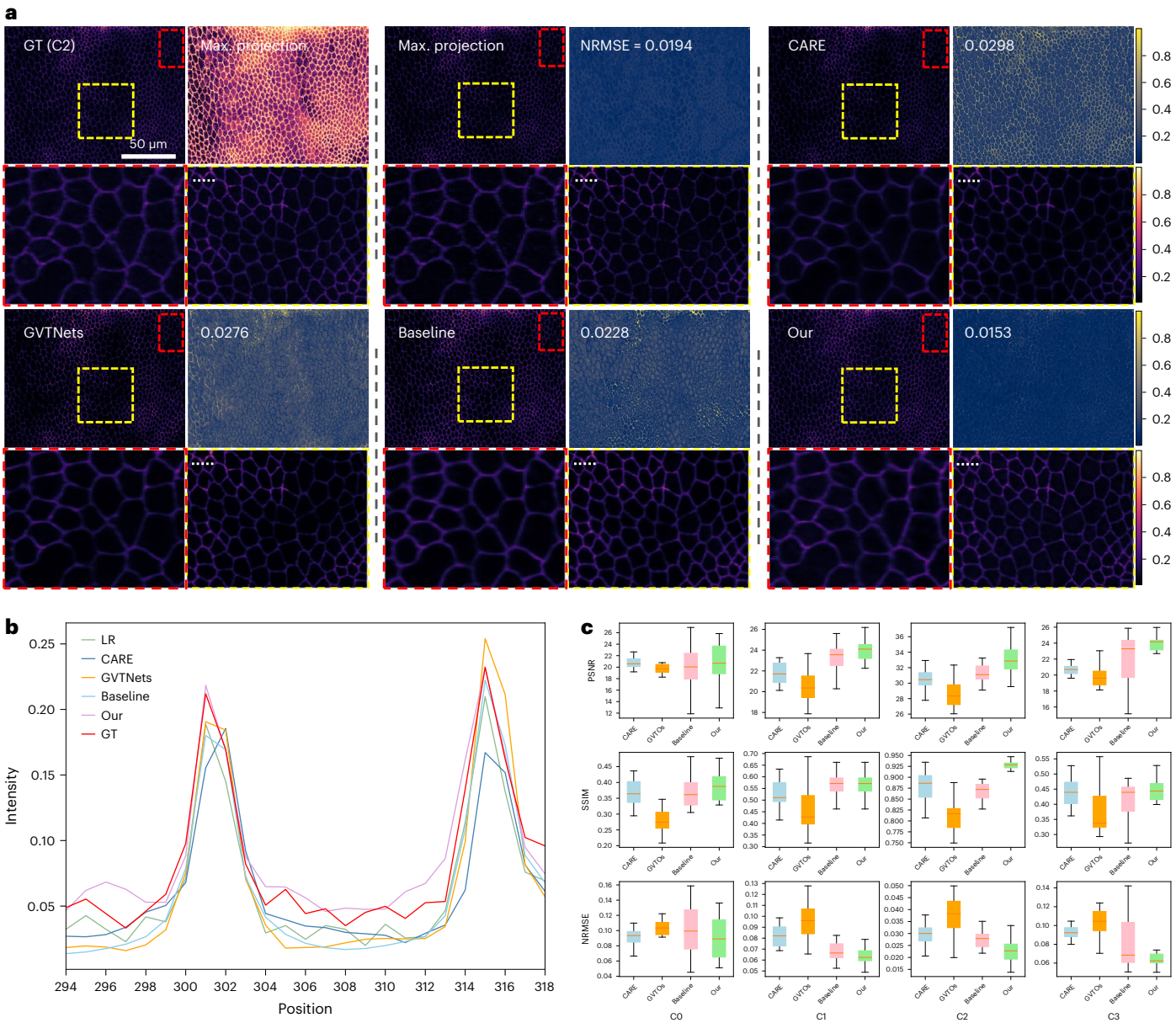


Fig. 4 | Applying UniFMIR to joint surface projection. a, First row, columns from left to right: high-SNR 2D image (GT), projection results of maximum projection and CARE[4]. Second row, columns from left to right: reconstruction results of GVTNets[7], baseline (same network structure as UniFMIR trained from scratch) and our fine-tuned UniFMIR model. The residual images are shown on the right, and the magnified regions (in the red and yellow dashed boxes) are shown below the projection results under the C2 condition. **b**, The line plot shows the pixel intensities along the dashed lines in the images in **a**. **c**, Box plots of the PSNR↑/SSIM↑/NRMSE↓, $n = 26$ results obtained on the Flywing dataset[4] under four imaging conditions (C0–C3). Scale bar, 50 μm.

and testing. Similarly to the *Planaria* and *Tribolium* datasets, the Flywing dataset also covers different laser power conditions (C1–C3). To project each 3D volume onto a 2D plane, we adopted a U-Net-shaped head for UniFMIR, which achieved promising results with end-to-end network calculations on the 3D *D. melanogaster* Flywing imaging task and yielded higher reconstruction accuracy in terms of the PSNR/SSIM/NRMSE metrics; meanwhile, the pixel intensities were closer to those of the GT (Fig. 4 and Supplementary Fig. 8).

Volumetric reconstruction

The volumetric reconstruction of light-field microscopy images, permitting the acquisition of artifact-free 3D image sequences with uniform spatial resolutions from 2D information, is significant for instantaneously imaging fast biological processes. As a demonstration, we verified the volumetric reconstruction ability of UniFMIR on the data provided by VCD-Net[8]. Each view of a reconstructed 3D volume can identify the motion trajectory of the imaging object (Fig. 5 and Supplementary Fig. 9), which is beneficial for revealing the underlying mechanisms of many complicated live-cell dynamics involving various subcellular structures. Since no GT was available for calculating more quantitative metrics, such as the PSNR/SSIM/NRMSE, to quantify the accuracy of the volumetric reconstruction results, we adopted decorrelation analysis[37] to measure the nanometer resolution of each reconstructed image sequence (Fig. 5 and Supplementary Fig. 9).

Generalization ability analysis

To demonstrate the generalization ability of our pretrained UniFMIR approach, we validated its image restoration performance on unseen data from DeepBacs[41] for bacterial image analysis purposes, including two denoising datasets (*E. coli*_H-NS-mScarlet-I, *E. coli*_MreB) and two SR

https://doi.org/10.1038/s41592-024-02244-3

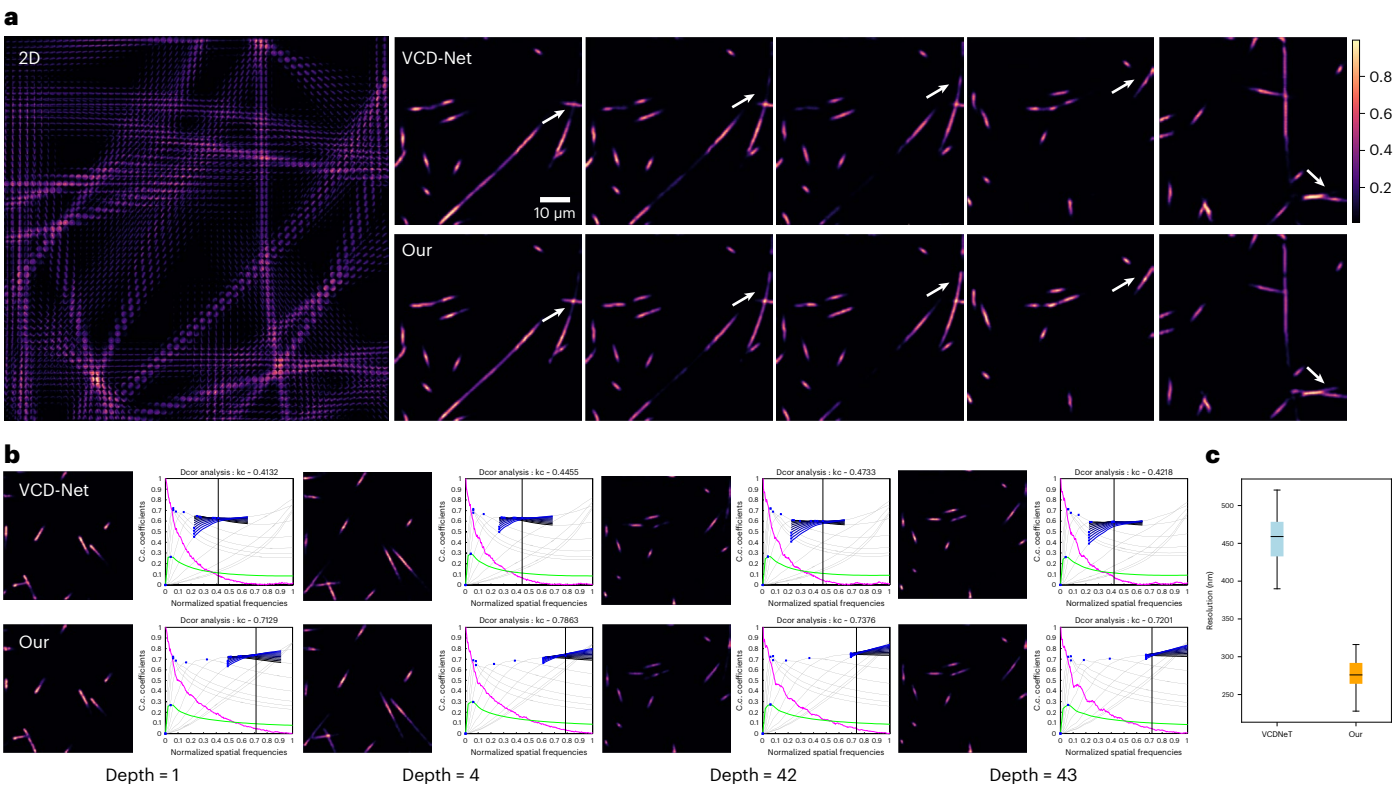


Fig. 5 | Applying UniFMIR to volumetric reconstruction. a, The image sequences of the artifact-free 3D volume reconstructed from a 2D input. **b**, More visual comparisons between the volumetric reconstruction results of our UniFMIR model and VCD-Net[8]. The resolution evaluation of the images (depth = 1, 4, 42, 43) in the reconstructed 3D volume performed with decorrelation analysis. **c**, Box plots of the resolutions (nm) of $n = 61$ image sequences. Scale bar, 10 μm.

datasets (*Escherichia coli*, *Staphylococcus aureus*). We fine-tuned UniFMIR on the new bacterial microscopy images and then conducted denoising and SR experiments. In addition to the denoising and SR models developed on DeepBacs, we also compared our results with those of a baseline model trained from scratch, which had the same structure as UniFMIR, and the pretrained UniFMIR model without fine-tuning. Our UniFMIR method restored clear prokaryote structures to enhance the low-phototoxicity live-cell microscopy data and predict accurate mappings of biological target shapes, obtaining higher PSNR/SSIM values (Fig. 6). Compared with the model trained from scratch, our UniFMIR approach achieved better performance on new datasets (Supplementary Fig. 13).

We also analyzed whether UniFMIR could be generalized to other SR modalities in addition to the SIM images used in the pretraining stage. First, we adopted single-molecule localization microscopy data from the Shareloc platform[42] and applied our model to direct stochastic optical reconstruction microscopy (dSTORM) images of MTs stained with Alexa Fluor 647 in U2OS cells incubated with nocodazole. Since the input WF images and GT were not well matched, we fine-tuned our UniFMIR model with the contextual bilateral (CoBi) loss[43] instead of the L1 or L2 loss, which requires pixel-wise alignment between the input and GT. We compared a U-Net-based SR single-molecule microscopy model (DeepSTORM[44]), a baseline model and a pretrained UniFMIR model without fine-tuning. In addition to the pretrained UniFMIR model, all competing models were trained with the CoBi loss on the same training data as those used by our method. Our UniFMIR model could restore accurate structures that were similar to the GTs, and the other models failed to learn mappings between the unaligned images (Extended Data Fig. 3).

Discussion

The focus of this paper was to present a UniFMIR solution for maximizing the potential of deep learning-based methods and circumventing the limitations exhibited by the existing fluorescence microscopy-based image restoration deep models. Here, we outlined how recent advances in foundation model research enable the development of FMIR. Inspired by the success of pretrained large-scale models in artificial intelligence, we developed a unified foundation model for fluorescence microscopy-based image restoration, facilitating high-quality image restoration in different fluorescence microscopy-based image restoration tasks with various imaging modalities by extending the strong transfer capabilities of large-scale pretrained models to fluorescence microscopy-based image restoration. We also collected 14 public fluorescence microscopy-based image restoration datasets with 196,418 training pairs covering various biological samples, microscopes and degradation conditions. The UniFMIR model pretrained on the collected data could be easily applied to different tasks and new image distributions through efficient fine-tuning, transferring the knowledge of a foundation model to a specific one.

The experimental results obtained in different fluorescence microscopy-based image restoration tasks suggested the excellent performance of UniFMIR in restoring high-fidelity microscopy images. The HR results obtained for the SR and isotropic resolution reconstruction problems could clearly resolve diffraction-limited image details to improve the resolutions of images by uncovering subtle biological structures (Figs. 1 and 2 and Supplementary Fig. 5). The denoising and projection results restored clean signals from the noisy inputs, achieving accurate reconstruction quality (Figs. 3 and 4 and Supplementary Figs. 6–8). The volume reconstruction results showed transient biological image dynamics with minimal artifacts (Fig. 5 and Supplementary Fig. 9). The generalization ability of the pretrained UniFMIR was also shown (Fig. 6 and Extended Data Fig. 3).

Key capabilities

We outline three key capabilities that distinguish UniFMIR from conventional FMIR models. (1) Higher restoration quality. UniFMIR

https://doi.org/10.1038/s41592-024-02244-3

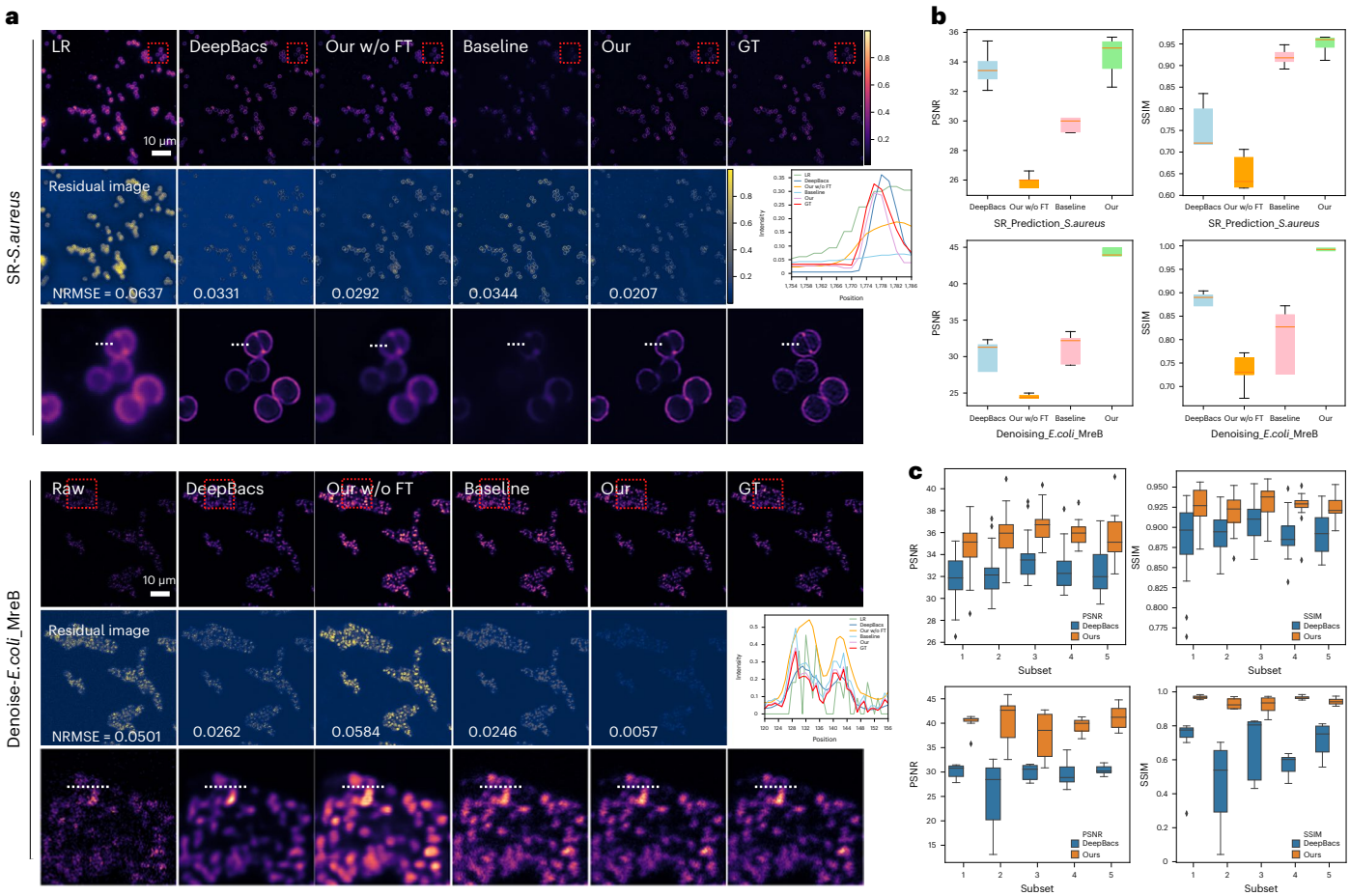


Fig. 6 | Generalization ability analysis conducted on unseen datasets[41]. **a**, SR and denoising results obtained on the *S.aureus*_MreB and *E.coli*_MreB datasets, respectively. Visual comparison among the outputs of the SOTA model (DeepBacs[41]), pretrained UniFMIR model without fine-tuning, baseline (same network structure as UniFMIR trained from scratch) and our fine-tuned UniFMIR model. The NRMSE is shown on each residual image. **b**, Box plots show the PSNR/SSIM results obtained on the test sets of the two datasets. **c**, *k*-fold validation ($k = 5$) on the *S.aureus*_MreB dataset (top) and the *E. coli*_MreB dataset (bottom). The PSNR/SSIM results obtained on the $n = 5$ test images of different subsets. Scale bar, 10 μm.

achieved the highest restoration precision and was often superior to the task-specific FMIR models. (2) Better generalization ability. UniFMIR displayed an impressive generalization ability and enabled efficient training processes on new FMIR datasets by transferring knowledge from the available datasets to the new data. (3) Unifying FMIR tasks. UniFMIR possessed applicability to handle multiple FMIR problems with one model and unified the restoration processes concerning different data modalities for different fluorescence microscopes. Our work also identified that publicly shared fluorescence images can be considered a tremendous resource that can be harnessed to develop foundation models for enhancing fluorescence microscopy images.

We demonstrated the ability of UniFMIR to achieve high precision with promising generalization performance. The key intuition behind this idea is that the foundation model, pretrained on more diverse data distributions, could learn more generalized representations[29] of different high-quality image modalities and biological structures. Existing works[32,33] have found that large-scale pretraining allows the training of deeper models with greater predictive potential and even promotes robustness by gaining a fundamental understanding of the knowledge in training data. As stated by Guo et al.[45], the foundation model may be effective in terms of acquiring informative global patterns that can improve the robustness of task-specific models.

Different fluorescence microscopy-based image restoration operations pushed the distribution spaces of the low-quality images to the distribution spaces of high-quality images. During pretraining, the UniFMIR gained a fundamental understanding of the high-quality images and regularized optimization direction for different tasks, unlike the task-specific models that learned a path from low quality to high quality. Thus, the pretraining knowledge benefits the fine-tuning toward diverse tasks, facilitating faster and better convergence and enabling the model to perform better than a model trained from scratch (Supplementary Fig. 15).

Limitations and future work

Despite the impressive results of UniFMIR, it is noteworthy that this paper is an exploratory and inspiring work that demonstrates the feasibility of foundation models in fluorescence microscopy-based image restoration. We hope that this work, accompanied by a brief review of the current deep learning-based fluorescence microscopy-based image restoration approaches (Supplementary Table 1), will provide new insights for more researchers. Looking to the future, much room remains for further improvement and evaluation to advance the frontiers of fluorescence microscopy-based image restoration. We discuss some limitations and future research directions as follows. First, a large model (>100 MB or even 1 TB) often requires a considerable number of training images (millions or even billions), a time-consuming pretraining process (weeks or even months) and costly calculation resources (graphics processing units, GPUs). To reduce the deployment cost and make UniFMIR more energy efficient, we optimized UniFMIR models by applying two model compression

methods (pruning and quantization). The UniFMIR model still takes a long calculation time, especially for those 3D image-related tasks. Therefore, faster inference and higher efficiency require further exploration. Second, we utilized existing public datasets for pretraining and found that the current data have not yet saturated the model's performance (Supplementary Fig. 14). Theodoris et al.[32] outlined that as the amount of publicly available data continues to expand, pretraining on larger-scale data may further enhance the model's performance, especially for tasks with increasingly limited task-specific data. Therefore, efforts are expected to contribute to a more diverse and larger dataset. In future work, we will continuously train the foundation model with new data to make the fluorescence microscopy-based image restoration foundation model stronger and share it with the community timely and freely.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-024-02244-3.

References

1. Preibisch, S. et al. Efficient bayesian-based multiview deconvolution. *Nat. Methods* **11**, 645–648 (2014).
2. Gustafsson, N. et al. Fast live-cell conventional fluorophore nanoscopy with ImageJ through super-resolution radial fluctuations. *Nat. Commun.* **7**, 12471 (2016).
3. Arigovindan, M. et al. High-resolution restoration of 3D structures from widefield images with extreme low signal-to-noise-ratio. *Proc. Natl Acad. Sci. USA* **110**, 17344–17349 (2013).
4. Weigert, M. et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090–1097 (2018).
5. Qiao, C. et al. Evaluation and development of deep neural networks for image super-resolution in optical microscopy. *Nat. Methods* **18**, 194–202 (2021).
6. Chen, J. et al. Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes. *Nat. Methods* **18**, 678–687 (2021).
7. Wang, Z., Xie, Y. & Ji, S. Global voxel transformer networks for augmented microscopy. *Nat. Mach. Intell.* **3**, 161–171 (2021).
8. Wang, Z. et al. Real-time volumetric reconstruction of biological dynamics with light-field microscopy and deep learning. *Nat. Methods* **18**, 551–556 (2021).
9. Li, X. et al. Reinforcing neuron extraction and spike inference in calcium imaging using deep self-supervised denoising. *Nat. Methods* **18**, 1395–1400 (2021).
10. Qiao, C. et al. Rationalized deep neural network for sustained super-resolution live imaging of rapid subcellular processes. *Nat. Biotechol.* **41**, 367–377 (2022).
11. Belthangady, C. & Royer, L. A. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nat. Methods* **16**, 1215–1225 (2019).
12. Wu, Y. & Shroff, H. Faster, sharper, and deeper: structured illumination microscopy for biological imaging. *Nat. Methods* **15**, 1011–1019 (2018).
13. Wu, Y. et al. Multiview confocal super-resolution microscopy. *Nature* **600**, 279–284 (2021).
14. Chen, R. et al. Single-frame deep-learning super-resolution microscopy for intracellular dynamics imaging. *Nat. Commun.* **14**, 2854 (2023).
15. Xu, Y. K. T. et al. Cross-modality supervised image restoration enables nanoscale tracking of synaptic plasticity in living mice. *Nat. Methods* **20**, 935–944 (2023).
16. Arigovindan, M. et al. High-resolution restoration of 3D structures from widefield images with extreme low signal-to-noise-ratio. *Proc. Natl Acad. Sci. USA* **110**, 17344–17349 (2013).
17. Bommasani, R. et al. On the opportunities and risks of foundation models. Preprint at https://arxiv.org/abs/2108.07258 (2021).
18. Fei, N. et al. Towards artificial general intelligence via a multimodal foundation model. *Nat. Commun.* **13**, 3094 (2022).
19. Zhang, Y. et al. DialoGPT: large-scale generative pre-training for conversational response generation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. 270–278 (2020).
20. Yang, Z. et al. Xlnet: generalized autoregressive pretraining for language understanding. In *Conference on Neural Information Processing Systems (NeurIPS)* (2019).
21. Dai, Z. et al. Coatnet: marrying convolution and attention for all data sizes. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2021).
22. Kirillov, A. et al. Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 4015–4026 (2023).
23. Achiam, J. et al. Gpt-4 technical report. Preprint at https://arxiv.org/abs/2303.08774 (2023).
24. Bao, F. et al. One transformer fits all distributions in multi-modal diffusion at scale. In *International Conference on Machine Learning (ICML)* (2023).
25. Bi, K. et al. Accurate medium-range global weather forecasting with 3D neural networks. *Nature* **619**, 533–538 (2023).
26. Singhal, K. et al. Large language models encode clinical knowledge. *Nature* **620**, 172–180 (2023).
27. Jiang, L. Y. et al. Health system-scale language models are all-purpose prediction engines. *Nature* **619**, 357–362 (2023).
28. Huang, Z. et al. A visual-language foundation model for pathology image analysis using medical twitter. *Nat. Methods* **29**, 2307–2316 (2023).
29. Zhou, Y. et al. A foundation model for generalizable disease detection from retinal images. *Nature* **622**, 156–163 (2023).
30. Moor, M. et al. Foundation models for generalist medical artificial intelligence. *Nature* **616**, 259–265 (2023).
31. Madani, A. et al. Large language models generate functional protein sequences across diverse families. *Nature Biotechnol.* **41**, 1099–1106 (2023).
32. Theodoris, C. V. et al. Transfer learning enables predictions in network biology. *Nature* **618**, 616–624 (2023).
33. Henighan, T. et al. Scaling laws for autoregressive generative modeling. Preprint at https://arxiv.org/abs/2010.14701 (2020).
34. Zamir, A. et al. Taskonomy: disentangling task transfer learning. In *Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI)*, 3712–3722 (2019).
35. Liu, Z. et al. Swin transformer: hierarchical vision transformer using shifted windows. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2021).
36. Xia, B. et al. Efficient non-local contrastive attention for image super-resolution. In *Association for the Advancement of Artificial Intelligence (AAAI)* (2022).
37. Descloux, A., Grubmayer, K. S. & Radenovic, A. Parameter-free image resolution estimation based on decorrelation analysis. *Nat. Methods* **16**, 918–924 (2019).
38. Nieuwenhuizen, R. et al. Measuring image resolution in optical nanoscopy. *Nat. Methods* **10**, 557–562 (2013).
39. Culley, S. et al. Quantitative mapping and minimization of super-resolution optical imaging artifacts. *Nat. Methods* **15**, 263–266 (2018).
40. Li, X. et al. Three-dimensional structured illumination microscopy with enhanced axial resolution. *Nat. Biotechnol.* **41**, 1307–1319 (2023).

41. Spahn, C. et al. DeepBacs for multi-task bacterial image analysis using open-source deep learning approaches. *Commun. Biol.* **5**, 688 (2022).
42. Ouyang, W. et al. ShareLoc—an open platform for sharing localization microscopy data. *Nat. Methods* **19**, 1331–1333 (2022).
43. Zhang, X. C. et al. Zoom to learn, learn to zoom. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019).
44. Nehme, E. et al. Deep-storm: super-resolution single-molecule microscopy by deep learning. *Optica* **5**, 458–464 (2018).
45. Guo, L. L. et al. EHR foundation models improve robustness in the presence of temporal distribution shift. *Sci. Rep.* **13**, 3767 (2023).

Methods

Data preparation process

To cover as many imaging modalities and fluorescence microscopy-based image restoration tasks as possible, we collected datasets from the literature (Supplementary Table 2) and grouped numerous datasets for different fluorescence microscopy-based image restoration tasks and imaging modalities. As these datasets vary significantly in terms of formats, domains and numerical ranges, we processed the images for convenient training and cross-dataset validation.

First, we wrote the input and GT images of existing datasets with different storage formats, including 'TIF', 'npz', 'png' and 'nii.gz', into an '.npz' file. In addition, we normalized the images to unify the numerical distributions of different datasets by following the data processing method in CARE[4]. Because the spatial sizes of the features in a deep neural network are fixed during training, we further cropped the training images into multiple patches with the same spatial size to facilitate simultaneous training on images from different datasets.

Network architectures

We designed a multihead and multitail network architecture for the UniFMIR model, which included three components, including multiple feature extraction modules, a Swin transformer-based feature enhancement module and multiple image reconstruction modules (Fig. 1a and Extended Data Fig. 1). More specifically, the multihead and multitail branches for different fluorescence microscopy-based image restoration tasks adopted different feature extraction and image reconstruction modules to extract task-specific shallow features and reconstruct images, respectively (Supplementary Note 3).

Different fluorescence microscopy-based image restoration calculations shared the same feature enhancement module. Inspired by SwinIR[46], a SOTA model for natural image restoration, the Swin transformer[35]-based feature enhancement module adopted several vision transformer-based blocks to enhance the feature representations and to restore the final features for high-quality image reconstruction. As shown in Extended Data Fig. 2, the feature enhancement module consisted of convolutional layers and a series of Swin transformer blocks, each of which included several Swin transformer layers, a convolutional layer and a residual connection. The Swin transformer layer was composed of layer normalization operations, a multihead self-attention mechanism and a multilayer perceptron. In the multihead self-attention mechanism, the input features f_{in} were first divided into multiple small patches with a moving window operation, and then the self-attention in each patch was calculated with the function in equation (1).

$$\begin{aligned} Q &= Conv_Q(f_{\text{in}}), K = Conv_K(f_{\text{in}}), V = Conv_V(f_{\text{in}}), \\ f_{\text{out}} &= Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \end{aligned} \tag{1}$$

where Q, K and V represent the query, key and value, respectively, which were separately obtained by three convolutional layers. d_k is the dimensionality of K. SoftMax(·) normalized the similarity between Q and K, and the output feature f_{out} was obtained by multiplying V. The multilayer perceptron was composed of two fully connected layers and Gaussian-error linear unit activation.

Training losses

We used a combination of the $\mathcal{L}1$ and $\mathcal{L}2$ losses during the pretraining stage to exploit the robustness of the $\mathcal{L}1$ loss and the stability of the $\mathcal{L}2$ loss. During fine-tuning, $\mathcal{L}1$ is adopted to pursue a higher quantitative metric (PSNR). Suppose that $(x_i, y_i)_{i=1:N}$ denotes N pairs of input and GT training data and that f_θ denotes the UniFMIR model with a parameter θ (equation (2)).

$$\begin{aligned} \mathcal{L} &= 0.5 \times \mathcal{L}1 + 0.5 \times \mathcal{L}2, \\ \mathcal{L}1 &= \sum_{i=1}^{N} |y_i - f_\theta(x_i)|, \\ \mathcal{L}2 &= \sum_{i=1}^{N} (y_i - f_\theta(x_i))^2. \end{aligned} \tag{2}$$

To apply our UniFMIR model to unmatched data, in which the input WF images and the GT single-molecule localization microscopy data[42] were not well aligned in a pixel-wise manner, we fine-tuned our model with the CoBi loss[43], which improved its robustness to mild misalignment in the input–output image pairs (equation (3)).

$$\begin{aligned} CoBi(y, \tilde{y}) &= \frac{1}{N}\sum_{i}^{N} \min_{j=1,\ldots,M}(D_{p_i,q_j} + w_s D'_{p_i,q_j}), \\ D_{p_i,q_j} &= \text{Distance}(p_i, q_j), \\ D'_{p_i,q_j} &= ||(h_i, w_i) - (h_j, w_j)||_2, \end{aligned} \tag{3}$$

where $(\tilde{y}, y)$ denotes a pair of restored and GT images and $p_{i=1,\ldots,N}$, $q_{j=1,\ldots,M}$ are the features of X and Y extracted by the pretrained Visual Geometry Group 19 (VGG-19)[47] network, respectively. 'Distance(·)' denotes a distance function for calculating the cosine similarity between features p_i, q_j, (h_i, w_i), (h_j, w_j), are the spatial coordinates of features p_i, q_j and $w_s = 0.1$ denotes a weight that is flexible to the degree of misalignment.

Training details

The UniFMIR model was based on a PyTorch implementation and optimized by adaptive moment estimation (Adam)[48] with $\beta1 = 0.9$ and $\beta2 = 0.999$ for 500 epochs. The initial learning rate started at 5×10^{-5} and was halved after 200 epochs. All experiments were conducted on a machine with an Nvidia GeForce RTX 3090 GPU (with 24 GB of RAM).

In the pretraining stage, we set the batch size to 1 and the patch size to 64×64. We fed all training data to the model and optimized different head and tail branches for different tasks with the corresponding data. The middle feature enhancement branch was optimized using all training data. During the fine-tuning stage, we set the batch size/patch size to 4/128, 32/64, 32/64, 4/64 and 1/16 for the SR, isotropic reconstruction, denoising, projection and volume reconstruction tasks, respectively, to produce a better learning effect (Extended Data Table 1).

Evaluation metrics

To evaluate the quantitative accuracy of the fluorescence microscopy-based image restoration results, we adopted common image quality assessment metrics as follows (Supplementary Note 4).

The PSNR, NRMSE and SSIM[49] are proposed to measure the pixel-level and structure-level similarities between a restored image $\tilde{y}$ and a GT image y.

The resolution evaluation with decorrelation analysis[37], a comprehensive measurement of the resolution and SNR, was performed to estimate the highest frequency from the local maxima of the decorrelation functions rather than the theoretical resolution stated by Abbe[50]. NanoJ-SQUIRREL, an ImageJ-based analytical approach, was proposed to quantitatively assess SR quality by comparing diffraction-limited reference images and SR equivalents of the same acquisition volume.

FRC. FRC-based resolution measures[38,51] can estimate image resolution without a reference image. We adopt the public GitHub codes and microscope image processing library (MIPLIB), a Python-based software library, for FRC-based image resolution analysis of fluorescence microscopy images.

Giga floating-point operation (GFLOPs), a common computational complexity measure, refers to the number of FLOPs, including the addition, subtraction, multiplication and division of floating-point

numbers, that the model requires to process the input data. GFLOPs can reflect the processing power needed to execute the model and can vary depending on the number of convolutional layers, the types of operations performed within each layer, and the size of the input. The GFLOPs of a deep model are composed of operations contained in all convolution layers. The number of calculations (FLOPs) of a convolutional layer with a kernel size of $K \times K$ is $2K^2 - 1$, and the total number of FLOPs for conducting a convolutional calculation on an input with a size of $H \times W$ can be calculated according to equation (4):

$$\mathrm{FLOP} = \left(\frac{H - K + P}{S} + 1\right) \times \left(\frac{W - K + P}{S} + 1\right), \quad (4)$$

where P and S denote the pooling and stride parameters of a convolution, respectively.

BOP (bit operations)[52], which stands for number of bits times FLOPs, is considered to quantify the computational complexity of a deep model on a GPU that supports 32-bit, 16-bit or lower arithmetic. Since GFLOPs cannot well measure the computational complexity of low-precision and high-precision networks composed of integer or float operations, we also calculated the BOPs by the following function in equation (5):

$$\mathrm{BOPs} = \mathrm{FLOP} \times bw \times ba, \quad (5)$$

where bw and ba denote the weight and activation bit-width and are set to 32 and 16 for the 32-bit and 16-bit models, respectively.

Image restoration and segmentation. To explore whether performing image restoration with UniFMIR could improve the downstream image analysis and segmentation tasks in live-cell imaging, we applied a common segmentation pipeline (trainable Weka segmentation[53]) to the raw and restored images in the CCP, ER, *Tribolium* and Flywing datasets. The resulting UniFMIR model improved the segmentation effect by performing denoising and increasing the image resolution (Supplementary Note 5 and Supplementary Figs. 10–12).

Optimization of memory and complexity

To make UniFMIR more energy efficient, we produced optimized models by adopting two model compression methods (pruning and quantization). First, we adopted structure pruning to cut some useless branches out of the original model. Specifically, we removed the redundant head and tail branches and only kept the head and tail branches for fine-tuning a task-specific model to reduce the number of redundant parameters, resulting in 'prune-UniFMIR'. Inspired by Jacob et al.[54], we also conducted model quantization, converting floating-point weights (float32) and activation values to low-precision numbers, to reduce the storage requirements and calculation time of the model; for this task, we adopted float16 quantization (Extended Data Tables 2–6).

Pixel size correction

As the features learned by deep models depend on the scale and resolution of the training data, the performance of the well-trained models is influenced by the pixel-wise scale of input images. Given input images with varying pixel sizes, to make the outputs of the UniFMIR consistent, we equipped the UniFMIR software platform with a pixel size correction option for the input images, enabling an automatic pixel size calibration feature. Specifically, the input image is resized into different scales and input into the model, and then the multiscale outputs of the model are fused to obtain the final result.

Competing methods

All competing models, accompanied by their quantitative results, are listed in Extended Data Tables 3–6. To conduct a fair comparison, we downloaded the codes and the saved model checkpoints of all competing approaches from their GitHub repositories (ENLCN, CARE, DFCAN, GVTNets, VCD-Net and XTC). We retrained the ENLCN and XTC methods on the BioSR dataset[5]. Because DFCAN, CARE and VCD-Net do not provide well-trained models, we also retrained these models using their codes on the datasets used in our experiments for different tasks. The results of GVTNets were obtained by directly using their public models and codes. All experiments were conducted on the same machine with Nvidia GeForce RTX 3090 GPU (24 GB of RAM).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All training and testing data involved in the experiments come from existing literature and can be downloaded from the corresponding links provided in Supplementary Table 2 or via Zenodo at https://doi.org/10.5281/zenodo.8401470 (ref. 55).

Code availability

The PyTorch code of our UniFMIR, together with trained models, as well as some example images for inference are publicly available at https://github.com/cxm12/UNiFMIR (https://doi.org/10.5281/zenodo.10117581)[56]. Furthermore, We also provide a live demo for UniFMIR at http://unifmir.fdudml.cn/. Users can also access the colab at https://colab.research.google.com/github/cxm12/UNiFMIR/blob/main/UniFMIR.ipynb or use the steps in our GitHub documentation to run the demo locally. This newly built interactive software platform facilitates users to freely and easily use the pretrained foundation model. It also makes it easy for us to continuously train the foundation model with new data and share it with the community. Finally, we shared all models on BioImage.IO at https://bioimage.io/#/. Data are available via Zenodo at https://doi.org/10.5281/zenodo.10577218, https://doi.org/10.5281/zenodo.10579778, https://doi.org/10.5281/zenodo.10579822, https://doi.org/10.5281/zenodo.10595428, https://doi.org/10.5281/zenodo.10595460, https://doi.org/10.5281/zenodo.8420081 and https://doi.org/10.5281/zenodo.8420100 (refs. 57–63). We used the Pycharm software for code development.

References

46. Liang, J. et al. Swinir: image restoration using swin transformer. In *IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, 1833–1844 (2021).
47. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. In *International Conference on Machine Learning (ICLR)* (2015).
48. Kingma, D. & Ba, J. Adam: a method for stochastic optimization. Preprint at https://arxiv.org/abs/1412.6980 (2014).
49. Wang, Z. et al. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).
50. Abbe, E. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv. f. Mikrosk. Anatomie* **9**, 413–418 (1873).
51. Koho, S. et al. Fourier ring correlation simplifies image restoration in fluorescence microscopy. *Nat. Commun.* **10**, 3103 (2019).
52. Baskin, C. et al. UNIQ: uniform noise injection for non-uniform quantization of neural networks. *ACM Transactions on Computer Systems (TOCS)*, **37** (1–4), 1–15 (2021).
53. Arganda, C. et al. Trainable weka segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics* **33**, 2424–2426 (2017).
54. Jacob, B. et al. Quantization and training of neural networks for efficient integer-arithmetic-only inference. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2704–2713 (2018).

55. Ma, C., Tan, W., He, R. & Yan, B. UniFMIR: pre-training a foundation model for universal fluorescence microscopy image restoration (2023.10.03). *Zenodo* https://doi.org/10.5281/zenodo.8401470 (2023).
56. Ma, C., Tan, W., He, R., & Yan, B. UniFMIR: pre-training a foundation model for universal fluorescence microscopy image restoration (version 2023.11.13). *Zenodo* https://doi.org/10.5281/zenodo.10117581 (2023).
57. Ma, C., Tan, W., He, R. & Yan, B. UniFMIRProjectionOnFlyWing. *Zenodo* https://doi.org/10.5281/zenodo.10577218 (2024).
58. Ma, C., Tan, W., He, R. & Yan, B. UniFMIRDenoiseOnPlanaria. *Zenodo* https://doi.org/10.5281/zenodo.10579778 (2024).
59. Ma, C., Tan, W., He, R. & Yan, B. UniFMIRDenoiseOnTribolium. *Zenodo* https://doi.org/10.5281/zenodo.10579822 (2024).
60. Ma, C., Tan, W., He, R. & Yan, B. UniFMIRVolumetricReconstructionOnVCD. *Zenodo* https://doi.org/10.5281/zenodo.10595428 (2024).
61. Ma, C., Tan, W., He, R. & Yan, B. UniFMIRIsotropicReconstructionOnLiver. *Zenodo* https://doi.org/10.5281/zenodo.10595460 (2024) .
62. Ma, C., Tan, W., He, R. & Yan, B. UniFMIRSuperResolutionOnMicrotubules. *Zenodo* https://doi.org/10.5281/zenodo.8420081 (2023).
63. Ma, C., Tan, W., He, R. & Yan, B. UniFMIRSuperResolutionOnFactin. *Zenodo* https://doi.org/10.5281/zenodo.8420100 (2023).

Acknowledgements

We gratefully acknowledge support for this work provided by the National Natural Science Foundation of China (NSFC) (grant nos. U2001209 to B.Y. and 62372117 to W.T.) and the Natural Science Foundation of Shanghai (grant no. 21ZR1406600 to W.T.).

Author contributions

B.Y. and W.T. supervised the research. C.M. and W.T. conceived of the technique. C.M. implemented the algorithm. C.M. and W.T. designed the validation experiments. C.M. trained the network and performed the validation experiments. R.H. implemented the interactive software platform and organized the codes and models. All authors had access to the study and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

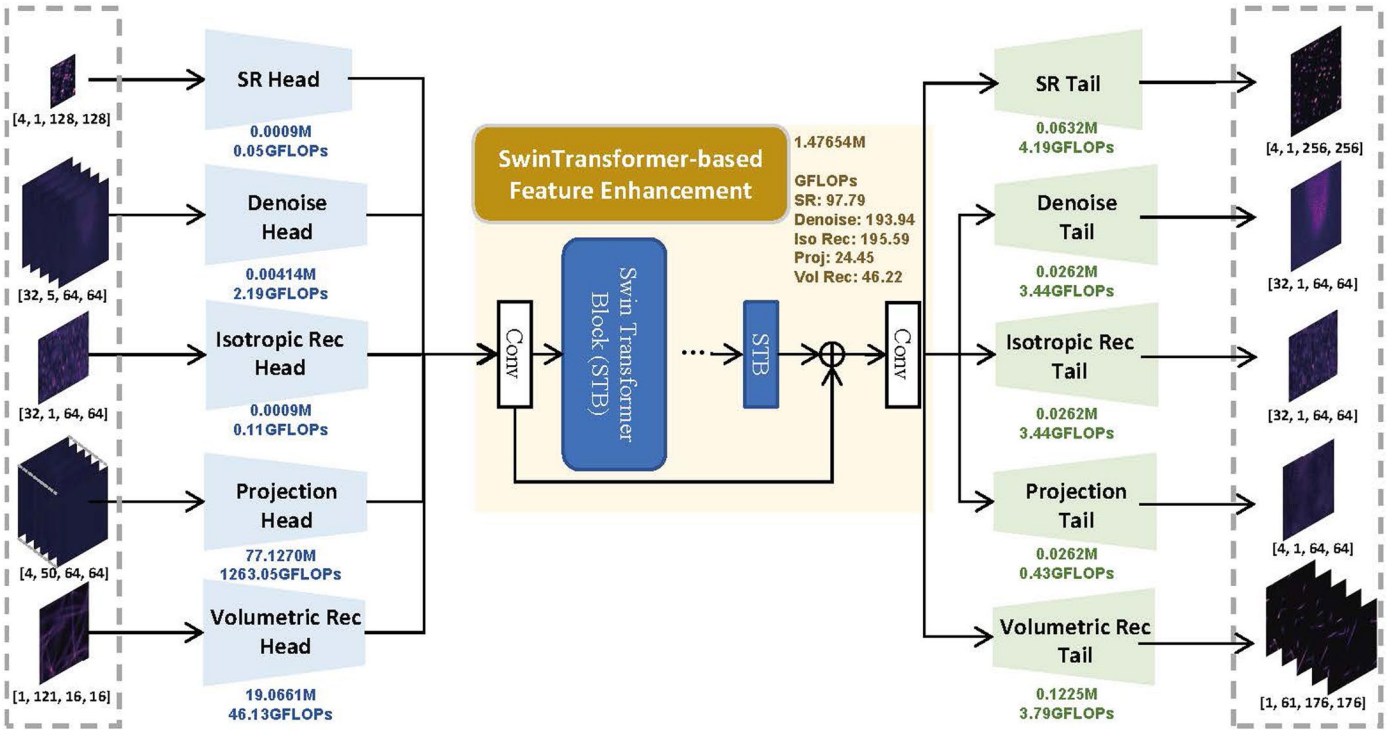
Extended data is available for this paper at https://doi.org/10.1038/s41592-024-02244-3.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41592-024-02244-3.

Correspondence and requests for materials should be addressed to Bo Yan.

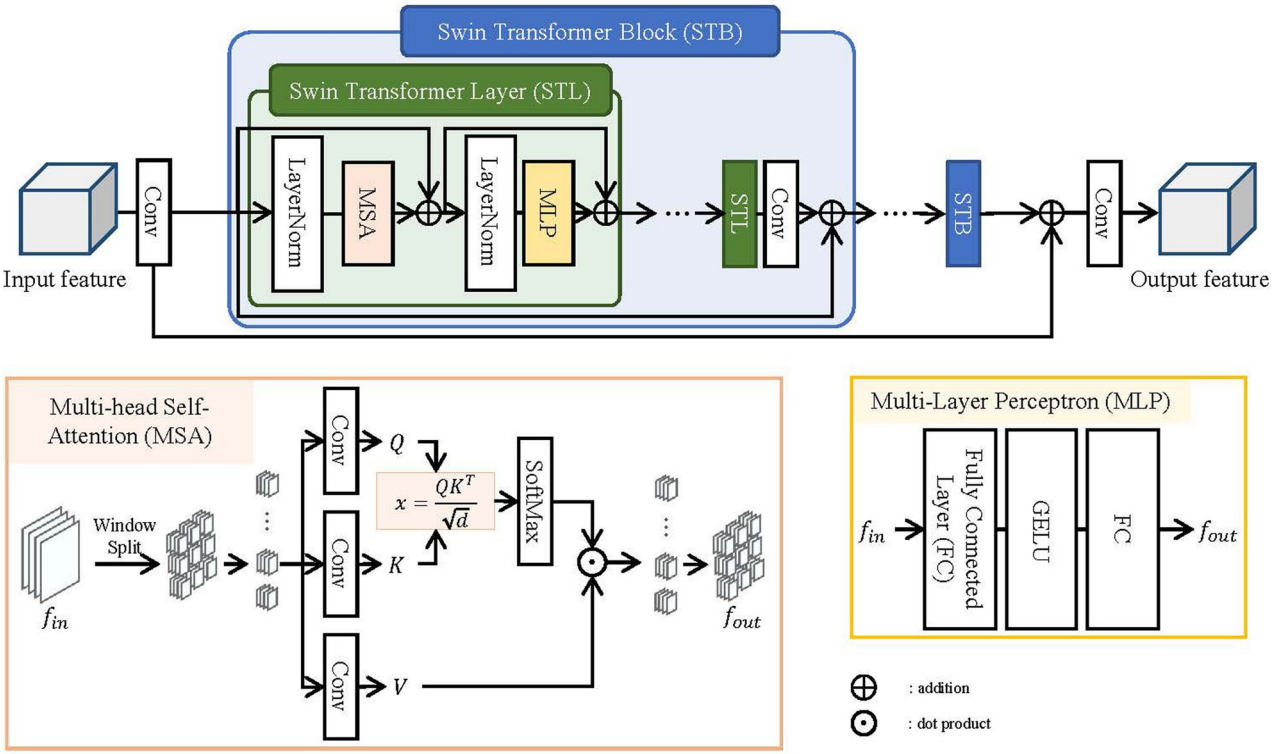
Peer review information *Nature Methods* thanks Ricardo Henriques and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Rita Strack, in collaboration with the *Nature Methods* team.

https://doi.org/10.1038/s41592-024-02244-3



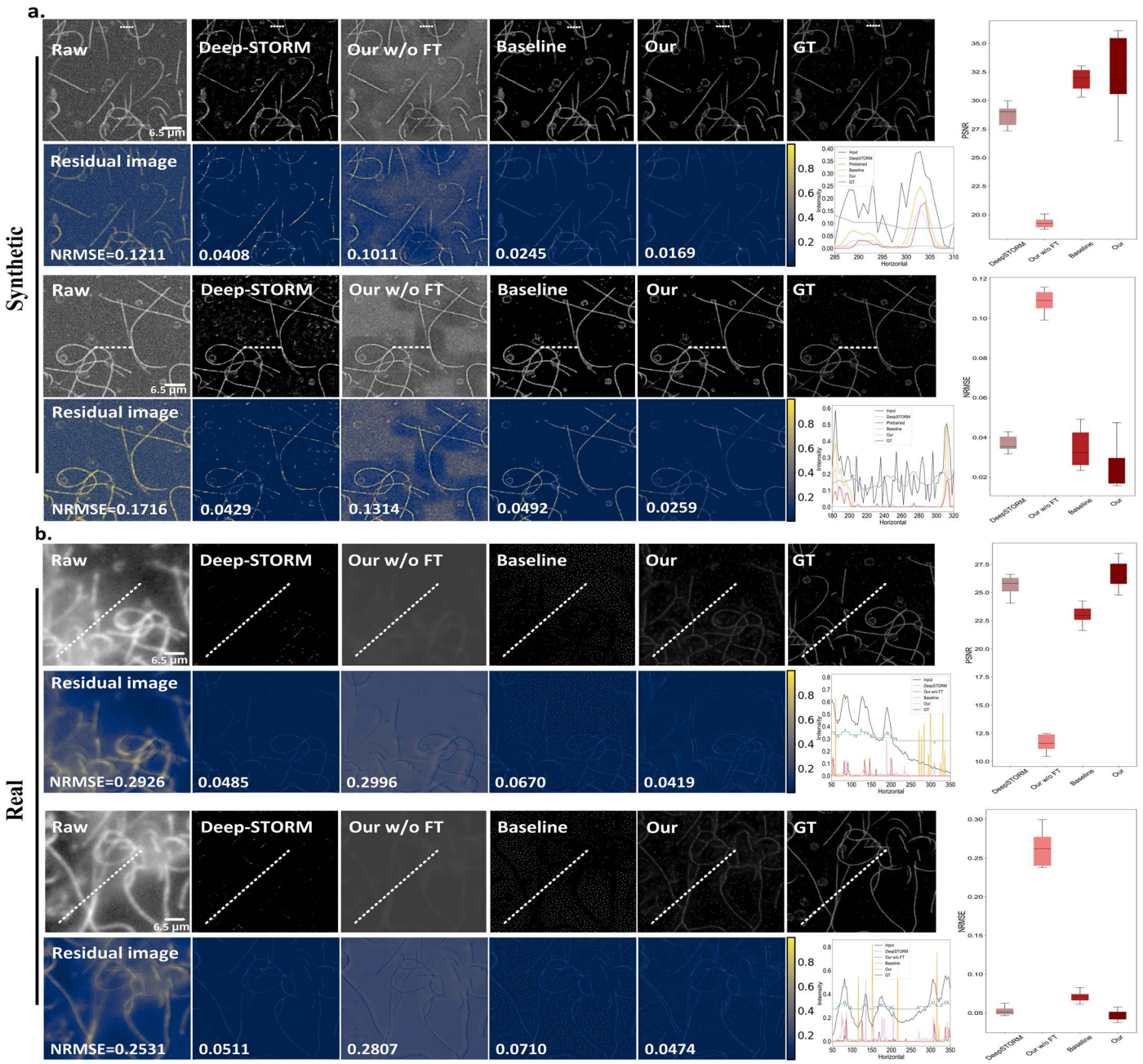
Extended Data Fig. 1 | Overall architecture of the UniFMIR. The proposed UniFMIR approach is composed of three submodules: a multihead module, a Swin transformer-based feature enhancement module, and a multitail module. The numbers of parameters (M) and calculations (GFLOPs) required for the head, feature enhancement and tail modules for different tasks are marked below the structures of the respective modules. The input sizes and output sizes of training batches for different tasks are also marked below the images.

https://doi.org/10.1038/s41592-024-02244-3



Extended Data Fig. 2 | Network architecture of the Swin transformer-based feature enhancement module[46]**.** The feature enhancement module consists of convolutional layers and a series of Swin transformer blocks (STB), each of which includes several Swin transformer layers (STL), a convolutional layer and a residual connection. The STL is composed of layer normalization operations, a multihead self-attention (MSA) mechanism and a multilayer perceptron (MLP). In the MSA mechanism, the input features are first divided into multiple small patches with a moving window operation, and then the self-attention in each patch is calculated to output features f_{out}. The MLP is composed of two fully connected layers (FCs) and Gaussian-error linear unit (GELU) activation.

https://doi.org/10.1038/s41592-024-02244-3



Extended Data Fig. 3 | Generalization ability analysis of super-resolution on unseen modality of single-molecule localization microscopy data from the Shareloc platform[52]**.** a, SR results obtained by the SOTA model (DeepSTORM[54]), the pretrained UniFMIR model without fine-tuning, Baseline (same network structure as UniFMIR trained from scratch), and our fine-tuned UniFMIR model. The GT dSTORM images of microtubules stained with Alexa 647 in U2OS cells incubated with nocodazole and the input synthesized LR images are also shown. The PSNR/NRMSE results of the SR outputs obtained on n = 16 synthetic inputs are shown on the right. b, SR results obtained on the real-world wide-field images. The NRMSE values are depicted on the residual images under different SR results and the raw input images. The PSNR/NRMSE results on n = 9 real-world inputs are shown on the right. Box-plot elements are defined as follows: center line (median); box limits (upper and lower quartiles); whiskers (1.5x interquartile range). The line plots show the pixel intensities along the dashed lines in the corresponding images. Scale bar: 6.5 μm.

https://doi.org/10.1038/s41592-024-02244-3

Extended Data Table 1 | Comparison between the training costs of the SOTA models and that of our model on different tasks

	Input Size	Model	GFLOPs	Param (M)	Model Size (MB)	Time	
						CPU (s)	GPU (ms)
SR	1, 1, 128, 128	ENLCN	12.9793	0.1980	0.82	53.82	128.03
	4, 1, 128, 128	DFCAN	125.8166	1.9308	8.10	15.12	80.34
		XTC	3.4163	0.3326	1.38	0.35	10.13
		UniFMIR	102.0431	100.4071	402.16	37.25	419.02
		prune-UniFMIR		1.5407	5.88		
Denoise	32, 1, 16, 64, 64	CARE	328.3071	0.9968	4.03	16.67	84.62
		GVTNets	1018.6811	1.4915	5.70	38.40	156.52
	32, 5, 64, 64	FMIR	199.5658	100.4071	402.16	82.98	142.30
		prune-UniFMIR		1.5069	5.75		
Isotropic Reconstruction	32, 1, 1, 64, 64	CARE	20.4976	0.9968	3.84	0.46	3.30
	32, 1, 64, 64	UniFMIR	199.1411	100.4071	402.16	82.02	148.69
		prune-UniFMIR		1.5037	5.74		
Projection	4, 50, 64, 64, 1	CARE	2.2755	1.3798	5.57	0.14	5.54
		GVTNets	3.7714	0.4932	1.91	0.82	18.65
	4, 50, 64, 64	UniFMIR	1287.9276	100.4071	402.16	224.28	273.04
		prune-UniFMIR		78.6000	325.17		
Volume Reconstruction	1, 121, 16, 16	VCD-Net	44.6016	19.0166	75.30	4.47	32.71
		UniFMIR	96.1466	100.4071	402.16	20.40	905.18
		prune-UniFMIR		23.1294	92.95		

The number of parameters (M), the model size (MB), the number of calculations (GFLOPs), and the running time on a CPU/GPU required by each model in each training iteration are shown.

https://doi.org/10.1038/s41592-024-02244-3

Extended Data Table 2 | Comparison between the prediction costs of the SOTA models and that of our model on different tasks

	Image Size	Patch size	Model	Param (M)	Model Size (MB)	GFLOPs	BOPs (T)	CPU time (s)		GPU time (ms)	
								Patch	Full Image	Patch	Full Image
SR	1,128,128	1, 1, 128, 128	DFCAN	1.9308	8.10	31.5098	32.2655	2.16	2.16	17.17	17.17
			XTC	0.3326	1.38	0.8568	0.8774	0.05	0.05	1.95	1.95
			ENLCN	0.1980	0.82	3.2580	3.3362	53.82	53.82	128.03	128.03
			prune-UniFMIR	1.5407	5.88	44.0695	45.1272	4.19	4.19	197.81	197.81
			prune-UniFMIR (float16)		2.94		11.2818	2.97	2.97	140.29	140.29
Denoise	41,626,1015	1, 1, 41,160,128	CARE	0.9968	4.03	4.2064e3	4.3073e3	7.49	239.68	36.42	1.17e3
			GVTNets	1.4915	5.70	1.3052e4	1.3364e4	17.13	548.16	66.68	2.13e3
		41, 5, 160, 128	**prune**-UniFMIR	1.5069	5.75	4.0003e4	4.0962e4	308.00	9856.00	1561.59	5.00e4
			prune-UniFMIR (float16)		2.88		1.0241e4	198.42	6349.66	996.30	3.19e4
Isotropic Reconstruction	301,752,752	1, 1, 188, 188, 301	CARE	0.9968	3.84	2.6559e4	2.7196e4	96.71	1547.36	508.45	8135.2
		301, 1, 188, 188	**prune**-UniFMIR	1.5037	5.74	2.6081e5	2.6707e5	3.44e3	5.51e4	1.79e4	2.87e5
			prune-UniFMIR (float16)		2.87		6.6767e4	2.45e3	3.92e4	1.28e4	2.04e5
Projection	50,520,692	1, 1, 50, 144, 176	CARE	1.3798	5.57	56.4251	57.7793	0.21	3.51	4.51	72.16
			GVTNets	0.4932	1.91	93.5141	95.7584	1.24	19.84	19.57	313.12
		1, 50, 144, 176	**prune**-UniFMIR	78.6000	325.17	2.8195e4	2.8872e4	121.91	1.95e3	1.01e3	1.61e4
			prune-UniFMIR (float16)		162.59		7.2179e3	43.33	693.28	358.13	5.73e3
Volume Reconstruction	121,59,59	1, 121, 15, 15	VCD-Net	19.0166	75.30	625.7100	640.73	2.07	33.12	5.77	92.32
			prune-UniFMIR	23.1294	92.95	1521.2671	1557.78	23.26	372.16	358.70	5.74e3
			prune-UniFMIR (float16)		46.48		389.44	22.67	362.87	349.75	5.60e3

The number of parameters (M), the model size (MB), the number of calculations (GFLOPs/BOPs), and the running time on a CPU/GPU required by different models for a low-quality image input are shown.

https://doi.org/10.1038/s41592-024-02244-3

Extended Data Table 3 | Quantitative comparison among different models for the x2 SR task

Model	Input size	Param (M)	Model Size (MB)	GFLOPs	BOPs (T)	Time		Dataset			
						CPU (s)	GPU (ms)	CCP	MTs	F-actin	ER
ENLCN	1, 1, 128, 128	0.1980	0.82	3.2580	3.3362	53.82	128.03	27.337/0.9353	24.102/0.7571	25.438/0.7018	30.189/0.9093
XTC		0.3326	1.38	0.8568	0.8774	0.05	1.95	26.357/0.4652	21.875/0.6226	22.450/0.5786	27.989/0.8125
DFCAN		1.9308	8.10	31.5098	32.2655	2.16	17.17	26.865/0.9005	25.044/0.7882	25.125/0.7102	30.974/0.9277
prune-UniFMIR		1.5407	5.88	44.0695	45.1272	4.19	197.81	28.195/0.9447	25.231/0.7985	25.791/0.7041	31.318/0.9255
prune-UniFMIR (float16)			2.94		11.2818	2.97	140.29	28.929/0.9424	25.702/0.7985	25.788/0.7046	31.618/0.9255

The number of parameters (M), the model size (MB), the number of calculations (GFLOPs/BOPs), and the running time required on a CPU/GPU for a given LR image input are shown, as are the accuracies (PSNR/SSIM values) achieved on the BioSR dataset.

https://doi.org/10.1038/s41592-024-02244-3

Extended Data Table 4 | Quantitative comparison among different models for the isotropic reconstruction task

Model	Image Size	Input size	Param (M)	Model Size (MB)	GFLOPs	BOPs (T)	Time		Planaria			Tribolium		
							CPU (s)	GPU (ms)	C1	C2	C3	C1	C2	C3
CARE	41,626,1015	1, 1, 41,160,128	0.9968	4.03	4.2064e3	4.3073e3	239.68	1.17e3	32.344/0.6904	29.041/0.5537	24.869/0.4099	25.812/0.3902	24.705/0.3788	24.284/0.3527
GVTNets			1.4915	5.70	1.3052e4	1.3364e4	548.16	2.13e3	35.796/0.7890	34.264/0.7289	28.511/0.5253	30.165/0.8441	28.448/0.6451	28.442/0.7692
prune-UniFMIR		41, 5, 160, 128	1.5069	5.75	4.0003e4	4.0962e4	9856.00	5.00e4	36.668/0.8340	36.644/0.8329	35.257/0.7786	32.260/0.8787	29.542/0.7414	28.699/0.8336
prune-UniFMIR (float16)				2.88		1.0241e4	6349.66	3.19e4	36.459/0.8306	36.653/0.8343	35.134/0.7747	32.378/0.8676	29.748/0.8457	28.363/0.8170

The number of parameters (M), the model size (MB), the number of calculations (GFLOPs/BOPs), and the running time required on a CPU/GPU for an anisotropic image are shown, as are the accuracies (PSNR/SSIM values) achieved on the Liver dataset.

https://doi.org/10.1038/s41592-024-02244-3

Extended Data Table 5 | Quantitative comparison among different models for the 3D image denoising task

Model	Image Size	Input size	Param (M)	Model Size (MB)	GFLOPs	BOPs (T)	Time		Liver	
							CPU (s)	GPU (ms)	XY-slices	XZ-slices
CARE	301,752,752	1, 1, 188, 188, 301	0.9968	3.84	2.6559e4	2.7196e4	1547.36	8135.2	24.999/0.5963	25.912/0.6322
prune-UniFMIR		301, 1, 188, 188	1.5037	5.74	2.6081e5	2.6707e5	5.51e4	2.87e5	26.181/0.6276	26.238/0.6351
prune-UniFMIR (float16)				2.87		6.6767e4	3.92e4	2.04e5	25.001/0.5887	26.482/0.6445

The number of parameters (M), the model size (MB), the number of calculations (GFLOPs/BOPs), and the running time required on a CPU/GPU for a noisy image are shown, as are the accuracies (PSNR/SSIM values) achieved on the *Planaria* and *Tribolium* datasets.

https://doi.org/10.1038/s41592-024-02244-3

Extended Data Table 6 | Quantitative comparison among different models for the 3D-to-2D projection task

Model	Image Size	Input size	Param (M)	Model Size (MB)	GFLOPs	BOPs (T)	Time		Flywing			
							CPU (s)	GPU (ms)	C0	C1	C2	C3
CARE	50,520,692	1, 1, 50, 144, 176	1.3798	5.57	56.4251	57.7793	3.51	72.16	21.016/0.3819	21.999/0.5243	30.676/0.8796	20.884/0.4431
GVTNets			0.4932	1.91	93.5141	95.7584	19.84	313.12	20.101/0.3104	20.572/0.4511	28.746/0.8077	20.556/0.3966
prune-UniFMIR		1, 50, 144, 176	78.6000	325.17	2.8195e4	2.8872e4	1.95e3	1.61e4	20.525/0.3762	22.489/0.5184	31.721/0.9118	21.477/0.4510
prune-UniFMIR (float16)				162.59		7.2179e3	693.28	5.73e3	21.009/0.3795	22.483/0.5163	31.694/0.9117	21.719/0.4531

The number of parameters (M), the model size (MB), the number of calculations (GFLOPs/BOPs), and the running time required on a CPU/GPU for a volumetric input are shown, as are the accuracies (PSNR/SSIM values) achieved on the Flywing dataset.

nature portfolio

Corresponding author(s): Bo Yan

Last updated by author(s): Feb 29, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All training and testing data involved in the experiments come from existing literatures. No software was used for data collection.
Data analysis	The following tools were used: Python 3.8.3 for performing enrichment tests PyCharm Professional 2020.02 for plotting figures Pytorch 1.6.0 for building and training deep learning models Numpy 1.19.2 for numerical calculation and analysis Seaborn 0.11.1 for data visualization TensorFlow 2.8.0 The pytorch code of our UniFMIR, together with several representative trained models, as well as some example images for testing are publicly available at https://github.com/cxm12/UNiFMIR.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All training and testing data involved in the experiments come from existing literatures and can be downloaded in https://zenodo.org/records/8401470.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	There is no research involving human participants, their data, or biological material.
Reporting on race, ethnicity, or other socially relevant groupings	There is no research involving race, ethnicity, or other socially relevant groupings.
Population characteristics	There is no research involving population characteristics.
Recruitment	There is no research involving recruitment.
Ethics oversight	There is no research involving ethics oversight.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size calculation was not performed. We used the datasets from existing published works, for example, we used the BioSR dataset from a paper in Nature Methods 2021. By following theses works, the sample sizes are sufficient for demonstrating our work. And other detailed sample size information is described in the corresponding legends of figures.
Data exclusions	There is no data excluded from the analyses.
Replication	In each experiment, all evaluation metrics of the results were averaged on n (described in the corresponding legends of figures) samples, which all show similar characteristics and performance.
Randomization	No randomization was used for experimental data. The network model parameter initialization uses the default settings of the Pytorch framework, which is also the most commonly used method.
Blinding	The models were blindly tested using the data that were not included in the training process.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	There is no source of seed stocks or other plant material.
Novel plant genotypes	There is no methods related to plant genotypes.
Authentication	There is no seed stock used or genotype generated.