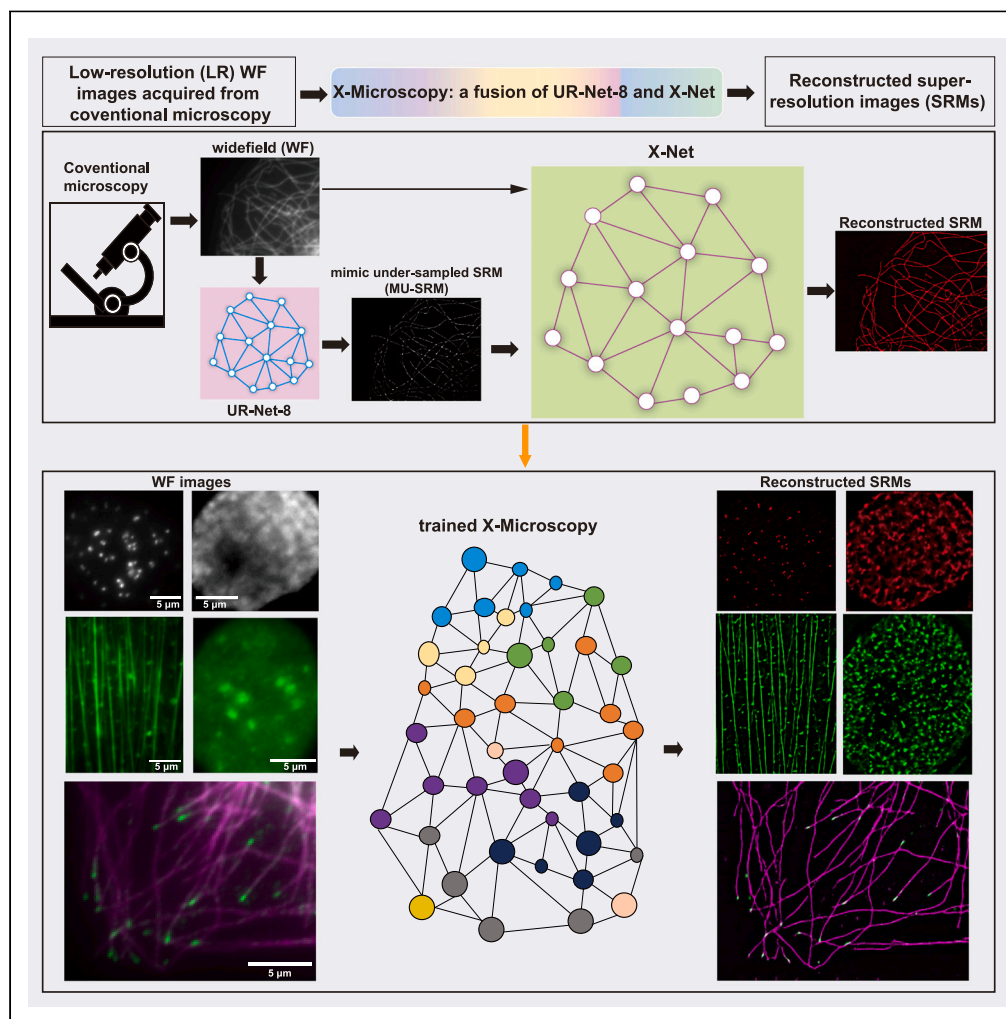## Article

Deep learning enables stochastic optical reconstruction microscopy-like superresolution image reconstruction from conventional microscopy

Lei Xu, Shichao Kan, Xiying Yu, ..., Yigang Cen, Changjun Zhu, Wei Jiang

wjiang6138@cicams.ac.cn

Highlights

X-Microscopy enables SRM reconstructions from low-resolution WF images

X-Microscopy comprises two deep learning-based subnetworks

The trained X-Microscopy predicts SRMs from WFs of various biological structures

X-Microscopy facilitates multicolor and cross-modality SRM reconstructions

iScience

Article

Deep learning enables stochastic optical reconstruction microscopy-like superresolution image reconstruction from conventional microscopy

Lei Xu,[1,2,6] Shichao Kan,[3,6] Xiying Yu,[1] Ye Liu,[5] Yuxia Fu,[1] Yiqiang Peng,[5] Yanhui Liang,[5] Yigang Cen,[4] Changjun Zhu,[2] and Wei Jiang[1,7,*]

SUMMARY

Despite its remarkable potential for transforming low-resolution images, deep learning faces significant challenges in achieving high-quality superresolution microscopy imaging from wide-field (conventional) microscopy. Here, we present X-Microscopy, a computational tool comprising two deep learning subnets, UR-Net-8 and X-Net, which enables STORM-like superresolution microscopy image reconstruction from wide-field images with input-size flexibility. X-Microscopy was trained using samples of various subcellular structures, including cytoskeletal filaments, dot-like, beehive-like, and nanocluster-like structures, to generate prediction models capable of producing images of comparable quality to STORM-like images. In addition to enabling multicolour superresolution image reconstructions, X-Microscopy also facilitates superresolution image reconstruction from different conventional microscopic systems. The capabilities of X-Microscopy offer promising prospects for making superresolution microscopy accessible to a broader range of users, going beyond the confines of well-equipped laboratories.

INTRODUCTION

Fluorescence microscopy (FM) plays a major role in monitoring cell physiology, especially subcellular structures, molecular localizations, and interactomes. Using highly specified molecular probes together with the development of noninvasive and multicolour imaging capabilities, FM allows the visualization of complex biological systems.[1] Conventional FM imaging, however, has a resolution barrier, as the resolution is limited by the wavelength of light λ and the numerical aperture (NA) of the objective lens. Developments in superresolution microscopy have overcome the long-standing diffraction limit and continued to push the resolution barrier to the nanometre level, thus enabling the visualization of previously invisible molecular details in biological systems that could previously be achieved only by electron microscopy (EM).[2–6]

Recently, superresolution microscopy has been improved on the foundation of wide-field (WF), confocal, and/or total internal reflection fluorescence (TIRF) microscope equipment. Although these setups in superresolution microscopy imaging retain some advantages of conventional microscopy, each group of superresolution techniques, such as photoactivated localization microscopy (PALM), stochastic optical reconstruction microscopy (STORM), stimulated emission depletion microscopy (STED) and structured illumination microscopy (SIM), has its own limitations.[7–10] Sophisticated optical settings, expert handling, labeling probe, and dye choices, the consideration of sample characteristics for preparations, aberration corrections, mathematical image postprocessing, and complex quantitative analyses are still required for superresolution microscopy, making fast and high-throughput superresolution microscopy imaging currently very challenging.

Deep learning (DL) has had stunning successes in biomedical research and promises to revolutionize the microscopy field.[11–25] Consequently, several studies have established quick transformations from low resolution (LR) images to superresolution microscopy images (SRMs) by DL. Prior works using DL were devoted to reconstructing superresolution views from LR images (i.e., confocal images) or undersampled superresolution microscopy images (U-SRMs) or to restoring SIM images under low signal-to-noise ratio conditions.[23,24,26–33] These studies demonstrated that DL technologies could help to overcome, at least in part, the disadvantages of superresolution microscopy. However, using DL to reconstruct high-quality STORM-like SRMs from wide-field images (WFs) with a variety of input sizes has been hampered by the challenges of acquiring large-scale training datasets, conducting exhaustive parameter searches, and the computational costs involved. It also remains uncertain whether the trained DL models can be generalized to reconstruct SRMs from multicolour WFs or WFs obtained from

[1]Department of Etiology and Carcinogenesis and State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China
[2]Key Laboratory of Molecular and Cellular Systems Biology, College of Life Sciences, Tianjin Normal University, Tianjin 300387, China
[3]School of Computer Science and Engineering, Central South University, Changsha, Hunan 410083, China
[4]Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China
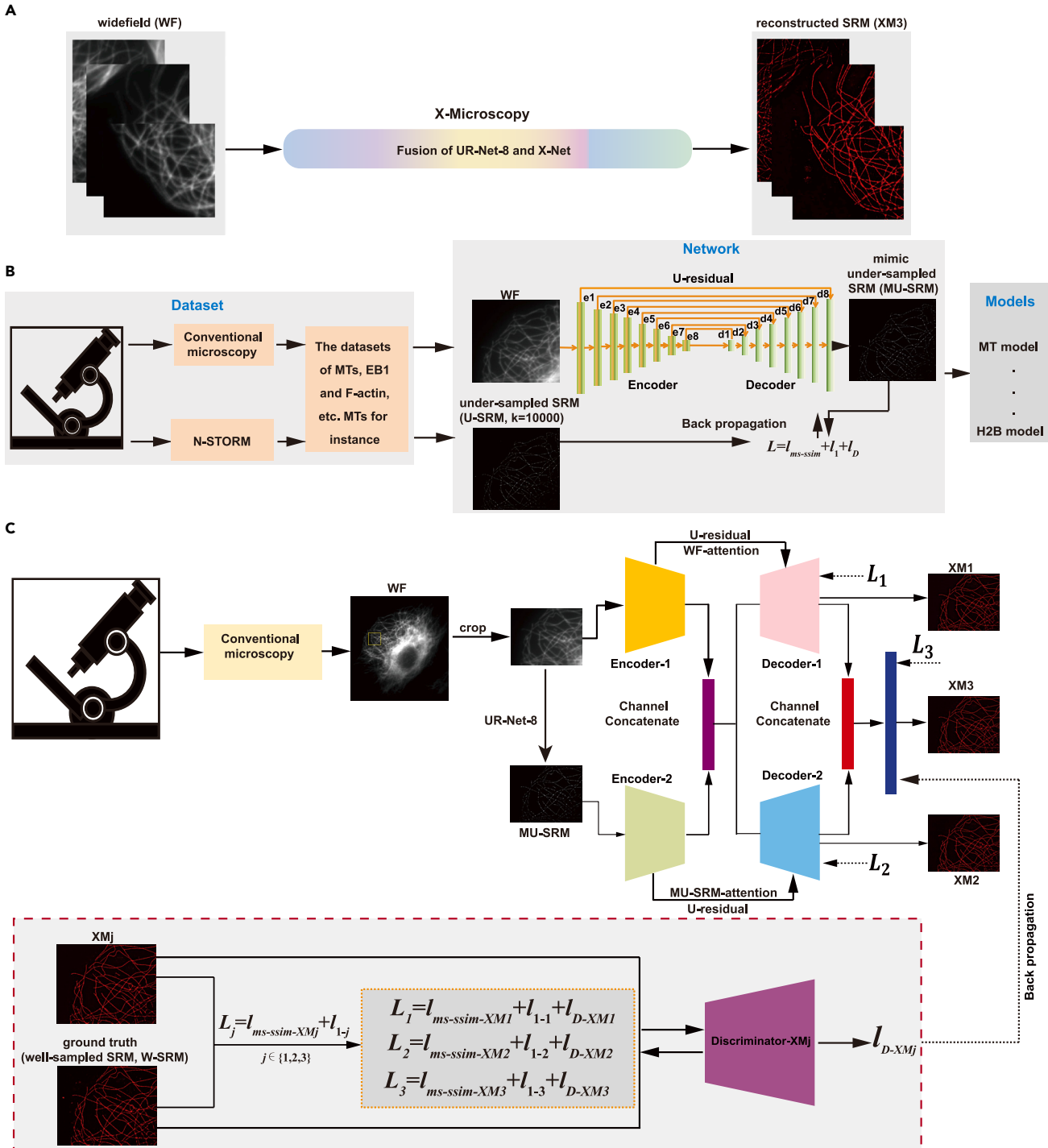[5]HAMD (Ningbo) Intelligent Medical Technology Co., Ltd, Ningbo 315194, China
[6]These authors contributed equally
[7]Lead contact
*Correspondence: wjiang6138@cicams.ac.cn
https://doi.org/10.1016/j.isci.2023.108145

Figure 1. The schematic architecture and training strategy of X-Microscopy

(A) An overview of X-Microscopy: X-Microscopy was constructed with UR-Net-8 and X-Net to reconstruct STORM-like SRM images (labeled XM3) from wide-field images (WFs). The inputs were WFs acquired from diffraction-limit conventional microscopy, and the outputs were the reconstructed SRMs.

(B) The architecture and training strategy of UR-Net-8. UR-Net-8 was constructed to generate mimic undersampled SRM images (MU-SRMs) from WFs. The training datasets were acquired simultaneously from conventional microscopy and N-STORM. UR-Net-8 comprised an encoder with 8 convolutional layers (encoding units) and a decoder with 8 deconvolutional layers (decoding units). A U-connected operation and a residual function were displayed between each encoding unit and decoding unit. For the MT model instance, WFs were defined as network inputs, and undersampled SRM images (k = 10000) served as the ground truth during training. The network outputs were mimic undersampled SRM images (MU-SRMs). The training loss functions were displayed as

Figure 1. *Continued*

the combination of MS-SSIM, l_1 norm and conditional generative adversarial loss (l_{D-XMj}) to determine the similarity between the MU-SRMs and the ground truth images.

(C) The architecture and training workflow of X-Net. X-Net was equipped with two computational paths. The input of the upper path was the WF, and the input of the lower path was the MU-SRM generated by UR-Net-8. X-Net was designed with two encoders and two decoders. The attention function and U-connection residual operation (U-residual, MU-SRM-attention, or WF-attention) modules were established between each pair of encoding units and decoding units. The output of the upper path was defined as XM1, and the output of the lower path was defined as XM2. These outputs were concatenated to obtain the output XM3. During training, XMj was compared to the ground truth image (well-sampled SRM, W-SRM) via three loss functions: L1, L2 and L3. Three loss values calculated by loss functions were used to update the learnable parameters through the back-propagation operation.

different conventional microscopic modalities. In this study, we report devising a deep network, X-Microscopy, a fusion of X-Net and UR-Net-8, to transform WFs to STORM-like SRMs.

RESULTS

X-Microscopy design and dataset preparation

Previously, Ouyang et al. demonstrated that ANNA-PALM could reconstruct SRMs of approximately similar information content as a standard PALM acquisition (with K frames) from a much smaller number of raw frames (k ≪ K, k = 300 frames and single-frame exposure time is 30ms) and/or WFs. We named these smaller number of raw frames as undersampled images, hereafter U-SRMs. If WFs and U-SRMs were used as inputs simultaneously, high-quality reconstructions of SRMs could be achieved compared with standard PALM/STORM images (full-sequenced SRMs, F-SRMs).[27] However, if WFs were directly used as inputs with the same dataset, the reconstructed SRMs displayed many erroneous features. These results suggested that the high-quality reconstruction of SRMs relied on the guidance of U-SRMs, which had to be obtained from STORM/PALM. For a desired model to achieve the same effects by using WFs alone as inputs, we hypothesized that a large-scale training dataset would be required. However, acquiring large amounts of training data is labor intensive and time consuming. To address these challenges, we developed X-Microscopy, a synthesized pixel-to-pixel[34,35] mapping network based on supervised DL[36–39] that can directly reconstruct high-quality SRMs from WFs by using mimic undersampled SRMs (MU-SRMs) and a small-scale training dataset. X-Microscopy is a fusion of two subnets,[40] UR-Net-8 and X-Net (Figure 1A), in which UR-Net-8 was trained to generate MU-SRMs from WFs (Figure 1B), while X-Net was designed to take the original WFs and the MU-SRMs as inputs to reconstruct SRMs (labeled as "XM3") (Figure 1C). Once trained, X-Microscopy was able to reconstruct STORM-like SRMs from WFs alone without a trade-off in resolution.

UR-Net-8, inspired by U-Net,[41,42] was developed with 8 U-connected residual units (UR units)[43] with the goal of finding predictive pixel relationships between U-SRMs and WFs. Each UR unit comprised 3 convolutional layers and 1 deconvolutional layer. There were two innovations in developing this algorithm: (1) The sizes of the feature maps for each convolutional layer and deconvolutional layer were calculated based on the sizes of the WFs (network inputs), which endowed UR-Net-8 with input-size flexibility. (2) A U-connected operation and a residual function in each UR unit were utilized to extract finer pixel-level features from WFs, allowing UR-Net-8 to reconstruct MU-SRMs (Figures S1A and S1B).

X-Net was utilized to reconstruct high-quality SRMs from WFs and MU-SRMs obtained by UR-Net-8. There were three innovations in X-Net: (1) X-Net was constructed with two computational paths: the upper path and the lower path comprised 8 convolutional layers and 8 deconvolutional layers, respectively. The upper path was devised to extract detailed features from WFs and fuse features of the lower path simultaneously. The lower path was applied to extract detailed pixel-level features from MU-SRMs and fuse WF features of the upper path simultaneously. Hence, these two paths were symmetrical and reciprocal to achieve interdependent fusion for acquiring, processing, and fusing feature maps (Figure S1C). (2) For effective feature extraction, a channel attention mechanism and a residual function between each convolutional layer and deconvolutional layer were adopted. Each channel attention unit was applied to refine important features of SRMs from the output of the deconvolutional layers and the input of the convolutional layers, which could determine what information should flow into the residual units (Figure S2B). Furthermore, each residual unit was utilized to filter out useful pixel-level features of MU-SRMs and WFs and to suppress useless features (Figure S2C). (3) Similar to UR-Net-8, X-Net provided input-size flexibility that enabled the reconstruction of SRMs from different sizes of LR biomedical images (Tables S1 and S2).

To train X-Microscopy, datasets were acquired consisting of WFs and paired SRMs of immunostained cytoskeletal microtubules (MTs), F-actin microfilament bundles, MT plus-end binding protein, EB1,[44] nuclear envelope structure protein Lamin B1,[45,46] nucleosome histone protein H2B and histone H3 modification marker, H3K9me3[47–50], epithelial cytoskeletal intermediate filament (IF) component, Cytokeratin 14 (CK14)[51] and kinetochore antigen, CREST[52] in different cell lines, namely, HeLa, U2OS, RPE1, U87MG, U373MG[53] and RNE-D3 (rat normal esophageal cell line).[54]

X-Net can reconstruct superresolution microscopy images from wide-field images plus undersampled superresolution microscopy images or undersampled superresolution microscopy images

A previous study showed that ANNA-PALM could reconstruct SRMs (i.e., MTs) using U-SRMs acquired from PALM and/or WFs, which were in good agreement with standard PALM SRMs.[27] To evaluate whether the X-Microscopy subnet, X-Net, could also quickly and robustly output high-quality SRMs from U-SRMs (k = 300–500) and/or WFs, we trained and tested X-Net and ANNA-PALM via our MT dataset. To this end, the
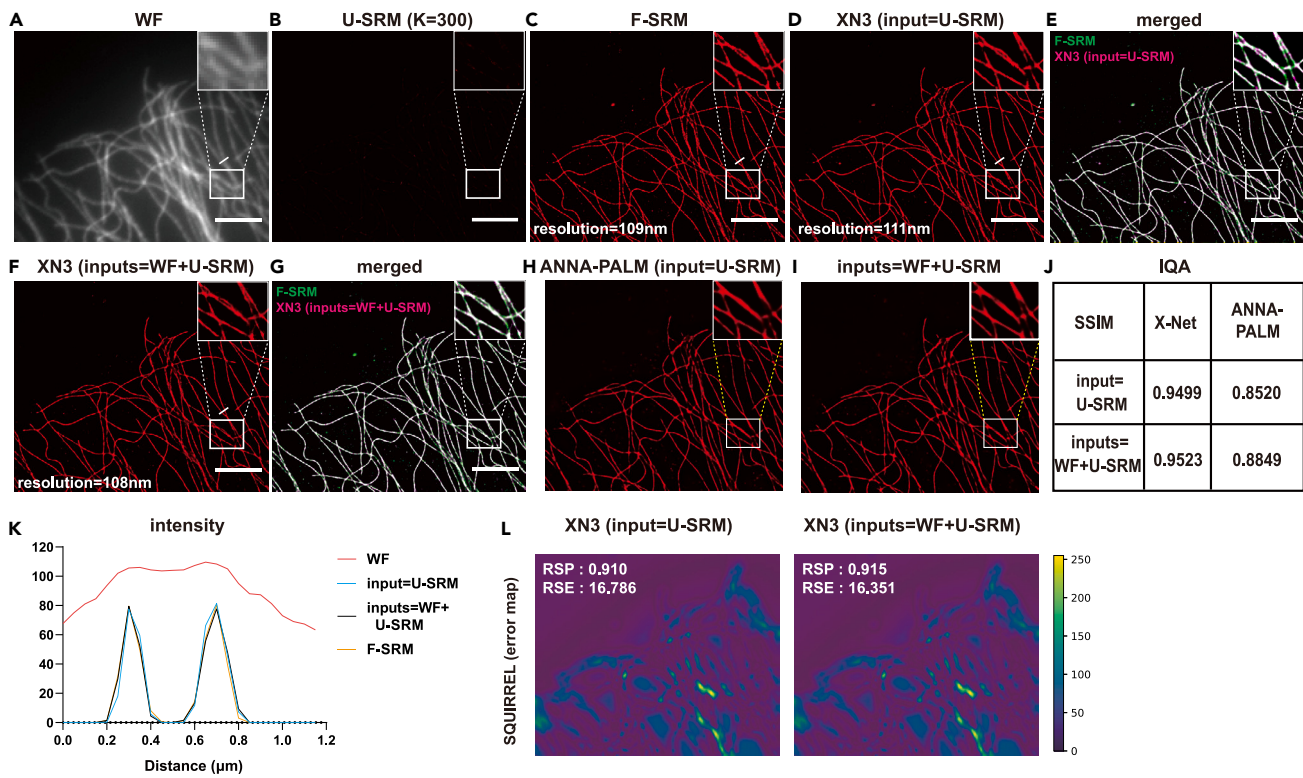
Figure 2. Comparison of performance between X-Net and ANNA-PALM

(A–I) Representative SRMs of MTs reconstructed from U-SRM (k = 300) and/or WF by X-Net and ANNA-PALM.

(A) WF acquired by conventional microscopy.

(B) U-SRM (k = 300) acquired by STORM.

(C) F-SRM (K = 60000) was acquired by STORM and used to calculate SSIM values with network predictions.

(D) The image labeled "XN3 (input = U-SRM)" was reconstructed from B only.

(E) C and D were merged to determine the similarity (magenta indicates XN3, green indicates F-SRM).

(F) The image labeled "XN3 (inputs = WF + U-SRM)" was reconstructed from A and B by X-Net. The resolution values of C, D and F were labeled in the lower left of each image.

(G) C and F were merged (magenta indicates XN3, green indicates F-SRM).

(H) The image labeled "ANNA-PALM (input = U-SRM)" was reconstructed from B.

(I) The image labeled "inputs = WF + U-SRM" was reconstructed from A and B. Scale bar, 5 μm.

(J) The image quality assessments (IQA) of the trained X-Net and ANNA-PALM in terms of SSIM metric were displayed.

(K) The intensity values along the white lines in the images in A, C, D and F were shown.

(L) The error maps, RSE scores, and RSP scores were calculated to compare the X-Net reconstructions with the WF image. The WF image was defined as a reference image. The color bar indicated magnitude of the error.

network inputs were defined as WFs plus U-SRMs or U-SRMs alone. The ground truth images were defined as well-sampled SRM images (W-SRMs), and the network outputs were labeled as XN3 or ANNA-PALM. The F-SRMs (Full-sequenced SRMs) acquired by STORM were used to perform image quality assessment (IQA) (Figures 2A–2J). As shown in Figures 2A–2G, SRMs reconstructed from the MT-trained X-Net revealed super-resolved features of MTs that were not presented in WFs and U-SRMs due to the low resolution and sparsity, consistent with those of F-SRMs obtained by STORM. Although the MT-trained ANNA-PALM also showed effectiveness in reconstructing SRMs, the SRMs obtained by ANNA-PALM displayed some reconstruction errors in the complex areas (Figures 2H and 2I). To assess the similarities between SRMs reconstructed by X-Net or ANNA-PALM and F-SRMs, we calculated the structural similarity (SSIM)[55] using our MT test dataset. The SSIM values of SRMs reconstructed by X-Net exceeded 0.95 in Figure 2J, and the average SSIM values were displayed in Figure S3A. To determine whether specific architectural designs could affect the MT-trained model performance, we performed ablation experiments involving the removal of loss functions, residual units, and channel concatenation, respectively (Figure S3B). The results showed that these loss functions, residual units, and channel concatenation played essential roles in the performance of X-Net.

We assessed the image quality of X-Net predictions. The intensity value quantifications of the reconstructed SRMs by plot profile indicated that X-Net provided high-fidelity predictions (Figure 2K). In addition to local quality assessment, the global image quality was evaluated using NanoJ-Squirrel analysis.[56] Reconstruction artifacts and local errors were indicated by error maps, RSE (Resolution Scaled Error), and RSP (Resolution Scaled Pearson coefficient) scores (Figure 2L). Moreover, the resolution quantifications[57,58] determined by decorrelation analyses also
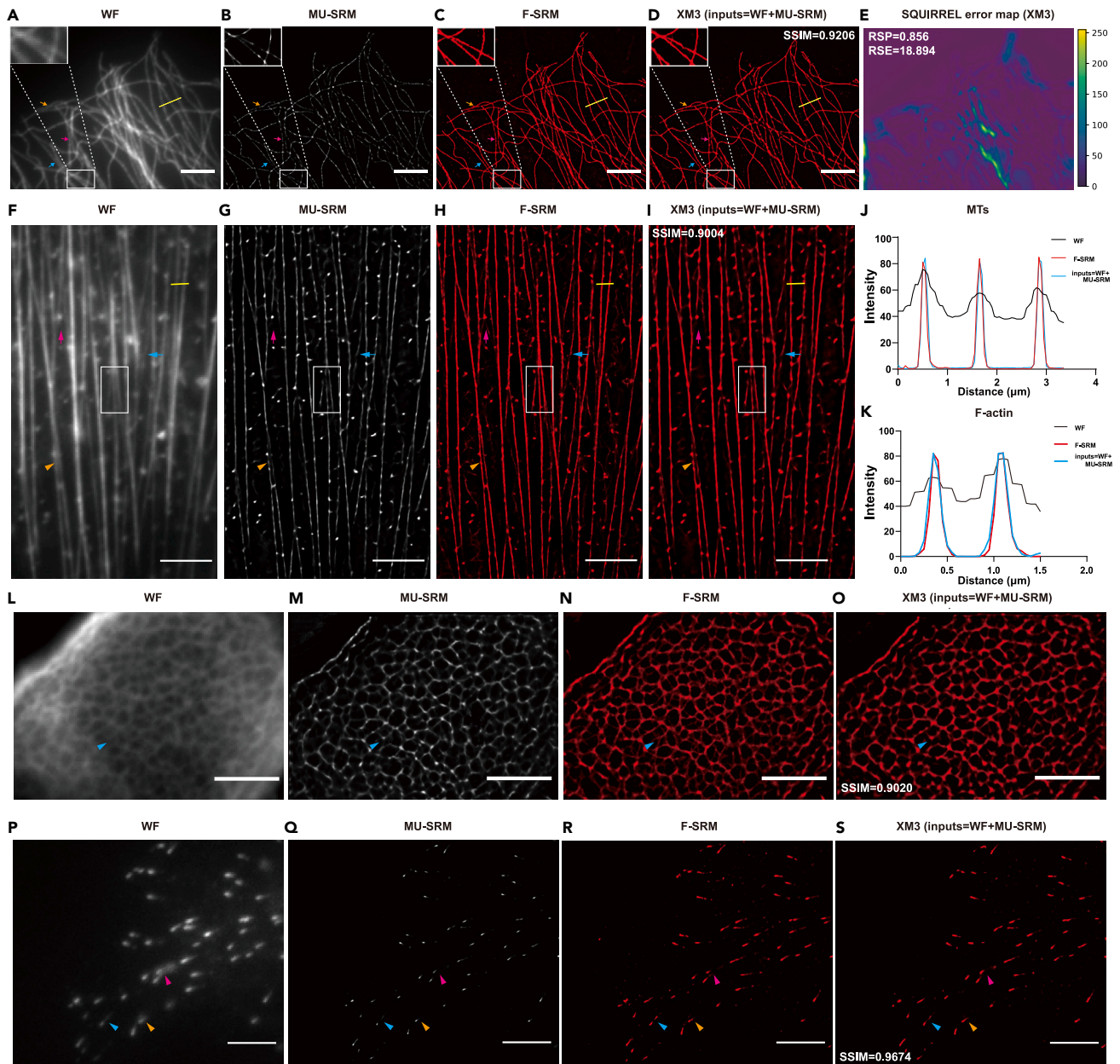
Figure 3. SRMs of MTs, F-actin, CK14, and EB1 were reconstructed from WFs using X-Microscopy

(A–D) Representative SRMs of MTs in U373MG cells reconstructed from WFs by X-Microscopy.

(A) WF acquired from conventional microscopy.

(B) MU-SRM reconstructed from UR-Net-8.

(C) F-SRM (K = 60000) was recorded from STORM.

(D) The image labeled "XM3 (inputs = WF + MU-SRM)" was reconstructed from A and B by X-Net. The super-resolved reconstruction details of B and D compared with C were highlighted by arrows (Magenta, blue and orange) and white boxes. SSIM values were calculated and labeled in the upper right of XM3 images. Scale bar, 5 μm.

(E) The error map, RSE score, and RSP score were used to assess the artifacts. The WF image was defined as the reference image and the XM3 image was defined as super-resolution reconstruction. The color bar indicated magnitude of the error.

(F–I) Representative SRMs of F-actin reconstructed from WFs by X-Microscopy in RPE1 cells stained with phalloidin-647. The reconstruction details were indicated by arrows (Magenta, blue, and orange) and white boxes. SSIM values were determined and displayed at the upper left of XM3 images. Scale bar, 5 μm.

(J) The intensity values of yellow lines in A, C, and D were quantified by line profile.

(K) The intensity values of yellow lines in F, H, and I were quantified by line profile.

Figure 3. Continued

(L–O) Representative SRMs of CK14 reconstructed from WFs by X-Microscopy in RNE-D3 cells. The reconstruction errors were indicated by blue arrows. SSIM values were displayed at the lower left of the XM3 images. Scale bar, 5 μm.

(P–S) Representative SRMs reconstructed from WFs immunostained with anti-EB1 antibody.

(P) WF obtained by conventional microscopy.

(Q) MU-SRM predicted by UR-Net-8.

(R) F-SRM acquired by STORM.

(S) The image labeled "XM3 (inputs = WF + MU-SRM)" was reconstructed from MU-SRM generated by UR-Net-8 and original WF using X-Net. Arrows (Magenta, blue and orange) show the reconstruction details. The reconstruction qualities were determined by the SSIM metric, and the calculated values were marked at the lower left of the XM3 images. Scale bar, 5 μm.

indicated that SRMs reconstructed by the X-Net were similar with F-SRMs (Figures S4A-S4F). To verify the importance of the input-size flexibility in SRM reconstructions, X-Net and ANNA-PALM were tested by WF images with irregular sizes, respectively (Figure S4G). The results showed that, when compared with ANNA-PALM, the trained X-Net could reconstruct super-resolved SRMs from these irregular-sized WF images, while preserving the original morphologies of subcellular structures (Figures S4H–S4L). Furthermore, we tested the robustness of X-Net by reconstructing SRMs from WFs of MTs obtained from different cell lines and fixed by different methods (Figure S5).

X-Microscopy can restore high-quality superresolution microscopy images from wide-field images alone

To restore high-quality SRMs from WFs alone with a small-scale training dataset, we next determined whether SRMs could be reconstructed directly from WFs via the trained UR-Net-8 or X-Net. Although SRMs reconstructed by UR-Net-8 and X-Net exhibited clear and curvilinear MT structures that were not displayed in WFs, reconstruction errors could be detected in areas where the MT structures were complex (Figures S6A-S6D). Similar results were also obtained previously by ANNA-PALM.[27] These results corroborated that applying U-SRMs acquired from STORM as guides was an essential factor for reconstructing high-quality SRMs from WFs with limited training datasets.

Therefore, we hypothesized if we could generate MU-SRMs from WFs, similar to U-SRMs acquired from STORM, to guide WFs for final reconstructions of SRMs. To test the idea, we applied UR-Net-8 to generate MU-SRMs from WFs and used the MU-SRMs as guides to provide extra information (resolution) for reconstructing SRMs of MTs from WFs alone by X-Net. We trained UR-Net-8 using WFs as inputs and paired STORM U-SRMs as the ground truth. The trained UR-Net-8 could transform WFs into superresolution MU-SRMs that matched with the U-SRMs obtained by STORM. Meanwhile, we determined that the use of multi-sequence U-SRMs (k = 10000) rather than few-shot U-SRMs (k = 300–500) for model training could achieve a better MU-SRM generation (Figures S7A–S7F). Furthermore, to analyze the reconstruction qualities of MU-SRMs, we merged the MU-SRM and U-SRM to determine the Pearson's coefficient. Although the MU-SRM and U-SRM did not completely overlap, the locations of reconstructed pixels and the true pixels were all located on the same microtubules. These results indicated that the generated MU-SRMs could simulate the role of real U-SRMs obtained by STORM, guiding WFs to learn the correct pixel distributions (resolution) during X-Net training (Figure S7G). Therefore, we integrated UR-Net-8 and X-Net into a single pipeline, named X-Microscopy.

X-Microscopy was trained to generate superresolution microscopy image models for various subcellular structures

X-Microscopy was trained with diverse datasets of subcellular structures, including cytoskeletal filaments, dot-like structures, beehive-like structures, and nanocluster-like structures, to create various SRM prediction models (e.g., the MT-trained model, the F-actin model, and so forth). After training, the SRMs of cytoskeletal structures, such as MTs, F-actin microfilament bundles, and Cytokeratin 14 intermediated filaments (CK14), reconstructed by the individual MT, F-actin or CK14 trained X-Microscopy displayed their structural details. For instance, the reconstructed MT or F-actin SRMs showed that super-resolved visualization was observed not only in areas with sparsely distributed structures but also in areas where bifurcations and crisscrosses were formed (Figures 3A–3D and 3F–3I; Figure S8). In addition, the CK14 trained X-Microscopy could also reconstruct SRMs from WFs (Figures 3L-3O). Quantification analyses using the plot profile demonstrated that the SRMs of MTs or F-actin filaments reconstructed by the MT or F-actin trained X-Microscopy were nearly the same as the F-SRMs obtained by STORM (Figures 3J and 3K). However, reconstruction errors could be observed. These reconstruction artifacts and structure errors were determined by NanoJ-Squirrel analysis (Figures 3E and S9A–S9E).

To further demonstrate that X-Microscopy could be applied to reconstruct SRMs from WFs of other subcellular structures, we established more structure-specific X-Microscopy models. SRMs of EB1 reconstructed by the EB1-trained X-Microscopy clearly revealed comet shapes of EB1 proteins that were not previously recognizable in WFs, while SRMs of Lamin B1 predicted by the Lamin B1-trained X-Microscopy exhibited clear and continuous structures of the nuclear envelope and eliminated many in-focus backgrounds present in WFs. The EB1 or Lamin B1 SRMs reconstructed by X-Microscopy achieved high SSIM values (Figures 3P–3S and 4A–4E; Figures S10A and S10B). The reconstruction artifacts of EB1 were determined by error maps (Figures S9F–S9J). Thus, the reconstructed SRMs of EB1 or Lamin B1 enable us to determine the structures of MT growing plus-ends or the shapes and sizes of nuclei at super-resolution levels.

We also reconstructed SRMs of kinetochores from WFs with the kinetochore-trained X-Microscopy. The results demonstrated that X-Microscopy could compensate for some STORM imaging limitations, such as phototoxicity and photobleaching, to precisely reconstruct SRMs for subcellular structures/macromolecules, i.e., paired sister kinetochores. Paired sister-kinetochores in superresolution images obtained by STORM often showed uneven acquisition due to the extended time of STORM imaging and photobleaching, although they could be separated well when compared with blurred paired sister-kinetochores in WFs. However, the SRMs reconstructed by the
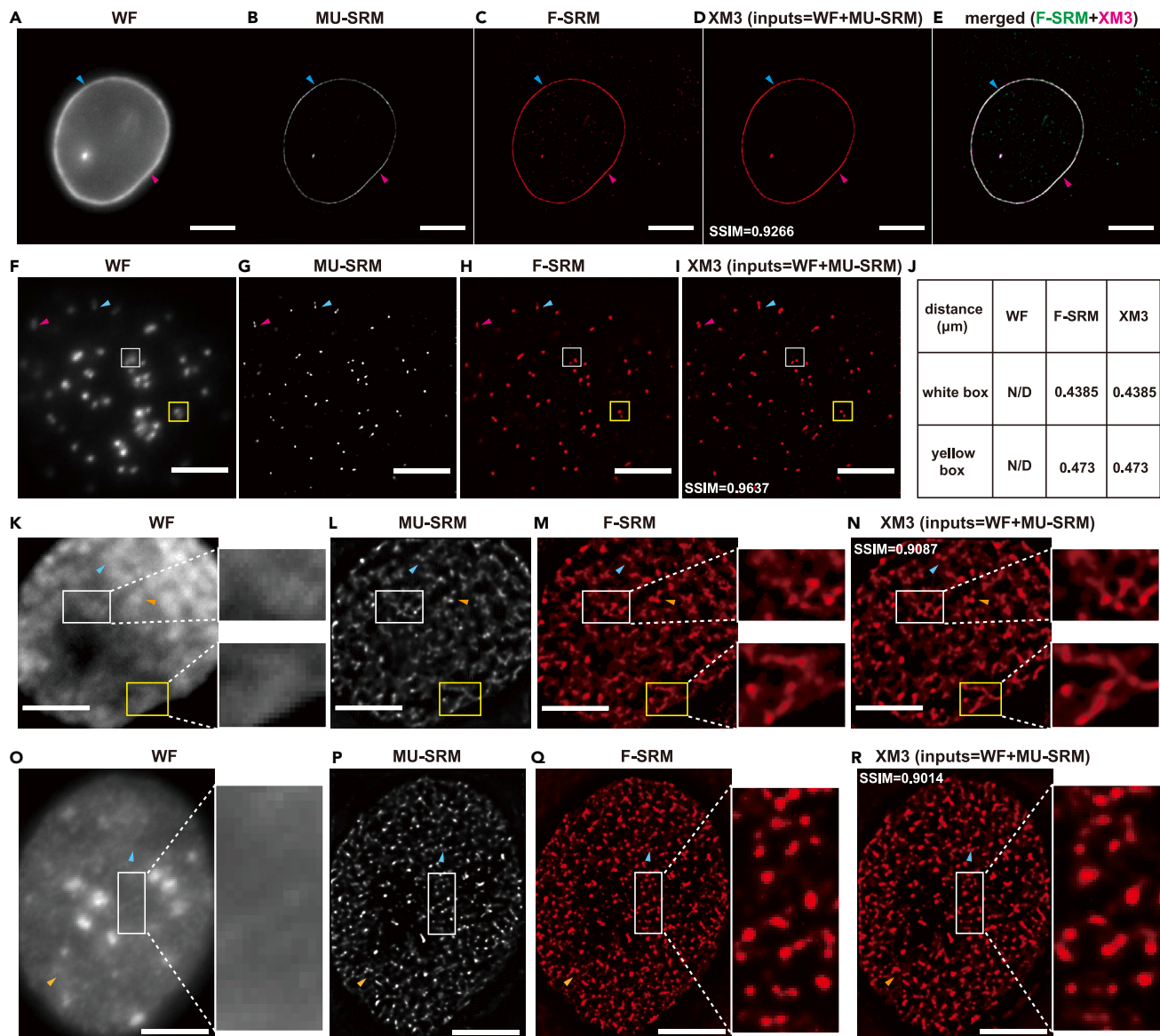
Figure 4. SRMs of Lamin B1, kinetochore, H2B and H3K9me3 reconstructed from WFs by X-Microscopy

(A–E) Representative SRMs reconstructed from the WFs of Lamin B1 in U2-OS cells by X-Microscopy. The calculated SSIM values between the F-SRM and XM3 images were displayed in XM3 images. Blue and magenta arrows indicate the reconstruction details. The reconstruction errors were indicated by the areas where red and green intersect in E. Scale bar, 5 μm.

(F–I) Representative images of WF, MU-SRM, F-SRM, and XM3 in U2-OS cells immunostained with anti-CREST antibody. Reconstruction details in XM3 images were indicated by blue and magenta arrows. The reconstruction qualities were determined by SSIM, and the SSIM values were marked in XM3 images. Scale bar, 5 μm.

(J) The distances of sister-kinetochores labeled with white boxes and yellow boxes in F, H and I were shown. The distances of sister-kinetochores in WFs could not be determined (N/D).

(K–N) Representative images of WF, MU-SRM, F-SRM, and XM3 in U2-OS cells stained with H2B antibody. The reconstruction details were indicated by arrows (blue and orange) and white boxes. The SSIM value was shown at the upper left of the XM3 images. Scale bar, 5 μm.

(O–R) Representative SRMs of H3K9me3 reconstructed from WF in U2-OS cells by X-Microscopy were displayed. The reconstruction details were indicated by arrows (blue and orange) and white boxes. The SSIM value was shown in the upper left part of the XM3 images. Scale bar, 5 μm.

kinetochore-trained X-Microscopy revealed paired sister-kinetochores not only as two clear splits but also as paired dot structures (Figures 4F–4I; Figures S11A and S11B). We also retrieved the distances between the sister kinetochores of WF, F-SRM and XM3 images. Given the fact that the examined cells were interphase cells, the distances of those kinetochores could be only measured from F-SRM and XM3 images but not from WF images due to the diffraction limitation. Figure 4J shows those quantification results. In addition, when

F-SRMs and WFs of MTs were acquired with a nearly quenching imaging buffer, SRMs reconstructed by X-Microscopy displayed more continuations of MT filaments with fewer reconstruction errors than F-SRMs obtained by STORM (Figure S11C). We also systematically verified the qualities of SRMs reconstructed from WFs and MU-SRMs using the test datasets obtained from different biological structures and calculated the SSIM values between F-SRMs and XM3 images (Figure S11D).

Next, X-Microscopy was used to reconstruct SRMs from WFs stained by anti-nucleosome histone H2B and H3 modification and H3K9me3 antibodies in interphase cells, aiming to determine whether SRMs could be reconstructed from WFs alone with even more complex subcellular structures, such as genome-wide higher-order chromatin structures. Consistent with published data,[47–50,59,60] F-SRMs obtained by STORM revealed that H2B and H3K9me3 displayed two major structural characteristics of higher-order chromatin, i.e., H2B staining that showed spatially segregated nucleosome nanoclusters and H3K9me3 staining that revealed highly condensed aggregates. The SRMs reconstructed from WFs by the Histone-trained X-Microscopy were very similar to the F-SRMs obtained by STORM. However, reconstruction errors could still be observed in areas where the WFs were too blurred. Qualitative analyses confirmed the structural similarity between F-SRMs and the reconstructed SRMs. The SSIM values of H2B and H3K9me3 SRMs reconstructed by the Histone-trained X-Microscopy achieved high scores (Figures 4K–4N and 4O–4R). Taken together, these results demonstrate that the structure-specific X-Microscopy could reconstruct SRMs from WFs alone.

Multicolour or cross-modality superresolution microscopy image reconstructions by X-Microscopy

Multicolour STORM imaging is challenging. One of the major limitations of multicolour STORM is that it is difficult to find a second fluorescent label that produces localizations of the same quality as far-red dyes such as Alexa Fluor 647. To determine whether X-Microscopy could predict high-quality multicolour SRMs from WFs, we first predicted SRMs of MTs and EB1 proteins with the MT and EB1 trained X-Microscopy from WFs. As shown in Figures 5A–5G, multicolour SRMs restored by the MT and EB1 trained X-Microscopy displayed superresolution structures where the MT signals gradually decayed from the peak position of EB1 toward their tips, which was consistent with the previous study.[61]

To further evaluate the biological significance of X-Microscopy, we tested whether the MT and EB1 trained X-Microscopy could reconstruct SRMs from WFs co-stained with MTs and EB1 proteins under different experimental perturbations to assess MT dynamic instabilities.[62,63] Cells treated with nocodazole or taxol displayed substantial alterations of the MT network in cells with drastic reductions of EB1 on MTs curved by nocodazole (depolymerization) or straightened by taxol (stabilization) (Figure S12). We also extended our study to reconstruct multicolour SRMs from WFs co-immunostained for MTs and F-actin, or WFs co-immunostained for MTs, Lamin B1, and the kinetochore antigen CREST. For MTs and F-actin filaments, multicolour SRMs predicted by the WFs of MT and F-actin trained X-Microscopy clearly displayed transformation configurations in which actin filaments were distributed around the periphery under the plasma membrane and formed bundles, while MTs radiated from the MT organization center (MTOC) to form a dense network (Figures 5H and 5I). Moreover, SRMs reconstructed by the MT, Lamin B1, and CREST trained X-Microscopy from WFs revealed that blurred and indistinguishable areas could be transformed to a superresolution level (Figures 5J and 5K). Decorrelation analyses of reconstructed SRMs and WFs indicated that the resolution of SRMs could be improved by X-Microscopy (Figures S13 and S14). Thus, these results demonstrated the good performance of X-Microscopy in the multicolour reconstruction of SRMs.

To demonstrate whether structure-specific X-Microscopy could reconstruct SRMs from WFs acquired using different microscopic platforms that were never trained, WFs of MTs recorded by Nikon or Leica conventional microscopes with 100× objectives were used for testing. Figures 6A and 6B showed that the MT trained X-Microscopy was capable of reconstructing SRMs to resolve the bundled undistinguishable microtubules in WF images. The possible artifacts in SRM images reconstructed with X-Microscopy were indicated with error maps, RSE and RSP scores (Figure 6C). These results indicated that X-Microscopy had the potential to reconstruct high-quality SRMs from different conventional microscopes.

Finally, we asked whether X-Microscopy could be trained using all the datasets acquired above to generate a generalized model, and whether this model could be used for SRM reconstructions from WFs, with or without subcellular structures included in the training. We based this inquiry on the inherent design of X-Microscopy, which had demonstrated the ability to generate high-quality SRMs from WFs using individualized models. To this end, all the datasets of various subcellular structures presented above were used to train X-Microscopy and the trained generalized X-Microscopy was applied to reconstruct SRMs from WFs of various subcellular structures above. Consistently, like the structure-specific X-Microscopy, the generalized X-Microscopy could reconstruct high-quality SRMs from WFs. When dealing with more complex structures such as H2B, the generalized model produced more reconstruction errors (Figure S15). We also tested the generalized X-Microscopy with outer mitochondrial membrane images that have never been trained (Figure S16). Thus, these results indicated that the generalized X-Microscopy holds the potential to offer a solution for testing new biological structures, showcasing its versatility in broader applications.

DISCUSSION

Deep learning is becoming an important tool for resolution enhancement, especially the reconstruction of SRMs in fluorescence microscopy. Thus far, several studies have demonstrated that networks built on U-Net and generative adversarial network (GAN) architectures leverage and transform LR images to superresolution (SR) images.[21,23,24,26–33,64] However, these "state-of-the-art" superresolution reconstruction methods show some limitations in reconstructing SRMs from WFs alone. For instance, it is difficult to establish accurate SRM reconstruction models with small-scale training datasets and it is challenging for these models to test various input sizes of WFs.
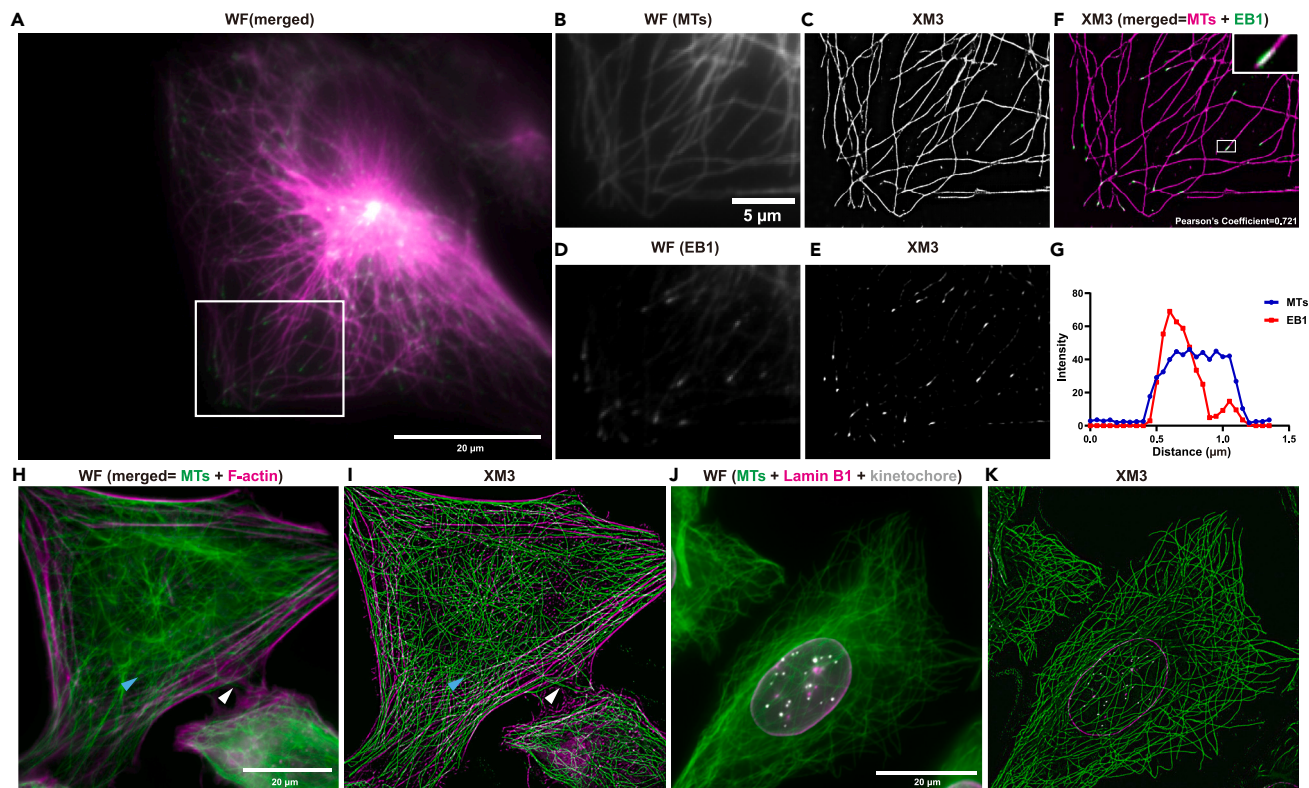
Figure 5. Multicolour SRMs reconstructed from WFs by X-Microscopy

(A–E) Representative images of WF and XM3 in U2-OS cells stained with anti-α-tubulin and anti-EB1 antibodies.

(A) The image labeled "WF (merged)" of EB1 and MTs was recorded by conventional microscopy. The merged image was displayed: magenta indicates MTs and green indicates EB1. Scale bars, 20 μm.

(B) The WF of MTs labeled with a white box as in A. Scale bar: 5 μm.

(C) The XM3 of MTs was generated from B by X-Microscopy.

(D) The WF of EB1 labeled with a white box as in A.

(E) The XM3 of EB1 generated from D by X-Microscopy was shown.

(F) Merged images predicted by X-Microscopy with the pretrained MT and EB1 models were shown. The reconstruction details in F were highlighted by white boxes, and magnified images were attached to the upper right of the XM3 images. The Pearson correlation coefficients of the EB1 and MT labeled with white boxes were determined and marked in the lower right of the XM3 images.

(G) Line profiles of fluorescence intensity plotted from the white boxes in F.
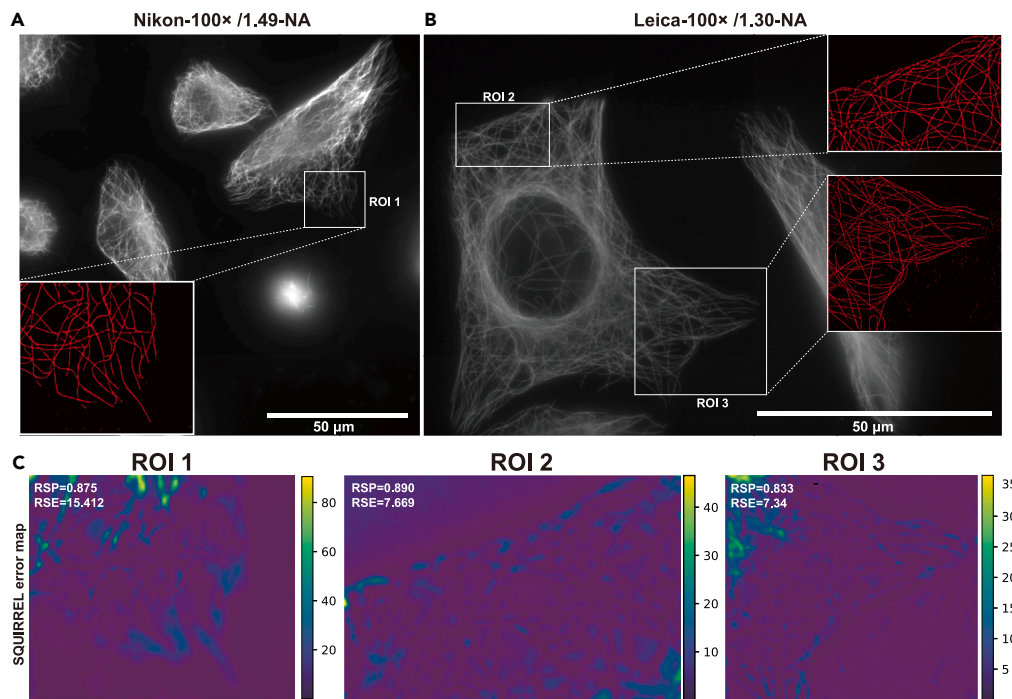
(H–I) Representative images of WF and XM3 in U2-OS cells costained with anti-α-tubulin and phalloidin-647.

(H) The image labeled "WF (merged = MTs + F-actin)" was obtained by conventional microscopy with a 100×/1.49-NA objective lens. Color map: magenta for F-actin, green for MTs.

(I) The XM3 images were predicted by X-Microscopy with two pretrained models of MTs and F-actin. The reconstruction details in the XM3 image compared with the WF image were marked by white and cyan arrowheads. Scale bar, 20 μm.

(J–K) Representative WF and XM3 images of U2-OS cells costained with anti-α-tubulin, anti-Lamin B1, and anti-CREST antibodies were shown. Magenta: Lamin B1, Green: MTs, Gray: Kinetochore. Scale bar, 20 μm.

In this study, we devised X-Microscopy to meet these challenges and showed that X-Microscopy could predict SRMs from WFs alone with limited training datasets and achieve comparable image quality to STORM. We performed quantitative analyses to define the limited dataset size for UR-Net-8 (Figure S17). By combining highly efficient deep learning architectures based on pix2pix, i.e., a U-SRM generation network (UR-Net-8) and a home-built two branching network (X-Net), we demonstrated that X-Microscopy was able to leverage binary branches symmetrically and reciprocally to reconstruct STORM-like SRMs from WFs. UR-Net-8 was applied to generate MU-SRMs, and then MU-SRMs were used to guide WFs to accomplish the final predictions of SRMs by X-Net. Hence, the implementation of UR-Net-8 could minimize the need for a large training dataset in SRM reconstructions. Additionally, X-Net that consisted of two encoders and two decoders effectively extracted intricate features from both WFs and MU-SRMs by learning distinct convolutional kernels and integrating them through connections between encoding and decoding units. Moreover, a channel attention mechanism and a residual function were also adopted in each cross-connection module that could transform input-related salient features to precise representations of high-frequency information. Furthermore, unlike other published networks, the input image sizes of UR-Net-8 and X-Net were flexible during the test stages, which was attributed to

Figure 6. SRMs reconstructed from WFs by X-Microscopy using different microscopic systems

(A and B) Representative images of WF and XM3 in U2-OS cells stained with tubulin were shown. The WFs were acquired by a Nikon Eclipse Ti2 microscope with a 100×/1.49-NA objective lens and a Leica DM5000B microscope equipped with a 100×/1.30-NA objective lens. The XM3 images were reconstructed by X-Microscopy with the pretrained MT model. The representative XM3 images of each region of interest (ROI) were displayed on the corresponding WF image. Scale bars: 50 μm.

(C) The XM3 image of each ROI was quantified with NanoJ-Squirrel analysis. The error maps, RSP scores and RSE scores were shown. The color bar indicated magnitude of the error.

calculating the sizes of feature maps automatically based on various inputs. In summary, X-Microscopy was tailor-made for the SRM reconstructions from WFs alone.

We used the MT dataset to demonstrate the necessity of the X-Microscopy pipeline instead of directly training X-Net and UR-Net-8 with WFs and SRMs (Figure S6). We also compared X-Microscopy with published networks, specifically a deep learning-based single-frame super-resolution microscopy method (SFSRM),[23] and a deep learning-based point-scanning super-resolution method (PSSR),[32] both trained directly using our MT dataset. The results revealed that X-Microscopy exhibited superior performance (Figure S18). We illustrate the robustness and generalization of X-Microscopy for SRM reconstructions of different macromolecular structures from WFs alone by the individual structural-specific trained X-Microscopy. These include SRM predictions of cytoskeletal structures, i.e., microtubules (MTs), F-actin microfilament bundles and CK14 intermediate filaments (IFs), MT plus-end binding protein EB1, nuclear lamina proteins, kinetochores, and chromatin nucleosomes in different cell lines. Moreover, X-Microscopy enabled the reconstruction of multicolour SRMs from WFs and could predict SRMs from WFs alone obtained by different conventional microscopic systems. Taken together, these results demonstrated that X-Microscopy allowed high-quality construction of SRMs from WFs alone of different subcellular macromolecular structures and paved the road for multi-dimensional applications of X-Microscopy in biomedical studies.

Limitations of the study

X-Microscopy, as a computational tool, also encounters challenges similar to those faced in other published works.[23,27,30] The data presented in the study most come from the individual structural-specific trained X-Microscopy. The "generalized" trained X-Microscopy provides a solution to predict SRMs from WFs of new subcellular structures (Figures S15 and S16). However, X-Microscopy requires new datasets for training or fine-tuning to effectively handle new structures with significant structural differences from the pre-trained model, thereby enhancing SRM reconstruction accuracy. In the future, more variable subcellular structure datasets need to be collected to train and refine the "generalized" trained X-Microscopy. In addition, more sophisticated quantification methods are required for more extensive characterizations of hallucination artifacts and reconstruction uncertainties during X-Microscopy training, validating and testing.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108145.

AUTHOR CONTRIBUTIONS

LX and WJ designed the research; LX, SK, and YC designed the deep neural network; LX and SK performed the research; LX, SK, XY, YP, YF, YL, CZ, YL, and WJ analyzed the data; and LX, SK, and WJ wrote the article.

DECLARATION OF INTERESTS

Wei Jiang, Ye Liu, and Yanhui Liang are the founders of HAMD (Ningbo) Intelligent Medical Technology Co. Ltd and members of its scientific advisory board. Wei Jiang et al. hold the patent " ZL 2021 1 1600867.7 " related to this work.

REFERENCES

1. Sigal, Y.M., Zhou, R., and Zhuang, X. (2018). Visualizing and discovering cellular structures with super-resolution microscopy. Science *361*, 880–887. https://doi.org/10.1126/science.aau1044.

2. Schermelleh, L., Heintzmann, R., and Leonhardt, H. (2010). A guide to super-resolution fluorescence microscopy. J. Cell Biol. *190*, 165–175. https://doi.org/10.1083/jcb.201002018.

3. Stewart, M., and Vigers, G. (1986). Electron microscopy of frozen-hydrated biological material. Nature *319*, 631–636. https://doi.org/10.1038/319631a0.

4. Schermelleh, L., Ferrand, A., Huser, T., Eggeling, C., Sauer, M., Biehlmaier, O., and Drummen, G.P.C. (2019). Super-resolution microscopy demystified. Nat. Cell Biol. *21*, 72–84. https://doi.org/10.1038/s41556-018-0251-8.

5. Liu, S., Hoess, P., and Ries, J. (2022). Super-Resolution Microscopy for Structural Cell Biology. Annu. Rev. Biophys. *51*, 301–326. https://doi.org/10.1146/annurev-biophys-102521-112912.

6. Fuhrmann, M., Gockel, N., Arizono, M., Dembitskaya, Y., Nägerl, U.V., Pennacchietti, F., Damenti, M., Testa, I., and Willig, K.I. (2022). Super-Resolution Microscopy Opens New Doors to Life at the Nanoscale. J. Neurosci. 42, 8488–8497. https://doi.org/10.1523/JNEUROSCI.1125-22.2022.

7. Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., and Hess, H.F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. Science 313, 1642–1645. https://doi.org/10.1126/science.1127344.

8. Rust, M.J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat. Methods 3, 793–795. https://doi.org/10.1038/nmeth929.

9. Hell, S.W., and Wichmann, J. (1994). Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. Opt. Lett. 19, 780–782. https://doi.org/10.1364/ol.19.000780.

10. Gustafsson, M.G.L. (2005). Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. Proc. Natl. Acad. Sci. USA 102, 13081–13086. https://doi.org/10.1073/pnas.0406877102.

11. Angermueller, C., Pärnamaa, T., Parts, L., and Stegle, O. (2016). Deep learning for computational biology. Mol. Syst. Biol. 12, 878. https://doi.org/10.15252/msb.20156651.

12. Arganda-Carreras, I., Kaynig, V., Rueden, C., Eliceiri, K.W., Schindelin, J., Cardona, A., and Sebastian Seung, H. (2017). Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. Bioinformatics 33, 2424–2426. https://doi.org/10.1093/bioinformatics/btx180.

13. Eulenberg, P., Köhler, N., Blasi, T., Filby, A., Carpenter, A.E., Rees, P., Theis, F.J., and Wolf, F.A. (2017). Reconstructing cell cycle and disease progression using deep learning. Nat. Commun. 8, 463. https://doi.org/10.1038/s41467-017-00623-3.

14. Hinton, G.E., and Salakhutdinov, R.R. (2006). Reducing the dimensionality of data with neural networks. Science 313, 504–507. https://doi.org/10.1126/science.1127647.

15. Kraus, O.Z., Grys, B.T., Ba, J., Chong, Y., Frey, B.J., Boone, C., and Andrews, B.J. (2017). Automated analysis of high-content microscopy data with deep learning. Mol. Syst. Biol. 13, 924. https://doi.org/10.15252/msb.20177551.

16. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539.

17. Lu, A.X., Kraus, O.Z., Cooper, S., and Moses, A.M. (2019). Learning unsupervised feature representations for single cell microscopy images with paired cell inpainting. PLoS Comput. Biol. 15, e1007348. https://doi.org/10.1371/journal.pcbi.1007348.

18. Shkolyar, A., Gefen, A., Benayahu, D., and Greenspan, H. (2015). Automatic detection of cell divisions (mitosis) in live-imaging microscopy images using Convolutional Neural Networks. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 2015, 743–746. https://doi.org/10.1109/embc.2015.7318469.

19. von Chamier, L., Laine, R.F., and Henriques, R. (2019). Artificial intelligence for microscopy: what you should know. Biochem. Soc. Trans. 47, 1029–1040. https://doi.org/10.1042/bst20180391.

20. David, D.J.V. (2022). Deep-learning for microscopy. Nat. Cell Biol. 24, 1321. https://doi.org/10.1038/s41556-022-00995-9.

21. Wang, K. (2021). Deep-learning-enhanced light-field microscopy. Nat. Methods 18, 459–460. https://doi.org/10.1038/s41592-021-01151-1.

22. Kobayashi, H., Cheveralls, K.C., Leonetti, M.D., and Royer, L.A. (2022). Self-supervised deep learning encodes high-resolution features of protein subcellular localization. Nat. Methods 19, 995–1003. https://doi.org/10.1038/s41592-022-01541-z.

23. Chen, R., Tang, X., Zhao, Y., Shen, Z., Zhang, M., Shen, Y., Li, T., Chung, C.H.Y., Zhang, L., Wang, J., et al. (2023). Single-frame deep-learning super-resolution microscopy for intracellular dynamics imaging. Nat. Commun. 14, 2854. https://doi.org/10.1038/s41467-023-38452-2.

24. Qiao, C., Li, D., Liu, Y., Zhang, S., Liu, K., Liu, C., Guo, Y., Jiang, T., Fang, C., Li, N., et al. (2023). Rationalized deep learning super-resolution microscopy for sustained live imaging of rapid subcellular processes. Nat. Biotechnol. 41, 367–377. https://doi.org/10.1038/s41587-022-01471-3.

25. Pärnamaa, T., and Parts, L. (2017). Accurate classification of protein subcellular localization from high-throughput microscopy images using deep learning. G3 (Bethesda) 7, 1385–1392. https://doi.org/10.1534/g3.116.033654.

26. Jin, L., Liu, B., Zhao, F., Hahn, S., Dong, B., Song, R., Elston, T.C., Xu, Y., and Hahn, K.M. (2020). Deep learning enables structured illumination microscopy with low light levels and enhanced speed. Nat. Commun. 11, 1934. https://doi.org/10.1038/s41467-020-15784-x.

27. Ouyang, W., Aristov, A., Lelek, M., Hao, X., and Zimmer, C. (2018). Deep learning massively accelerates super-resolution localization microscopy. Nat. Biotechnol. 36, 460–468. https://doi.org/10.1038/nbt.4106.

28. Qiao, C., Li, D., Guo, Y., Liu, C., Jiang, T., Dai, Q., and Li, D. (2021). Evaluation and development of deep neural networks for image super-resolution in optical microscopy. Nat. Methods 18, 194–202. https://doi.org/10.1038/s41592-020-01048-5.

29. Wagner, N., Beuttenmueller, F., Norlin, N., Gierten, J., Boffi, J.C., Wittbrodt, J., Weigert, M., Hufnagel, L., Prevedel, R., and Kreshuk, A. (2021). Deep learning-enhanced light-field imaging with continuous validation. Nat. Methods 18, 557–563. https://doi.org/10.1038/s41592-021-01136-0.

30. Wang, H., Rivenson, Y., Jin, Y., Wei, Z., Gao, R., Günaydın, H., Bentolila, L.A., Kural, C., and Ozcan, A. (2019). Deep learning enables cross-modality super-resolution in fluorescence microscopy. Nat. Methods 16, 103–110. https://doi.org/10.1038/s41592-018-0239-0.

31. Wu, Y., Rivenson, Y., Wang, H., Luo, Y., Ben-David, E., Bentolila, L.A., Pritz, C., and Ozcan, A. (2019). Three-dimensional virtual refocusing of fluorescence microscopy images using deep learning. Nat. Methods 16, 1323–1331. https://doi.org/10.1038/s41592-019-0622-5.

32. Fang, L., Monroe, F., Novak, S.W., Kirk, L., Schiavon, C.R., Yu, S.B., Zhang, T., Wu, M., Kastner, K., Latif, A.A., et al. (2021). Deep learning-based point-scanning super-resolution imaging. Nat. Methods 18, 406–416. https://doi.org/10.1038/s41592-021-01080-z.

33. Park, H., Na, M., Kim, B., Park, S., Kim, K.H., Chang, S., and Ye, J.C. (2022). Deep learning enables reference-free isotropic super-resolution for volumetric fluorescence microscopy. Nat. Commun. 13, 3297. https://doi.org/10.1038/s41467-022-30949-6.

34. Huang, X.P., Ye, T.T., Zhang, L., Liu, R.F., Lai, X.J., Wang, L., Yang, M., Zhang, B., Li, X.Y., Liu, Z.W., et al. (2018). Multimodal unsupervised image-to-image translation. Surg. Oncol. 27, 172–176.

35. Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A.A. (2017). Image-to-image translation with conditional adversarial networks. Proceedings of the IEEE conference on computer vision and pattern recognition, 1125–1134.

36. He, K., Zhang, X., Ren, S., and Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. Proceedings of the IEEE international conference on computer vision, 1026–1034.

37. Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2012). Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems.

38. Shelhamer, E., Long, J., and Darrell, T. (2017). Fully convolutional networks for semantic segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39, 640–651. https://doi.org/10.1109/TPAMI.2016.2572683.

39. Zamir, S.W., Arora, A., Khan, S., Hayat, M., Khan, F.S., and Yang, M.-H. (2022). Restormer: Efficient transformer for high-resolution image restoration. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 5728–5739.

40. Zhang, H., Zhao, Y., Fang, C., Li, G., Zhang, M., Zhang, Y.-H., and Fei, P. (2020). Exceeding the limits of 3D fluorescence microscopy using a dual-stage-processing network. Optica 7, 1627–1640. https://doi.org/10.1364/OPTICA.402046.

41. Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III, 18Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III (Springer), pp. 234–241.

42. Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., and Liang, J. (2020). UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. IEEE Trans. Med. Imaging 39, 1856–1867. https://doi.org/10.1109/TMI.2019.2959609.

43. Kan, S., Zhang, Y., Zhang, F., and Cen, Y. (2022). A GAN-based input-size flexibility model for single image dehazing. Signal Process. Image Commun. 102, 116599.

44. Bieling, P., Laan, L., Schek, H., Munteanu, E.L., Sandblad, L., Dogterom, M., Brunner, D., and Surrey, T. (2007). Reconstitution of a microtubule plus-end tracking system in vitro. Nature 450, 1100–1105. https://doi.org/10.1038/nature06386.

45. Padiath, Q.S., Saigoh, K., Schiffmann, R., Asahara, H., Yamada, T., Koeppen, A., Hogan, K., Ptácek, L.J., and Fu, Y.H. (2006). Lamin B1 duplications cause autosomal dominant leukodystrophy. Nat. Genet. 38, 1114–1123. https://doi.org/10.1038/ng1872.

46. Nmezi, B., Xu, J., Fu, R., Armiger, T.J., Rodriguez-Bey, G., Powell, J.S., Ma, H., Sullivan, M., Tu, Y., Chen, N.Y., et al. (2019). Concentric organization of A- and B-type lamins predicts their distinct roles in the spatial organization and stability of the nuclear lamina. Proc. Natl. Acad. Sci. USA 116, 4307–4315. https://doi.org/10.1073/pnas.1810070116.

47. Xu, J., Ma, H., Jin, J., Uttam, S., Fu, R., Huang, Y., and Liu, Y. (2018). Super-Resolution Imaging of Higher-Order Chromatin Structures at Different Epigenomic States in Single Mammalian Cells. Cell Rep. 24, 873–882. https://doi.org/10.1016/j.celrep.2018.06.085.

48. Xu, J., Ma, H., Ma, H., Jiang, W., Mela, C.A., Duan, M., Zhao, S., Gao, C., Hahm, E.R., Lardo, S.M., et al. (2020). Super-resolution imaging reveals the evolution of higher-order chromatin folding in early carcinogenesis. Nat. Commun. 11, 1899. https://doi.org/10.1038/s41467-020-15718-7.

49. Zessin, P.J.M., Finan, K., and Heilemann, M. (2012). Super-resolution fluorescence imaging of chromosomal DNA. J. Struct. Biol. 177, 344–348. https://doi.org/10.1016/j.jsb.2011.12.015.

50. Ricci, M.A., Manzo, C., García-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. Cell 160, 1145–1158. https://doi.org/10.1016/j.cell.2015.01.054.

51. Nahidiazar, L., Kreft, M., van den Broek, B., Secades, P., Manders, E.M.M., Sonnenberg, A., and Jalink, K. (2015). The molecular architecture of hemidesmosomes, as revealed with super-resolution microscopy. J. Cell Sci. 128, 3714–3719. https://doi.org/10.1242/jcs.171892.

52. Miller, J.M., Wang, W., Balczon, R., and Dentler, W.L. (1990). Ciliary microtubule capping structures contain a mammalian kinetochore antigen. J. Cell Biol. 110, 703–714. https://doi.org/10.1083/jcb.110.3.703.

53. Camand, E., Peglion, F., Osmani, N., Sanson, M., and Etienne-Manneville, S. (2012). N-cadherin expression level modulates integrin-mediated polarity and strongly impacts on the speed and directionality of glial cell migration. J. Cell Sci. 125, 844–857. https://doi.org/10.1242/jcs.087668.

54. Yang, Y., Deng, G., Qiao, L., Yuan, H., Yu, X., Xu, L., Lu, S.H., Jiang, W., and Yu, X. (2022). Identification and characterization of stem cells in mammalian esophageal stratified squamous epithelia. J. Mol. Cell Biol. 14, mjac038. https://doi.org/10.1093/jmcb/mjac038.

55. Wang, Z., Bovik, A.C., Sheikh, H.R., and Simoncelli, E.P. (2004). Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13, 600–612. https://doi.org/10.1109/tip.2003.819861.

56. Culley, S., Albrecht, D., Jacobs, C., Pereira, P.M., Leterrier, C., Mercer, J., and Henriques, R. (2018). Quantitative mapping and minimization of super-resolution optical imaging artifacts. Nat. Methods 15, 263–266. https://doi.org/10.1038/nmeth.4605.

57. Descloux, A., Grußmayer, K.S., and Radenovic, A. (2019). Parameter-free image resolution estimation based on decorrelation analysis. Nat. Methods 16, 918–924. https://doi.org/10.1038/s41592-019-0515-7.

58. Descloux, A., Grußmayer, K.S., and Radenovic, A. (2020). Addendum: Parameter-free image resolution estimation based on decorrelation analysis. Nat. Methods 17, 1061–1063. https://doi.org/10.1038/s41592-020-0963-0.

59. Liu, W., Padhi, A., Zhang, X., Narendran, J., Anastasio, M.A., Nain, A.S., and Irudayaraj, J. (2022). Dynamic Heterochromatin States in Anisotropic Nuclei of Cells on Aligned Nanofibers. ACS Nano 16, 10754–10767. https://doi.org/10.1021/acsnano.2c02660.

60. Chapman, K.B., Filipsky, F., Peschke, N., Gelléri, M., Weinhardt, V., Braun, A., Hausmann, M., and Cremer, C. (2023). A comprehensive method to study the DNA's association with lamin and chromatin compaction in intact cell nuclei at super resolution. Nanoscale 15, 742–756. https://doi.org/10.1039/d2nr02684h.

61. Nakamura, S., Grigoriev, I., Nogi, T., Hamaji, T., Cassimeris, L., and Mimori-Kiyosue, Y. (2012). Dissecting the nanoscale distributions and functions of microtubule-end-binding proteins EB1 and ch-TOG in interphase HeLa cells. PLoS One 7, e51442. https://doi.org/10.1371/journal.pone.0051442.

62. Akhmanova, A., and Steinmetz, M.O. (2015). Control of microtubule organization and dynamics: two ends in the limelight. Nat. Rev. Mol. Cell Biol. 16, 711–726. https://doi.org/10.1038/nrm4084.

63. Zhang, R., Alushin, G.M., Brown, A., and Nogales, E. (2015). Mechanistic Origin of Microtubule Dynamic Instability and Its Modulation by EB Proteins. Cell 162, 849–859. https://doi.org/10.1016/j.cell.2015.07.012.

64. Weigert, M., Schmidt, U., Boothe, T., Müller, A., Dibrov, A., Jain, A., Wilhelm, B., Schmidt, D., Broaddus, C., Culley, S., et al. (2018). Content-aware image restoration: pushing the limits of fluorescence microscopy. Nat. Methods 15, 1090–1097. https://doi.org/10.1038/s41592-018-0216-7.

65. He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

66. Glorot, X., Bordes, A., and Bengio, Y. (2011). Deep sparse rectifier neural networks. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (JMLR Workshop and Conference Proceedings), pp. 315–323.

67. He, K., Zhang, X., Ren, S., and Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans. Pattern Anal. Mach. Intell. 37, 1904–1916. https://doi.org/10.1109/TPAMI.2015.2389824.

68. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., and Isard, M. (2016). TensorFlow: a system for Large-Scale machine learning. In 12th USENIX symposium on operating systems design and implementation (OSDI 16), pp. 265–283.

69. Kingma, D.P., and Ba, J. (2014). Adam: A Method for Stochastic Optimization. Preprint at arXiv. https://doi.org/10.48550/arXiv.1412.6980.

70. Wang, Z., Simoncelli, E.P., and Bovik, A.C. (2003). Multiscale Structural Similarity for Image Quality Assessment. The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003 (Ieee), pp. 1398–1402.

71. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nat. Methods 17, 261–272. https://doi.org/10.1038/s41592-020-0772-5.

72. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682. https://doi.org/10.1038/nmeth.2019.

73. Bolte, S., and Cordelières, F.P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. J. Microsc. 224, 213–232. https://doi.org/10.1111/j.1365-2818.2006.01706.x.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-α-tubulin	Sigma-Aldrich	Cat# T5168; RRID:AB_477579
Rat anti-α-tubulin	Sigma-Aldrich	Cat# MAB1864; RRID: AB_2890657
Mouse anti-EB1	BD Biosciences	Cat# 610535; RRID: AB_397892
Human anti-crest	Erba Diagnostics	Cat# HCT-0100, RRID:AB_2744669
Rabbit Anti-Cytokeratin 14 antibody	Abcam	Cat# ab181595, RRID:AB_2811031
Rabbit anti-H2B	Abcam	Cat# ab1790, RRID:AB_302612
Rabbit anti-H3K9me3	Abcam	Cat# ab8898, RRID:AB_306848
Rabbit anti-LaminB1	Abcam	Cat# ab16048, RRID:AB_443298
Alexa Fluor 647 phalloidin	Thermo Fisher Scientific	Cat# A22287, RRID:AB_2620155
Goat anti mouse Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-21236, RRID:AB_2535805
Goat anti rat Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-21247, RRID:AB_141778
Goat anti mouse IgG Atto 488	Sigma-Aldrich	Cat# 62197, RRID:AB_1137649
Goat anti rabbit IgG Alexa Fluor 546	Thermo Fisher Scientific	Cat# A-11010, RRID:AB_2534077
Goat anti human IgG Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-21445, RRID:AB_2535862
Goat anti mouse IgG Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-21236, RRID:AB_2535805
Goat anti rabbit cy5	Thermo Fisher Scientific	Cat# A10523, RRID:AB_2534032
Goat anti rat IgG(H + L) Alexa Fluor 568	Thermo Fisher Scientific	Cat# A-11077, RRID:AB_2534121
Chemicals, peptides, and recombinant proteins		
DMEM/F-12	Thermo Fisher Scientific	Cat# 11330032
RPMI 1640	Thermo Fisher Scientific	Cat# 11875119
DMEM	Thermo Fisher Scientific	Cat# 11965092
Penicillin-Streptomycin	Thermo Fisher Scientific	Cat# 10378016
Phosphate Buffer Saline	Thermo Fisher Scientific	Cat# 20012050
Trypsin-EDTA	Thermo Fisher Scientific	Cat# 25200072
DMSO	Sigma-Aldrich	Cat# D8418
Fetal bovine serum	Thermo Fisher Scientific	Cat# 10091148
Cholera Toxin	CELL technologies	Cat# Cc104
insulin	CELL technologies	Cat# Cc101
hydrocortisone	CELL technologies	Cat# Cc103
EGF	PEPROTECH	Cat# AF-100-15
Y27632	Topscience	Cat# T1725
paraformaldehyde (PFA)	Sigma-Aldrich	Cat# 158127
Glutaraldehyde	Sigma-Aldrich	Cat# G6257
NaBH4	Sigma-Aldrich	Cat# 71320
Triton X-100	Sigma-Aldrich	Cat# T8787
Bovine Serum Albumin	Solarbio	Cat# A8020
Goat serum	Solarbio	Cat# SL038
Sodium choride	Sigma-Aldrich	Cat# S9888
Tris-cl	Sigma-Aldrich	Cat# TRIS-RO
Glucose	Sigma-Aldrich	Cat# D9434
HCl	Sigma-Aldrich	Cat# 258148

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
acetone	Sigma-Aldrich	Cat# 179124
ethanol	Sigma-Aldrich	Cat# 51976
methanol	Sigma-Aldrich	Cat# 34860
glucose oxidase	Sigma-Aldrich	Cat# G7141
catalase	Sigma-Aldrich	Cat# C9322
β-mercaptoethanol	Sigma-Aldrich	Cat# M3148
mercaptoethylamine (MEA)	Sigma-Aldrich	Cat# 30070
NaOH	Sigma-Aldrich	Cat# 655104
NaCl	Sigma-Aldrich	Cat# S9888
$MgCl_2$	Sigma-Aldrich	Cat# M2393
Experimental models: Cell lines		
HeLa	ATCC	CRM-CCL-2
U87MG	ATCC	HTB-14
U2-OS	ATCC	HTB-96
RPE1	ATCC	CRL-4000
RNE-D3	Preserved by our lab	N/A
U373MG	ATCC	HTB-17
Software and algorithms		
ImageJ	National Institutes of Health (NIH)	https://ImageJ.nih.gov/ij
Python2.7	Python Software Foundation	https://www.python.org/download/releases/2.7/
Scipy package for Python	Scipy package for Python	https://www.scipy.org
Tensorflow deep learning framework for Python	Google Brain	https://tensorflow.google.cn/
Cuda	NVIDIA	https://developer.nvidia.com/cuda-downloads
CuDNN	NVIDIA	https://developer.nvidia.com/cudnn
GraphPad Prism 8	GraphPad	https://www.graphpad.com/scientific-software/prism/
Adobe Illustrator 2020	Adobe	https://www.adobe.com/
Adobe Photoshop 2019	Adobe	https://www.adobe.com/
Other		
Datasets and trained models for X-Microscopy	This paper	https://www.scidb.cn/en/s/BzuuMz
Source code	This paper	https://github.com/kanshichao/X-Microscopy

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Wei Jiang (wjiang6138@cicams.ac.cn).

Materials availability

This study did not generate new unique regents.

Data and code availability

- The datasets of X-Microscopy have been deposited at the Science Data Bank public platform, and DOIs are listed in the key resources table. They are available in this link: https://www.scidb.cn/en/s/BzuuMz. If you have any questions, please contact wjiang6138@cicams.ac.cn.
- All original code has been deposited at the Github and is publicly available as of the date of publication. DOIs are listed in the key resources table.

- Any additional information required to reanalyze the data reported in this paper is available from the lead contact (wjiang6138@cicams. ac.cn) upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Cell culture

Human cell lines were obtained from ATCC and cultured at 37°C with 5 % CO_2 in humidified incubators. HeLa, U87MG and U2-OS cells were cultured in DMEM supplemented with 10 % FBS and 1 % penicillin–streptomycin antibiotics. RPE1 cells were maintained in DMEM/F12 supplemented with 10 % FBS and 1 % penicillin–streptomycin. U373MG cells were grown in RPMI 1640 medium supplemented with 10 % FBS and 1 % penicillin/streptomycin. The normal rat oesophageal epithelial cell line RNE-D3 was established and preserved by our laboratory.[54] RNE-D3 cells were cultured in DMEM/F12 (3:1) supplemented with 10 % FBS, 8 ng/mL cholera toxin, 5 ng/mL insulin, 25 ng/mL hydrocortisone, 0.1 ng/mL EGF and 10 μM Y27632 in a humidified 37°C incubator supplemented with 5 % CO_2. All cell lines were authenticated and proven to be free of mycoplasma.

METHOD DETAILS

Immunofluorescence

For immunostaining, cells were cultured on glass-bottomed dishes (World Precision Instruments, Singapore) overnight and then washed three times with PBS. Cells were fixed either by adding 3 % paraformaldehyde and 0.1 % glutaraldehyde solution in PBS at room temperature for 10 minutes or by adding cold acetone and methanol (1:1 v/v) solution for 5 min. After washing with PBS three times, the cells were treated with freshly prepared 0.1 % $NaBH_4$ solution to reduce the autofluorescence from PFA fixation. Fixed samples were incubated overnight in blocking buffer (10 % goat serum, 3 % BSA, 0.2 % Triton X-100 in PBS). Primary antibodies (mouse anti-α-tubulin 1:50000, mouse anti-EB1 1:5000, human anti-CREST 1:3000, rabbit anti-Cytokeratin-14 1:1000, rabbit anti-H3K9me3 1:500, rabbit anti-H2B 1:500 and rabbit anti-Lamin B1 1:200) were added to the dish in blocking buffer. After 1-2 h incubation in room temperature or overnight in 4°C, the cells were washed five times with washing buffer (0.2 % BSA, 0.02 % Triton X-100 in PBS) with at least 15 minutes of incubation between washes. Secondary antibodies (goat anti-mouse/rat/human Alexa Fluor 647/cy5 and goat anti-mouse Atto-488) were added at a dilution of 1:200 and incubated for 2 hours in blocking buffer on an orbital shaker. Before STORM imaging, the storage buffer was switched to STORM imaging buffer containing 10% (w/v) glucose, 0.56 mg/mL glucose oxidase, 0.17 mg/mL catalase and 0.14 M β-mercaptoethanol, which were diluted in 50 mM Tris-Cl, pH=8.0, 10 mM NaCl, 10 % glucose. The glucose oxidase and catalase were diluted with 10mM Tris-Cl, pH=8.0 and 50mM NaCl, respectively.

For the drug treatments in Figure S12, U2-OS cells were seeded in 35 mm glass-bottomed dishes until the cell confluence reached 60 %. The cells were treated with different concentrations of nocodazole (100 nM) and Taxol (1 μM) for 30 min or treated with a vehicle control. After drug treatments, the cells were fixed and co-stained with anti-α-tubulin and anti-EB1 antibodies and the WFs were acquired using a 100✕ /1.49-NA objective lens.

WF and STORM image acquisition

Our microscopic system from Nikon Instruments was equipped with Nikon Eclipse Ti2 conventional microscopy and commercial N-STORM (Nikon) at the same time so that WFs and SRMs could be recorded by the shared scientific CMOS (sCMOS) camera. The WFs and SRMs were obtained by PlanApo TIRF 100✕/1.49 NA oil immersion objective (CFI SR HP Apochromat TIRF 100x/1.49NA, Nikon). All WF-SRM training pairs were well matched, and there was no need for pixel registration, which simplified the training process. WF images were recorded on 512 ✕ 512 or 256 ✕ 256 pixel before SRM imaging. For SRM imaging, raw SRM data were acquired under the highly inclined illumination of a 50 mW 647 nm laser using a Nikon N-STORM inverted microscope. All raw SRMs were recorded at 512 ✕ 512 or 256 ✕ 256 pixels. The temperature of the imaging environment was controlled at 23°C.

For STORM imaging of MTs or F-actin filaments, 50000-60000 frames were acquired at an exposure time of 10-20 ms. For STORM imaging of EB1 or kinetochores, 30000-50000 frames were acquired at an exposure time of 5-8 ms. For STORM imaging of Lamin B1 and Lamin A/C, 40000-50000 frames were recorded at an exposure time of 10 ms. For STORM imaging of H2B, 40000-50000 frames were recorded at an exposure time of 10-20 ms. For STORM imaging of H3K9me3, 40000-50000 frames were recorded at an exposure time of 10 ms. For STORM imaging of Keratin-14, 40000-50000 frames were recorded at an exposure time of 20 ms. All raw SRM images were reconstructed and analysed using the N-STORM analysis module in NIS-Elements AR Analysis 5.01.00 (Nikon). Quantitative information for each localization, including the x- and y-position, standard deviation (σ), background noise, and localization precision, is determined by this analysis modules according to N-STORM user manual.

To reduce image storage and increase the training speed, all images were exported and saved as 8-bit images by commercial NIS-Elements AR Analysis 5.01.00 software, as in previous SRM reconstruction work.[26,27,30] To evaluate the robustness of X-Microscopy in Figure 6, The WF images of microtubule were captured by Nikon Eclipse Ti2 microscopy equipped with a 100✕/1.49NA objective lens (CFI Apochromat TIRF 100✕/1.49NA) and another conventional microscope (Leica DM5000 B, Lecia Microsystems) using a 100✕/1.30NA objective lens (HCX PL FL100✕/1.30 Oil 0.17/D, Leica Microsystems).

The microtubule and mitochondrial imaging data in Figure S16 were kindly provided by NOVEL optics (China) and obtained from the NIB 900 and NSR950 microscopes equipped with a 100✕/1.49 NA objective lens.

The training and test datasets for UR-Net-8

No more than 30 FOVs (field of view) of different structures were acquired for training. For SRM reconstructions with deep learning, the size of the training dataset should be as large as possible to cover the distribution of images in the task domain for deep learning. To achieve good performance with a small-scale dataset, we therefore manually cropped the original image FOVs into smaller ROIs to generate high-quality training samples for all experiments. To ensure the variety of the training dataset, only one high-quality ROI was captured from each FOV and each FOV was acquired from a different cell. Many ROIs were used for training and each ROI contained 30 U-SRMs and 1 WF. The WFs were used as inputs, and the U-SRMs were defined as ground truth. To define U-SRMs, random subsets of 10000 consecutive frames from all the image sequences were created. The details of each training dataset and test dataset were as follows. We trained our MT model with 25 ROIs of STORM images and corresponding WFs. The F-actin model was established with 14 ROIs by transfer learning with the MT model. The number of training samples for the EB1 model was 15 ROIs. We trained the Lamin B1 model and the kinetochore model with 17 ROIs and 15 ROIs, respectively. The CK14 model was constructed with 15 ROIs. The H3K9me3 model was established with 6 ROIs, and the H2B model was trained with 8 ROIs. During training, a separate validation set was not specifically designated in our approach. The training set was utilized as the validation set to determine the model parameters. At the test stage, the test sets consisted of 14 ROIs for microtubules, 7 ROIs for F-actin, 17 ROIs for Lamin B1, 13 ROIs for EB1, 21 ROIs for kinetochores, 7 ROIs for CK14, 9 ROIs for H2B, and 5 ROIs for H3K9me3.

The training and test datasets for X-Net

To validate the feasibility of X-Net for reconstructing SRMs from U-SRMs and/or WFs, 25 ROIs of immunostained MT images were used for X-Net training. To define U-SRMs, random subsets of 300-500 consecutive frames from all the image sequences were created. To define W-SRMs, random consecutive subsets of 95 % of all available frames were reconstructed. Each ROI of images for training contained 30 W-SRMs, 30 U-SRMs, 1 WF and 1 F-SRM. For testing, only U-SRMs (k=300-500) and/or WFs were used as inputs, and the F-SRMs were used to calculate SSIM with the desired outputs. For Figure S3, 14 ROIs were used to test the X-Net, each ROI contained 1 WF and 6 U-SRMs.

To reconstruct high-fidelity SRMs from WFs, we used the MU-SRMs predicted from WFs by UR-Net-8 and/or original WFs as inputs for X-Net training. The detailed interpretation for each training dataset was similar to that of UR-Net-8.

X-Microscopy: The architecture of UR-Net-8

UR-Net-8, constructed based on skip-connection[41] and residual,[65] was devised for generating U-SRMs (Figure 1B). UR-Net-8 contains a generator network $\mathcal{G}_s$ and a discriminator network $\mathcal{D}_s$.'U' indicates U-connection, 'R' indicates residual, and '8' indicates 8 units of UR-Net used between the encoder and decoder (Figures S1A and S1B).

The generator network $\mathcal{G}_s$ comprised an encoder with 8 convolutional layers (encoding units) and a decoder with 8 deconvolutional layers (decoding units) (Figure 1B). A residual unit was utilized between each encoding unit and decoding unit. The architecture and operating principle of the residual module were illustrated as follows. In the encoder, the number of feature channels was increased constantly by the first four convolutional layers, while it was decreased by the last four deconvolutional layers in the decoder. From the input layer to the output layer, the spatial resolution of feature maps in the encoder was halved layer by layer, while it was doubled layer by layer in the decoder. In this way, the input of UR-Net-8 passed through a deeper convolutional layer of the encoder, and more important feature maps were output. Then, all output feature maps flowed directly into the adjacent encoding unit and the corresponding residual unit. Finally, the output of the residual unit flowed into the decoding unit. Convolutional kernels were applied to convolve feature maps in each convolutional layer. The feature maps of UR-Net-8 were adaptively zero-padded (i.e., the size of output feature maps was calculated based on the size of input image to decide whether zero-padding should be used). The kernel size of the convolutional layers in the encoder and the deconvolutional layers in the decoder was 5×5, and the stride was 2. Moreover, the convolutional layers in the residual units comprised convolutions of size 3×3 with stride 1 (Figure S1B). Importantly, each convolution was followed by a rectified linear unit (ReLU)[66] and batch normalization.

$\mathcal{D}_s$ contains 4 convolutional layers (i.e., s1, s2, s3, and s4), 1 spatial pyramid layer (s5) and 1 linear classification layer (s6). Here, s5 is a spatial pyramid pooling layer,[67] conferring the input-size flexibility by sampling feature maps of different input sizes to vectors of the same length for the discriminator. The input of $\mathcal{D}_s$ is a concatenation of the input WFs of $\mathcal{G}_s$ and the reconstructed U-SRMs along the channel dimension or a concatenation of two input images of $\mathcal{G}_s$ and the ground truth image along the channel dimension. The output of $\mathcal{D}_s$ can be determined by training data and predefined ground truth values (Figure S2D).

X-Microscopy: The architecture of X-Net

X-Net was constructed to have a generator network $\mathcal{G}_r$ and a discriminator network $\mathcal{D}_r$ for adversarial training (Figure 1C). The generator network $\mathcal{G}_r$ displayed two parallel networks with an X-shape to share information between encoders and decoders (Figure S1C). The two desired inputs of $\mathcal{G}_r$, which could be set as the MU-SRMs predicted from UR-Net-8 plus WFs, two same U-SRMs acquired from superresolution microscopy, two same WFs or U-SRMs plus WFs were brought into two encoders and output feature maps. Then, feature maps could be concatenated along the channel dimension and propagated to two decoders. Finally, the outputs were concatenated along the channel dimension, and the SRMs were reconstructed by a convolution operation with one convolutional layer.

The detailed architecture of X-Net is illustrated as follows. For the generator network $\mathcal{G}_r$, $\mathcal{G}_r$ comprised two encoders and two decoders. Each encoder and each decoder were devised with 8 convolutional layers and 8 deconvolutional layers, respectively (Figure S1C). The kernel sizes of the first two layers in the encoder and the last two layers in the decoder were 5×5, and the stride was 2. Other layers had convolutional kernels of size 3×3 and the stride of 2. The rectified linear unit (ReLU) activation function was used to comply with nonlinear mapping in all layers except for e8 and d1 (Tables S1 and S2). The batch-normalization function was applied to carry out fast convergence in all layers except e1 and d8. The procedures for extracting and processing feature maps of X-Net were similar to those of UR-Net-8. The architecture of discriminator $\mathcal{D}_r$ was akin to $\mathcal{D}_s$ (Figure S2D).

The attention and U-connection residual modules were the key design of X-Net. As shown in Figure S1C, they were established between each pair of encoding units ei and decoding units d(9-i) ($1 \leq i \leq 8$) based on skip connection (ei and di are marked in Figure S2A). They played an important role in concentrating on the major features and information extracted from WFs and MU-SRMs to reconstruct the main and detailed maps of the SRMs. The detailed architecture of the attention and residual modules was as follows.

Each attention module had two inputs (A and B in Figure S2B). For the attention between layer i of the decoder and layer 9-i of the encoder, A is the output of d(i-1), and B is the input of ei. The attention module consisted of 5 convolutional layers and 1 deconvolutional layer. Four convolutional layers (a1, a3, a5 and a8) had kernels of size 1×1 with stride 1. The kernel size of the convolutional layer (a2) was 2×2, and the stride was 2 (Table S3). The deconvolutional layer (a6) had convolutions of size 3×3 with stride 2. Importantly, we calculated the elementwise addition (a4) for the convolution outputs of A and B (the outputs of a3 and a2) and calculated the elementwise multiplication (a7) for the output of deconvolution (a6) and the convolution output of B (a2) (Figure S2B). The residual module had four inputs (A, B, C and D in Figure S2C) for the residual between layer i of the decoder and layer 9-i of the encoder. A is the output of d(i-1), and B and C were the inputs of e(9-i) in the upper and lower encoders, respectively. D was the output of the attention module located between layer i of the decoder and layer 9-i of the encoder. The residual unit included 1 deconvolutional layer (r1) and 1 convolutional layer (r6), where the kernel size of the deconvolutional layer was 3×3 and the stride was 2. The convolutional layer had kernels of size 1×1 with stride 1 (Table S4). One elementwise additional operation was used to calculate the residual between the output of the last convolutional layer (r6) and the input B.

The training strategy of X-Net

Experiments were conducted on a computer with 64 GB of memory and one 11 GB GTX Titan 1080 GPU. The algorithms were implemented on a 64-bit Ubuntu 16.04 operating system with Python 3.6.4 and the 1.13.1 version of TensorFlow.[68] X-Net was trained end-to-end using stochastic gradient descent (SGD) with Adam[69] and a batch size of 1 with 100,000 or more iterations (back-propagation steps). The momentum term of Adam was set to 0.5. The initial learning rate was set to 2×10^{-4}. During training, the parameters of the generator and discriminator were updated alternately. We updated the parameters of the discriminator once after updating the parameters of the generator 4 times. To speed up training, the model was first trained from scratch with fixed size images and then fine-tuned with different sizes of images. Training the network from scratch typically took from hours to days on a single GTX Titan 1080 GPU. Once trained, the network took only a few seconds to reconstruct SRMs with input-size flexibility.

In training the network, three loss functions were applied, i.e., L_1, L_2, and L_3 in Figure 1C. The training loss functions were a combination of MS-SSIM loss,[70] l_1 norm loss and conditional generative adversarial (cGAN) loss. The MS-SSIM loss was calculated by using 1 minus the MS-SSIM between the reconstructed SRM image and the ground truth image. The l_1 norm loss (l_{1-j}, $j \in \{1,2,3\}$) was the mean of the absolute difference between the reconstructed SRM image and the ground truth image. The cGAN loss served as a regularization technique to improve the performance and stability of the GAN training process in image superresolution tasks,[30,33] which contained one generator loss and one discriminator loss. The generator loss was calculated based on the output value of the discriminator. The discriminator loss l_{D-XMj} was the cross-entropy between the value of the ground truth (0 or 1) and the output value of the corresponding discriminator.

The MS-SSIM assessed the similarity index of the overall structure between the reconstructed SRMs $\widehat{O}$ and the ground truth image O, which was calculated as follows: an image was resized to different scales, and the structure $s_j(\widehat{O}, O)$, luminance $l_j(\widehat{O}, O)$, and contrast $c_j(\widehat{O}, O)$ were calculated at each scale. Here, $s_j(\widehat{O},O)$, $l_j(\widehat{O},O)$, and $c_j(\widehat{O},O)$ were formulated as follow:

$$s_j(\widehat{O}, O) = \frac{\sigma_{\widehat{O}O} + C_1}{\sigma_{\widehat{O}}\sigma_O + C_1},$$ (Equation 1)

$$l_j(\widehat{O}, O) = \frac{2\mu_{\widehat{O}}\mu_O + C_2}{\mu_{\widehat{O}}^2 + \mu_O^2 + C_2},$$ (Equation 2)

$$c_j(\widehat{O}, O) = \frac{2\sigma_{\widehat{O}}\sigma_O + C_3}{\sigma_{\widehat{O}}^2 + \sigma_O^2 + C_3},$$ (Equation 3)

where $C_1 = 0.00045$, $C_2 = 0.0001$ and $C_3 = 0.0009$. μ and σ are the mean and standard deviation in a Gaussian sliding window (11 × 11). Then, the MS-SSIM is defined as:

$$\mathrm{MS-SSIM}(\widehat{O}, O) = [l_M(\widehat{O}, O)]^{\alpha_M} \cdot \prod_{j=1}^{M} [c_j(\widehat{O}, O)]^{\beta_j} [s_j(\widehat{O}, O)]^{\gamma_j},$$ (Equation 4)

where M is the highest scale factor, and α_M, β_j and γ_j are used to adjust the relative weighting of the measurements at different scales (luminance, contrast and structure). Their values are set according to the ref. [70] in the experiments:

$$M = 5, \boldsymbol{\alpha} = \boldsymbol{\beta} = \gamma = [0.0448, 0.2856, 0.3001, 0.2363, 0.1333].$$

Finally, the MS-SSIM loss is defined as follows:

$$l_{ms-ssim} = 1 - MS - SSIM(\widehat{O}, O). \qquad \text{(Equation 5)}$$

The l_1 norm loss is defined as:

$$l_1 = \mathbb{E}(|\widehat{O} - O|), \qquad \text{(Equation 6)}$$

where $\mathbb{E}$ is the average function.

Suppose that the output of the discriminator is $\widehat{y}$, and the ground truth value is 0 and 1 for the input of $\widehat{O}$ and O, respectively. The discriminator loss l_D is defined as:

$$l_d = \mathbb{E}[\log(\widehat{y}(O)) + \log(1 - \widehat{y}(\widehat{O}))]. \qquad \text{(Equation 7)}$$

Moreover, in all the experiments, the generative adversarial loss for the generator is defaulted to be the same as all other GAN-based losses, defined as $\mathbb{E}[\log(1 - \widehat{y}(\widehat{O}))]$.

The training strategy of UR-Net-8

UR-Net-8 was trained based on cGAN.[35] The training loss function was the combination of MS-SSIM loss, l_1 norm loss and cGAN loss. The calculation formulas for these losses were the same as those calculation formulas in X-Net training.

In the training process of UR-Net-8, the parameter settings and training strategies were consistent with those of X-Net except for the following settings. The initial learning rate was set to 1×10^{-4}. We alternated between updating the parameters of the generator 10 times and updating the parameters of the discriminator once.

Performance comparisons with the leading methods

We compared our X-Microscopy with the following state-of-the-art methods recently developed in the literature: a deep-learning-based single-frame super-resolution microscopy (SFSRM) method and a deep-learning-based point-scanning super-resolution (PSSR) method (Figure S18). We implemented these algorithms using PyTorch on a single GeForce GTX 1080 GPU with 11GB of memory and used the Adam optimization algorithm for all experiments, similar to X-Microscopy. The detailed settings of these parameters were based on their published works.[23,32]

Input-size flexibility

To realize input-size flexibility, the input size and output size of each convolutional layer and each deconvolutional layer were calculated based on the size of the input image. Suppose that the size of the input image is h×w and the computed output size of the feature map is $H_i \times W_i$ at layer ei or di, where i is the index number of convolutional layers or deconvolutional layers. The size calculating functions were defined as follows:

$$H_i = \left\lfloor \frac{h-1}{2^i} + 1 \right\rfloor, W_i = \left\lfloor \frac{w-1}{2^i} + 1 \right\rfloor, \qquad \text{(Equation 8)}$$

where $\lfloor \cdot \rfloor$ is the floor function (e.g., $\lfloor 1.3 \rfloor = 1$, $\lfloor 1.9 \rfloor = 1$).

The hyperparameters of our networks were shown in Tables S1–S7. Tables S1 and S2 show the hyperparameters of the encoders and decoders in X-Net. Tables S2–S5 show the hyperparameters of the attention unit, the residual unit, and the discriminator in X-Net. Tables S6 and S7 show the hyperparameters of the encoder and decoder in UR-Net-8. Table S8 shows the hyperparameters of the residual module in UR-Net-8. In each table, the input size and output size of each layer were calculated based on the size of the input images. Table S9 shows the training and test details between ANNA-PALM and X-Net.

Data augmentation

Data augmentation was used to increase the amount or diversity of data and to improve the robustness of the trained model. It was performed at each iteration during training. One iteration contained three steps. (1) One image was randomly selected from all the training images and then cropped. (2) This cropped image was defined as input to the network for training. In this stage, the loss values were calculated between the output and the ground truth to assist the training process. (3) Based on the loss values output from the previous step, the parameters of the network were updated. After completing the above three steps, an iteration was completed.

Data augmentation was conducted after data preparation, and images smaller than 100×100 were directly trained without data augmentation. If the image size was larger than 100×100, data augmentation was used to process these images for training. Our detailed explanation of the data augmentation program is illustrated as follows:

During training, we first randomly sample 4 integer numbers between 0 and 49, denoted as w_0, h_0, w_1, and h_1 in each iteration. Then, for images with sizes larger than 100×100, we cropped a patch from the input images using the following formula:

$$I \leftarrow I[h_0 : h - h_1, w_0 : w - w_1], \tag{Equation 9}$$

where h and w are the height and width of the image tensor I. The arrowhead pointed to the left indicates that the values of I were updated. After random cropping, we normalized the values of the cropped patches between 0 and 1 by performing $I \leftarrow I/255.0$. Here, 255.0 means that the pixel range of our experimental images was 8 bits. It should be noted that we did not consider images with sizes smaller than 100×100 in our experiments. Suppose that the input image size was smaller than 100×100, e.g., 80×80. The randomly sampled integer numbers were $w_0 = 48$, $h_0 = 48$, $w_1 = 48$, and $h_1 = 48$. Finally, $w - w_1 = 32 < w_0$, $h - h_1 = 32 < h_0$. Note: the starting and ending coordinates of random cropped images were determined by a computer program rather than by human operation. If the starting coordinates randomly fell on the edge of the image without content, useless and low-quality content (background without any image content) in the cropped images would interfere with training performance. Hence, data augmentation was not applied for images with sizes below 100×100 due to the possibility of useless information from the image background.

Multicolour superresolution image alignments

We found that the multicolour SRMs reconstructed from WFs by X-Microscopy exhibited misalignments. This is because the range of one pixel of WF may map to a range of multiple pixels of the SRM image. Therefore, the multi-color super-resolution images obtained by image reconstruction algorithm sometimes be misaligned. To align these reconstructed images, an image alignment and calibration algorithm in the Fourier space was adopted. Suppose that two reconstructed image tensors I_1 and I_2 were misaligned and the height and width of them were h and w, the height and width of tensor I_2 were expanded to $h + m$ and $w + n$ by zero padding, respectively. Then the cross-power spectrum was computed between I_1 and I_2 using the fft2 and ifft2 functions implemented in scipy package[71] and the maximum value was set as the peak correlation. Furthermore, the position of the peak correlation was the aligned position exactly. Finally, we moved one tensor to the matched position to align another tensor. This method could effectively align multicolour SRM.

QUANTIFICATION AND STATISTICAL ANALYSIS

During the test stage, quantitative metrics were used to evaluate the performance of X-Microscopy. The resolution of each cropped image was estimated using the Image Decorrelation Analysis[57] plugin of ImageJ.[72] For SSIM, we used the F-SRMs as the ground truth for calculation.

The SSIM between image $\widehat{O}$ and image O was calculated as follows:

$$\mathrm{SSIM}(\widehat{O}, O) = \frac{(2\mu_{\widehat{O}}\mu_O + C_2)(2\sigma_{\widehat{O}O} + C_3)}{(\mu_{\widehat{O}}^2 + \mu_O^2 + C_2)(\sigma_{\widehat{O}}^2 + \sigma_O^2 + C_3)}, \tag{Eqaution 10}$$

where $\mu_{\widehat{O}}$ and μ_O are the mean values of the pixels of image $\widehat{O}$ and image O, respectively. $\sigma_{\widehat{O}}$ and σ_O are the standard deviations of the pixels of image $\widehat{O}$ and image O, respectively. We applied a Gaussian sliding window (11×11, $\sigma = 1.5$) to move pixel by pixel across the whole image to calculate the SSIM value. At each position of the Gaussian window, the SSIM value was calculated individually. The final SSIM value was averaged from the values of all window positions. Note that the SSIM metric was used only to evaluate structural similarity during the test. During the training stage, we calculated the MS-SSIM loss by applying the SSIM metric to the reconstructed SRM and ground truth. The MS-SSIM metric is described above in detail.

For the Pearson r[73] value, suppose that we have two images X and Y, each with n pixels. We calculated the Pearson r value based on the following formula:

$$r(X, Y) = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2 \sum_{i=1}^{n} (Y_i - \overline{Y})^2}}, \tag{Equation 11}$$

where image X and image Y grey values of voxel will be noted as X_i and Y_i, respectively, and the corresponding average intensities over the full image as $\overline{X}$ and $\overline{Y}$.

The intensity of the line profile and colour boxes were measured using the Plot Profile (plugin of ImageJ). The error maps and RSE (resolution-scaled error) and RSP (resolution scaled Pearson coefficient) scores were determined by NanoJ-Squirrel.[56] Statistical analyses were performed using GraphPad Prism 8.0.