

Instant multicolor super-resolution microscopy with deep convolutional neural network

Songyue Wang^{1,2}, Chang Qiao³, Amin Jiang^{1,4}, Di Li¹, Dong Li^{1,2,5} ✉

¹ National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

² College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

³ Department of Automation, Tsinghua University, Beijing 100084, China

⁴ Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei 230026, China

⁵ Bioland Laboratory (Guangzhou Regenerative Medicine and Health Guangdong Laboratory), Guangzhou 510005, China

Received: 9 June 2021 / Accepted: 21 July 2021

Abstract Multicolor super-resolution (SR) microscopy plays a critical role in cell biology research and can visualize the interactions between different organelles and the cytoskeleton within a single cell. However, more color channels bring about a heavier budget for imaging and sample preparation, and the use of fluorescent dyes of higher emission wavelengths leads to a worse spatial resolution. Recently, deep convolutional neural networks (CNNs) have shown a compelling capability in cell segmentation, super-resolution reconstruction, image restoration, and many other aspects. Taking advantage of CNN's strong representational ability, we devised a deep CNN-based instant multicolor super-resolution imaging method termed IMC-SR and demonstrated that it could be used to separate different biological components labeled with the same fluorophore, and generate multicolor images from a single super-resolution image *in silico*. By IMC-SR, we achieved fast three-color live-cell super-resolution imaging with ~100 nm resolution over a long temporal duration, revealing the complicated interactions between multiple organelles and the cytoskeleton in a single COS-7 cell.

Keywords Multicolor imaging, Super-resolution, Convolutional neural network

INTRODUCTION

There are countless complicated yet well-regulated interactions between the organelles and the cytoskeleton in every cell, which are aimed at maintaining cellular homeostasis and executing various physiological functions (Phillips and Voeltz 2016; Prinz 2014). Observing the intracellular dynamics of multiple

organelles and the cytoskeleton is of vital importance for cell biology. Unfortunately, there are still several barriers for long-term multicolor and high spatial-temporal resolution live-cell imaging, which mainly lie in the following two aspects: the limited spatial resolution by optical diffraction theory and the negative effects created by excitation lasers, *i.e.*, photobleaching and phototoxicity. To break through the optical diffraction limits, several super-resolution methods have been proposed, *e.g.*, structured illumination microscopy (SIM) (Gustafsson *et al.* 2008), stimulated emission depletion microscopy (STED) (Willig *et al.* 2006), photoactivated localization microscopy (PALM)

Songyue Wang, Chang Qiao and Amin Jiang contributed equally to this work.

✉ Correspondence: lidong@ibp.ac.cn (D. Li)

(Shroff *et al.* 2008) and stochastic optical reconstruction microscopy (STORM) (Rust *et al.* 2006). However, although superior spatial resolution is achieved, these SR methods suffer from respective defects, *i.e.*, repetitive acquisitions for SIM, heavy photon budget for STED, and low temporal resolution for PALM and STORM. Furthermore, when performing multicolor imaging, one region of interest (ROI) will be illuminated two or more times for each time point, thus causing heavier degradation of cytoactivity and restricting the maximal frame rate, which distinctly limits the widespread adoption of various SR methods in multicolor live-cell imaging experiments.

In recent years, deep learning has become an increasingly popular approach to be applied in multiple biomedical imaging processing or analysis tasks, *i.e.*, extracting quantitative information from biological images (Isensee *et al.* 2021; Kermany *et al.* 2018; Moen *et al.* 2019; Pan *et al.* 2019), performing restoration or cross-modality transformation for microscopic images (Wang *et al.* 2019; Weigert *et al.* 2018) and reconstructing super-resolution images from a reduced number of corrupted raw data (Qiao *et al.* 2021b). In particular, with the help of elaborate convolutional neural network models such as U-net (Ronneberger *et al.* 2015), researchers can perform cell or organelle segmentation with high accuracy and efficiency (Isensee *et al.* 2021; Ronneberger *et al.* 2015), indicating that deep convolutional neural networks have a significant capability for image feature extraction and representation. Inspired by this, we made a step forward in using a deep CNN model to directly separate various subcellular structures from a single SR image, rather than only performing segmentation as others, and then recombining them again, which is equivalent to an instant multicolor SR (IMC-SR) procedure. The major benefits of this multicolor imaging framework are threefold: first, all of the biological structures can be labeled in the same fluorescent channel of relatively short wavelength, *i.e.*, the green channel, thus simplifying the sample preparation and potentially enhancing the spatial resolution; second, the IMC-SR scheme avoids repetitive exposure and acquisitions for each channel, which can relieve photobleaching and phototoxicity; third, the contents of each channel are captured simultaneously, thereby avoiding misinterpretations due to conventional sequential multicolor acquisitions. Furthermore, considering that U-net is a kind of wide but shallow neural network model, and it has been found that deeper CNN models may lead to a better performance than shallow models, especially when dealing with computationally complicated tasks (Lim *et*

al. 2017), we adopted a modified deep CNN model based on residual learning (Szegedy *et al.* 2017) and a channel attention mechanism (Hu *et al.* 2018; Zhang *et al.* 2018) to perform multicolor image separation.

In this work, we first presented the conception and workflow of our IMC-SR method and explored and optimized the application strategies of our deep CNN model. Then, we performed statistical evaluation according to the application strategies, different types of subcellular structures and input signal-to-noise-ratio (SNR) conditions. Finally, we demonstrated its effectiveness in experimental usage by observing three different types of intracellular structures simultaneously in a single COS-7 cell of ~100 nm resolution over a long temporal duration, visualizing the complicated interactions between different organelles and cytoskeleton.

RESULTS

Network architecture

Our IMC-SR neural network separates components of different types of biological structures from a single monochrome grayscale SR image in which all the information is superimposed. Compared with cell or organelle segmentation tasks, the separation of different biological components is a more complicated and underdetermined problem because separation means the network needs to determine the intensity values of every pixel for each output channel instead of only providing a binary mask in segmentation tasks. Furthermore, SR images contain many subtler details that are finer than the diffraction limits, thus increasing the structure separation difficulty. Therefore, we constructed our CNN model following two criteria: first, the network model should be deep enough to be competent in informative feature extraction and representation; second, the network model should be trainable and computationally efficient. Inspired by the SOTA single image super-resolution architecture residual channel attention network (RCAN) (Zhang *et al.* 2018), we devised a shallower version of RCAN, especially for multicolor image separation (hereafter named the IMC-SR network), which exhibits higher robustness and efficiency than the RCAN baseline model in our tasks.

Our IMC-SR network begins with a conv-GELU module for shallow feature extraction, which consists of one 3×3 convolutional layer and a GELU activation layer (Hendrycks and Gimpel 2016). The output of the conv-GELU module is fed into five consecutive residual

groups (RG). Each RG is composed of ten residual channel attention blocks (RCABs) and a skip connection pointing from the RG input to the output of the last RCAB. RCAB is the core component of the IMC-SR network, which consists of two conv-GELU modules and a channel attention module. In channel attention modules, each feature channel is adaptively rescaled according to the interdependency of all the feature channels, thus rendering the network focused on more meritorious information (Zhang *et al.* 2018). The output of the last RG is followed by an extra conv-GELU module to reorganize the feature channels, and a convolutional layer with sigmoid activation is used to generate the final monochrome image containing specific types of biological structure.

Data acquisition and preprocessing

Training and testing datasets of the endoplasmic reticulum (ER), microtubules (MTs) and lysosomes (Lyso) were acquired using our home-built multimodality structured illumination microscopy (TIRF-SIM and GI-SIM) (Guo *et al.* 2018; Li *et al.* 2015). For each type of specimen, we acquired 30 sets of raw SIM images (512×512) of different regions of interest (ROIs) and reconstructed them into 30 SR-SIM images (1024×1024) using a conventional SIM reconstruction algorithm (Gustafsson *et al.* 2008; Li *et al.* 2015). Each SR-SIM image was then augmented into 50 image patches of 128×128 pixels by flipping, random cropping and random rotation, constituting three sets of image patches corresponding to three types of biological specimens. To obtain paired data for training and testing, we randomly selected one image patch from each patch set separately and added them together as the input to the IMC-SR network. Correspondingly, the three original SR image patches serve as the ground truth (GT) in the training process.

Network training and evaluation

To evaluate the effectiveness of our IMC-SR method, we trained three individual IMC-SR network models for separating ER, MTs and Lyso from a single superimposed image of these three subcellular structures. The objective function we used in the training process was the combination of mean square error (MSE, M) and structural similarity (SSIM, S) as follows:

$$L(\hat{y}, y) = \lambda \cdot M(\hat{y}, y) + \mu \cdot S(\hat{y}, y), \quad (1)$$

where “ λ ” and “ μ ” are scalar weighting factors set as 1 and 0.1 in our experiments, respectively, “ $\hat{y}$ ” represents the output image of the network, and “ y ” is the

ground-truth image.

After being well trained, the three IMC-SR models were applied to the same input SR image containing both ER, MTs and Lyso. It can be clearly seen that our IMC-SR network models successfully separated the tubular structure of MTs, the reticulate structure of ER and the punctate structure of Lyso, and these details were observed together in an integrated multicolor image (Fig. 1). To statistically evaluate the performance of IMC-SR methods, we calculated the normalized root mean square error (NRMSE), SSIM and Pearson correlation coefficient (PCC) on 100 output images and 100 corresponding GT images for each model (Fig. 2A). The upper and lower whiskers of the Tukey box-and-whisker plots show the performance boundary of each model, and the median values (lines in the middle of the boxes) indicate that IMC-SR methods achieve fairly good multicolor separation quality. Although the output images of IMC-SR methods were perceptually satisfactory in general, there did exist, however, some noticeable false-positive errors in the separated images (yellow arrows in Fig. 2B).

Combination loss to alleviate false-positive errors

After carefully comparing the outputs of IMC-SR modes with the GT images, we found that for a certain model, false-positive errors are actually the misrecognition of contents from other color channels. To alleviate them, we explored a new training strategy, that is, training a sole model directly outputting all channels of biological structure (hereafter termed IMC-SR-All) instead of training models for every single type of structure (hereafter termed IMC-SR-Single). The most conspicuous benefit of this strategy is that we also utilized the original input image to supervise the iterative updating during the training procedure (hereafter termed summation loss) with the objective function formulated as

$$L(x, \hat{y}, y) = \lambda \cdot M(\hat{y}, y) + \mu \cdot S(\hat{y}, y) + \alpha \{ \lambda \cdot M[\text{sum}(\hat{y}), x] + \mu \cdot S[\text{sum}(\hat{y}), x] \}, \quad (2)$$

where “ $\text{sum}(\cdot)$ ” indicates the channelwise summation operator, “ x ” is the input image of the IMC-SR-All model and “ α ” is another weighting scalar to balance the summation loss, which is empirically set as 0.02 in our experiments.

To evaluate the practicability and robustness of IMC-SR-All under different input SNR conditions, we trained one IMC-SR-All model for clathrin-coated pits (CCPs), MTs and ER using the data from our previously published dataset BioSR (Qiao *et al.* 2021b) following the training procedure described above. It is

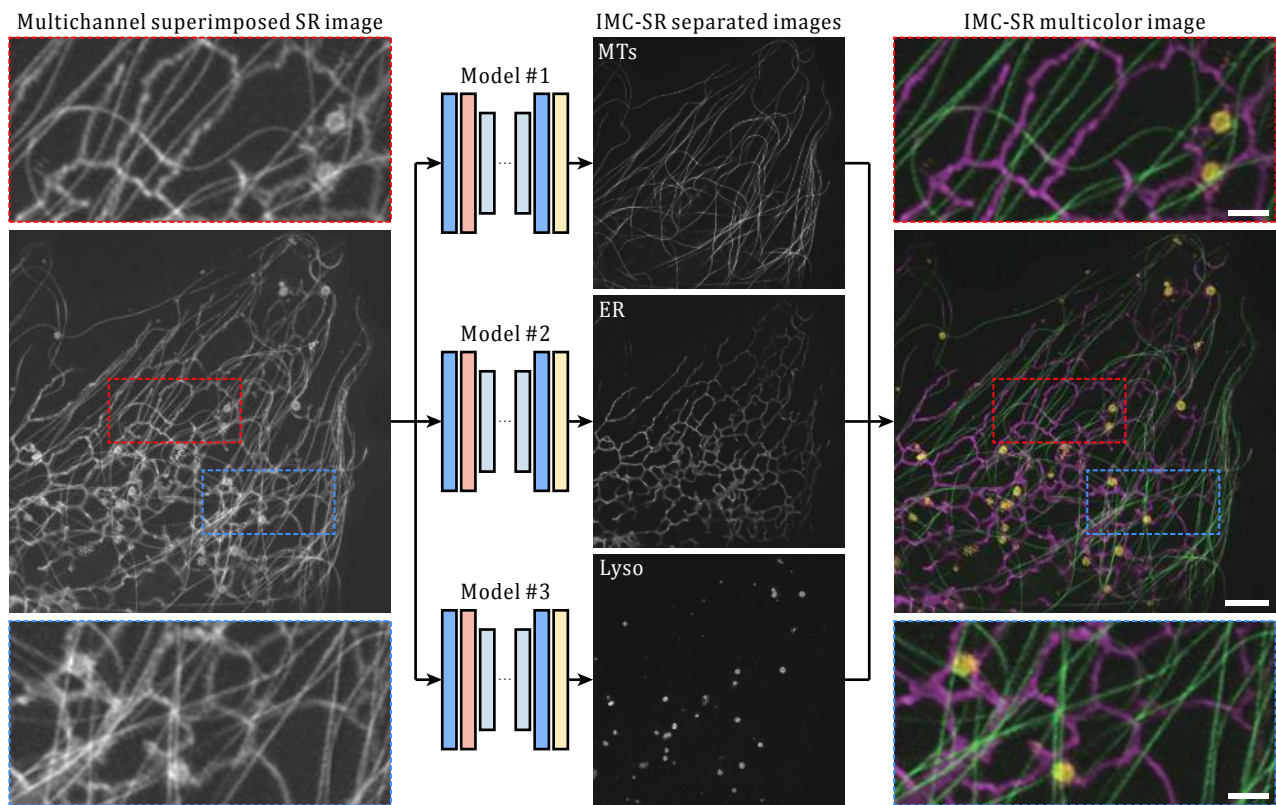


Fig. 1 Schematic and representative results of IMC-SR methods

noteworthy that we trained only one IMC-SR-All model for separating all subcellular structures and SNR conditions, which covered medium SNR (relatively low quality with slight artifacts after SIM reconstruction), high SNR (most common imaging conditions in daily experiments) and very high SNR (superior quality without reconstruction artifacts). We also trained three IMC-SR-Single models for CCPs, MTs and ER individually for comparison. The results shown in Fig. 3A indicate that with the regularization of the combination loss described in Eq. 2, the false-positive errors were reduced considerably (arrows in Fig. 3A), and the statistical analysis in terms of NRMSE, SSIM and PCC in Fig. 3B shows that the IMC-SR-All model achieved better performance than IMC-SR-Single models in all SNR conditions.

Enhancing performance of IMC-SR for time-lapse data by virtue of temporal continuity

We next determined that IMC-SR-All network models, trained solely on static images, can be used to separate the intracellular structure of time-lapse data, enabling instant multicolor live-cell imaging. To further enhance the performance of IMC-SR methods when dealing with

time-lapse data, we harnessed temporal continuity (TC) to conditionally instruct the training and inference progress of network models. For each training or inference trial, IMC-SR models were fed with three multichannel-superimposed frames of sequential timepoints but only performed channel separation for the median timepoint (Fig. 4A). Resorting to this scheme, IMC-SR models resolved the movements and interactions of organelles and cytoskeleton more stably along the temporal dimension without undesired blinks (arrows in Fig. 4B). We demonstrated that our IMC-SR method achieved instant three-color SR imaging by visualizing the delicate interaction dynamics of MT, ER and Lyso in COS-7 cells (Fig. 4C, supplementary Video 1). One time-lapse example shows that a lysosome moves along MTs and continuously contacts the ER, concomitantly changing the morphological distribution of the ER (Fig. 4D) and verifying its hitchhiking remodeling mechanism (Guo *et al.* 2018).

DISCUSSION

In this work, we introduced IMC-SR, a deep CNN-based analytical method, to perform channel separation

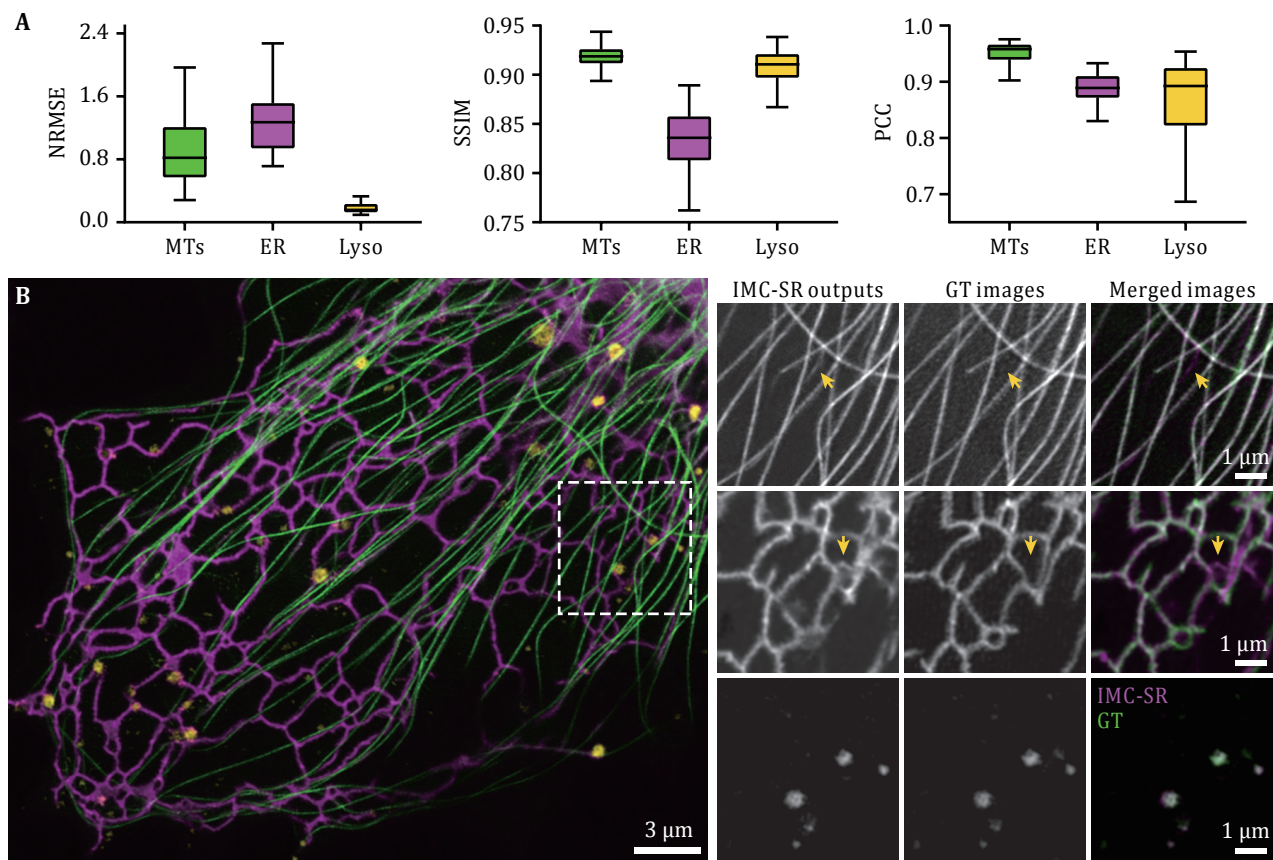


Fig. 2 Evaluation and representative results of IMC-SR-Single models. **A** Statistical analysis of IMC-SR-Single models trained for MT, ER and Lyso in terms of NRMSE, SSIM and PCC. **B** Representative results of IMC-SR-Single models trained for MT (green), ER (magenta) and Lyso (yellow). Yellow arrows indicate the false-positive errors of network output images. Scale bar: 3 μ m and 1 μ m (zoom-in regions)

directly from a single multistructure superimposed SR image, enabling fast multicolor super-resolution live-cell imaging. To improve the performance of IMC-SR models, we devised the combination loss that leverages the original input grayscale but informative images to regularize the training process of the model, thus effectively reducing the false-positive errors; to utilize the temporal continuity of time-lapse data, we conditionally provided the IMC-SR models with multiple consecutive timepoints as its input during training and inference procedures, leading to a more robust performance when handling time-lapse data in daily experiments. Finally, we applied the IMC-SR method to achieve instant three-color super-resolution live-cell imaging of subcellular structures at 2-second intervals over a duration of 100 timepoints, which reveals the sophisticated interaction between different organelles and the cytoskeleton.

Furthermore, although our IMC-SR method was only demonstrated and evaluated based on linear SIM data in this paper, apparently, it is appropriate for SR images of other imaging modalities, *e.g.*, nonlinear SIM, STED,

STORM, and PALM. Moreover, IMC-SR methods have the potential to perform channel separation and SR image restoration simultaneously if trained with artifact-corroded inputs, *i.e.*, SIM SR images reconstructed from noise raw images and corresponding clear GT images of high quality, which could be a desirable application when there is difficulty in acquiring ideal multichannel-superimposed data for network input.

Despite the inspiring multicolor imaging performance, it is worth noting that there are some restrictions for IMC-SR methods. First, if the biological structures to be separated in the input SR image are too similar to each other, *i.e.*, punctate CCPs and lysosomes, IMC-SR models will struggle to tell them apart. Second, in live-cell imaging experiments, the signal intensities of different biological structures are not always approximately comparable due to the susceptible expression of fluorescent protein, in which circumstance IMC-SR models cannot correctly identify certain structures if their fluorescence is too much weaker than others in a single sCMOS frame. Third,

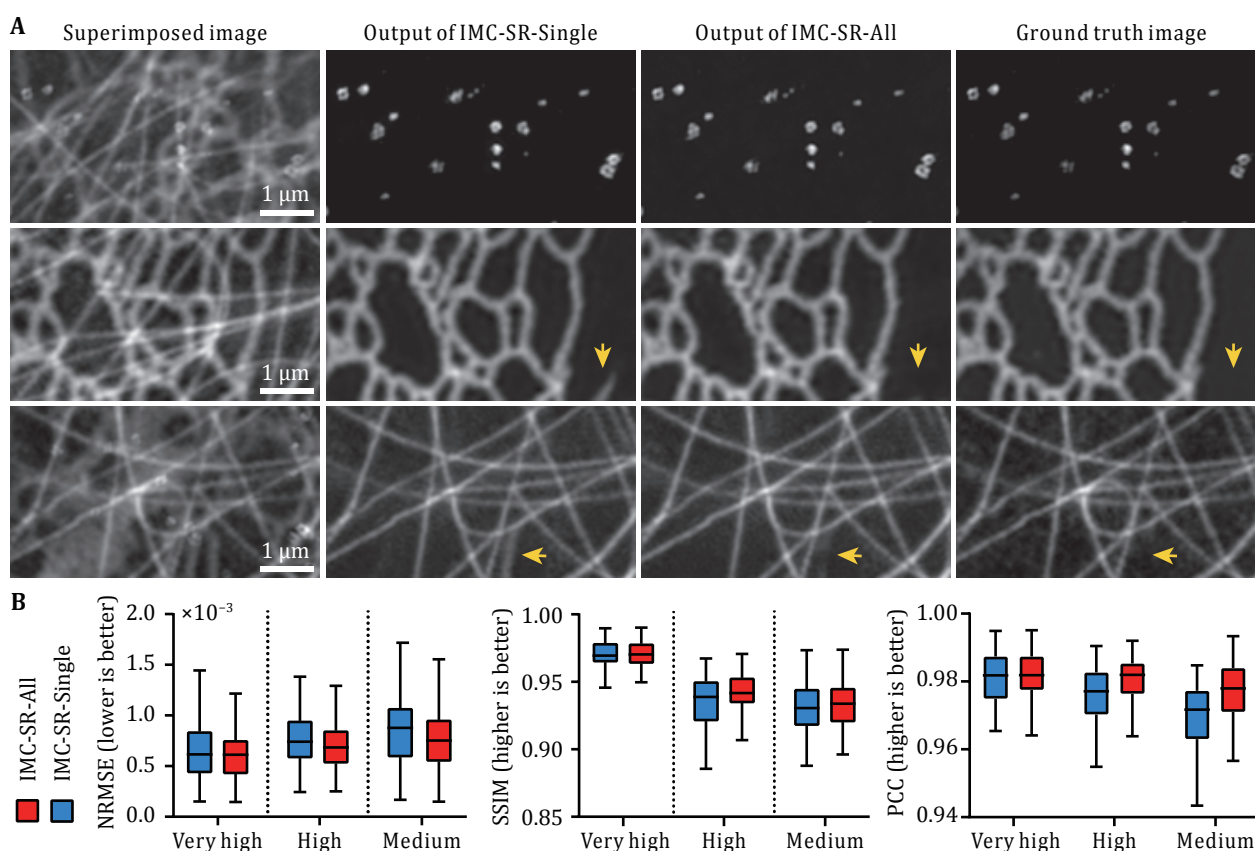


Fig. 3 Comparisons of IMC-SR-All models and IMC-SR-Single models. **A** Representative results of the IMC-SR-All model and IMC-SR-Single models trained for CCPs, ER and MTs. Arrows show that the false-positive errors can be reduced considerably with combination loss. Scale bar: 1 μ m. **B** Statistical analysis of the IMC-SR-All model and IMC-SR-Single models trained for ER in terms of NRMSE, SSIM and PCC at different input SNR conditions

imputing to the limited generalization ability, which is the case for the majority of CNN-based methods, individual IMC-SR models should be trained for every different combination of biological structures.

Equipped with IMC-SR methods, traditional multicolor fluorescence microscopy can achieve faster multicolor imaging of high spatiotemporal resolution at prominently lower photobleaching and phototoxicity. We hope that the superior performance and extensive application potential of the IMC-SR method can shed light on new discoveries in life science research necessitating multicolor SR imaging in the future.

METHODS

Detailed information of network training and testing

The IMC-SR models in this paper were trained and applied on a computer workstation equipped with a

Xeon Gold 6134 CPU@ 3.20 GHz (Intel) and two RTX 3090Ti graphic processing cards (NVIDIA) using Python 3.6, Tensorflow 2.4.0. In the training phase, the dataset was augmented into 20,000 image patch pairs of 128×128 pixels. The initial learning rate was set to 1×10^{-4} , and reduced by half if the validation NRMSE doesn't decrease for 10,000 iterations. We adopted a batch size of 3, and the training process typically took 10 hours and $\sim 200,000$ mini batch iterations. Once trained, it typically takes less than 250 milliseconds to process a multi-channel superimposed SR image of 512×512 pixels for both IMC-SR-All models and IMC-SR-Single models.

Calculation of different evaluation metrics

In this paper, we used three quantitative metrics, NRMSE, SSIM and PCC to evaluate the performance of IMC-SR models according to GT. Before calculating them, we performed the percentile normalization and linear transformation (LT) for IMC-SR model output

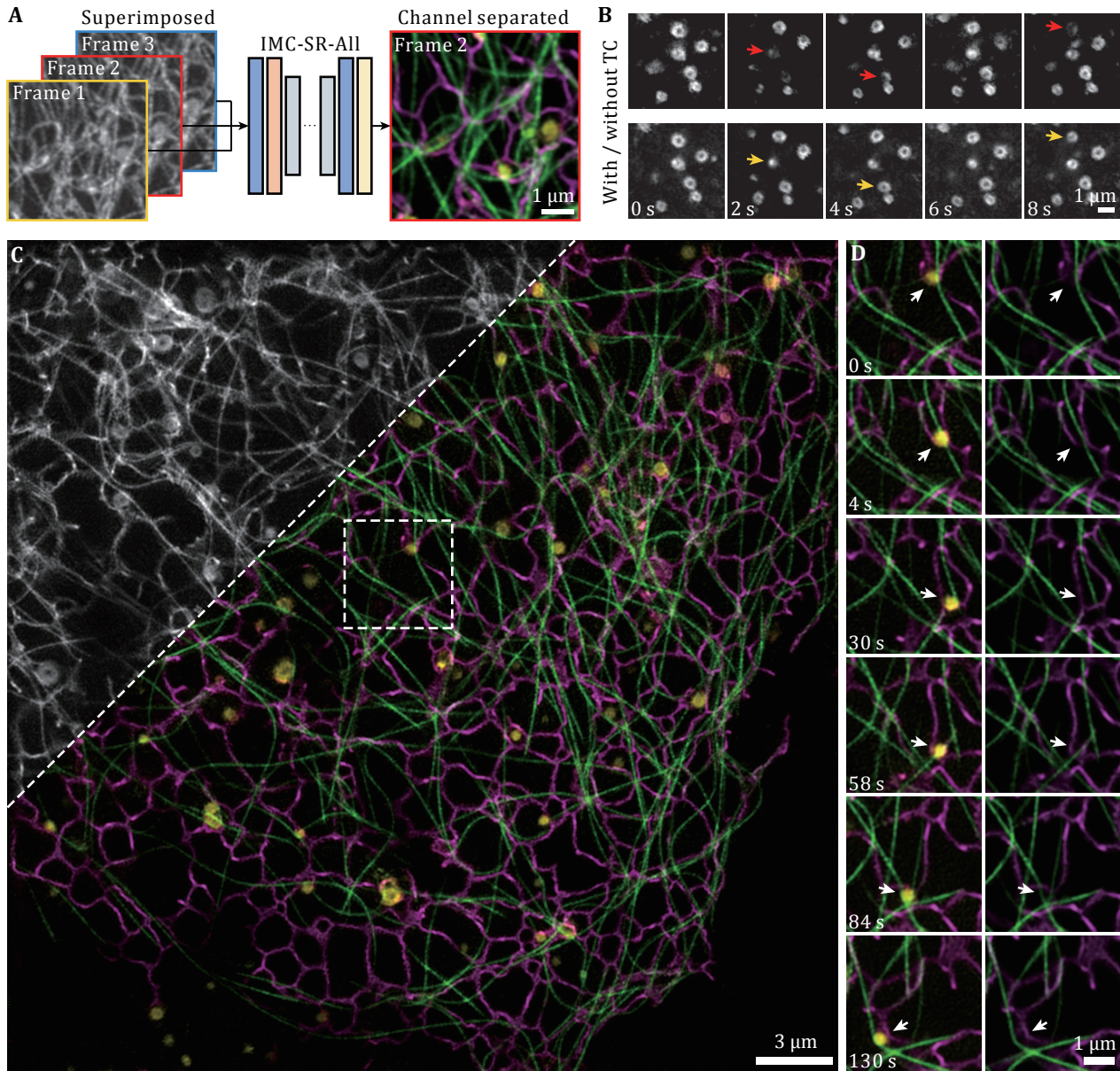


Fig. 4 Instant three-color live-cell SR imaging by IMC-SR. **A** Schematic of IMC-SR-All models applied with temporal continuity (TC). **B** Comparisons of IMC-SR-All models applied with (lower row) or without (upper row) temporal continuity. **C** Representative three-color images of MT (green), ER (magenta) and Lyso (yellow) generated by the IMC-SR-All model. Top left: a fraction of the corresponding multichannel-superimposed input image. **D** Time-lapse three-color images showing a lysosome moving along MTs that continuously contacted the ER and the hitchhiking remodeling mechanism of the ER. Gamma value: 0.8 for MT channel (**C** and **D**). Scale Bar: 1 μ m (**A**, **B** and **D**), 3 μ m (**C**)

images and GT images at first. This is a commonly used data preprocessing procedure (Qiao *et al.* 2021a; Weigert *et al.* 2018), which can be formulated as

$$Y_N = \frac{Y - \text{perct}(Y, p_{\text{low}})}{\text{perct}(Y, p_{\text{high}}) - \text{perct}(Y, p_{\text{low}})}, \quad (3)$$

$$\hat{Y}_N = \frac{\hat{Y} - \text{perct}(\hat{Y}, p_{\text{low}})}{\text{perct}(\hat{Y}, p_{\text{high}}) - \text{perct}(\hat{Y}, p_{\text{low}})}, \quad (4)$$

$$\hat{Y}_T = \alpha \hat{Y}_N + \beta, \quad (5)$$

where “ $\hat{Y}$ ”, “ $\hat{Y}_N$ ”, “ $\hat{Y}_T$ ”, “ Y ”, “ Y_N ” indicate the network output image, percentile normalized output image, linear transformed output image, GT image and percentile normalized GT image, respectively; “ $\text{perct}(Y, p)$ ” represents the intensity value of “ Y ” ranking “ $p\%$ ” with “ p_{low} ” and “ p_{high} ” are 0.1 and 0.99 in

our experiments; “ α ” and “ β ” are the hyper parameters of LT, which can be determined by solving the least square problem:

$$\arg \min_{\alpha, \beta} \|\alpha \hat{Y}_N + \beta - Y_N\|_2. \quad (6)$$

Then the NRMSE, SSIM and PCC can be calculated as:

$$\text{NRMSE}(\hat{Y}_T, Y_N) = \sqrt{\frac{1}{w \times h} \sum_{i=1}^{w \times h} (\hat{Y}_{T,i} - Y_{N,i})^2}, \quad (7)$$

$$\text{SSIM}(\hat{Y}_T, Y_N) = \frac{(2\mu_{\hat{Y}_T} \mu_{Y_N} + c_1)(\sigma_{\hat{Y}_T Y_N} + c_2)}{(\mu_{\hat{Y}_T}^2 + \mu_{Y_N}^2 + c_1)(\sigma_{\hat{Y}_T}^2 + \sigma_{Y_N}^2 + c_2)}, \quad (8)$$

$$\text{PCC}(\hat{Y}_T, Y_N) = \frac{\sigma_{\hat{Y}_T Y_N}}{\sigma_{\hat{Y}_T} \sigma_{Y_N}}, \quad (9)$$

where “ μ_X ”, “ σ_X ” and “ σ_{XY} ” indicate the mean value of “ X ”, standard deviation of “ X ” and the covariation of “ X ” and “ Y ”; “ c_1 ” and “ c_2 ” are non-zero constants of small value used to prevent the denominator from being close to zero, which are set to 0.01 and 0.03 by default in MATLAB 2017b.

Imaging conditions of live-cell experiments

In the three-color live-cell SR imaging experiment shown in Fig. 4, we acquired raw images in SIM mode of three-phase $\times$ three-orientation for 100 frames at 2-s intervals. Each raw image was acquired with a 488 nm illumination intensity of 40 W/cm² and 30 ms exposure time for 3xmEmerald-Ensconsin, 560 nm illumination intensity of 15 W/cm² and 30 ms exposure time for KDEL-mCherry, and 640 nm illumination intensity of 18 W/cm² and 30 ms exposure time for LAMP1-Halo.

Cell culture, transfection and staining

The COS-7 cell line was grown in Dulbecco’s Modified Eagle Medium (DMEM) (Gibco, Cat. No.: 11965-092) and supplemented with 10% fetal bovine serum (Gibco, Cat. No.:16140071) and 1% penicillin/streptomycin at 37 °C and 5% CO₂ until 60%–80% confluency was reached. After coating 25-mm coverslips with 50 mg/mL collagen for 1 h, cells were seeded onto the coverslips to achieve ~70% confluence before transfection. The manufacturer’s protocol of Lipofectamine 3,000 (Invitrogen) was employed for transient transfections. The plasmid constructs used in

this study include Clathrin-mEmerald, 3xmEmerald-Ensconsin, KDEL-mCherry and LAMP1-Halo.

Statistics and reproducibility

Figures 2A and 3B were plotted in Tukey box-and-whisker format. The box plots extend from the 25th and 75th percentiles and the line in the middle of the box represents the median value. The upper whisker indicates the larger value between the largest data point and the 75th percentiles plus 1.5 $\times$ the interquartile range, and the lower whisker indicates the smaller value between the smallest data point and the 25th percentiles minus 1.5 $\times$ the interquartile range. Data points larger than the upper whisker or smaller than the lower whisker are identified as outliers. In the evaluation experiments in Fig. 2A and 3B, the statistics for each type of intracellular structure and SNR were calculated from 100 image pairs of 512 $\times$ 512 pixels, which were augmented following similar procedure to training phase from more than 20 different ROIs of 1024 $\times$ 1024 pixels.

Acknowledgements This work was financially supported by grants from Ministry of Science and Technology (2017YFA0505301), the National Natural Science Foundation of China (31827802, 31770930), the Chinese Academy of Sciences (ZDBS-LY-SM004).

Compliance with Ethical Standards

Conflict of interest Dong Li, Chang Qiao and Songyue Wang have a pending patent application on the presented framework.

Human and animal rights and informed consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Gustafsson MG, Shao L, Carlton PM, Wang CJ, Golubovskaya IN, Cande WZ, Agard DA, Sedat JW (2008) Three-dimensional

- resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophys J* 94: 4957–4970
- Guo Y, Li D, Zhang S, Yang Y, Liu JJ, Wang X, Liu C, Milkie DE, Moore RP, Tulu US, Kiehart DP, Hu J, Lippincott-Schwartz J, Betzig E, Li D (2018) Visualizing intracellular organelle and cytoskeletal interactions at nanoscale resolution on millisecond timescales. *Cell* 175: 1430–1442
- Hendrycks D, Gimpel K (2016) Gaussian error linear units (gelus). arXiv preprint arXiv: 160608415
- Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. Paper presented at: Proceedings of the IEEE conference on computer vision and pattern recognition
- Isensee F, Jaeger PF, Kohl SA, Petersen J, Maier-Hein KH (2021) nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods* 18: 203–211
- Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, Dong J, Prasadha MK, Pei J, Ting MYL, Zhu J, Li C, Hewett S, Dong J, Ziyar I, Shi A, Zhang R, Zheng L, Hou R, Shi W, Fu X, Duan Y, Huu VAN, Wen C, Zhang WD, Zhang CL, Li O, Wang X, Singer MA, Sun X, Xu J, Tafreshi A, Lewis MA, Xia H, Zhang K (2018) Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 172: 1122–1131
- Li D, Shao L, Chen BC, Zhang X, Zhang M, Moses B, Milkie DE, Beach JR, Hammer JA, 3rd Pasham M, Kirchhausen T, Baird MA, Davidson MW, Xu P, Betzig E (2015) Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics. *Science* 349: aab3500
- Lim B, Son S, Kim H, Nah S, Mu Lee K (2017) Enhanced deep residual networks for single image super-resolution. Paper presented at: Proceedings of the IEEE conference on computer vision and pattern recognition workshops
- Moen E, Bannon D, Kudo T, Graf W, Covert M, Van Valen D (2019) Deep learning for cellular image analysis. *Nat methods* 16: 1233–1246
- Pan C, Schoppe O, Parra-Damas A, Cai R, Todorov MI, Gondi G, von Neubeck B, Böğürçü-Seidel N, Seidel S, Sleiman K (2019) Deep learning reveals cancer metastasis and therapeutic antibody targeting in the entire body. *Cell* 179: 1661–1676
- Phillips MJ, Voeltz GK (2016) Structure and function of ER membrane contact sites with other organelles. *Nat Rev Mol Cell Biol* 17: 69–82
- Prinz WA (2014) Bridging the gap: membrane contact sites in signaling, metabolism, and organelle dynamics. *J Cell Biol* 205: 759–769
- Qiao C, Chen X, Zhang S, Li D, Guo Y, Dai Q, Li D (2021a) 3D Structured illumination microscopy via channel attention generative adversarial network. *IEEE J Selc Topics Quant Electr* 27: 1–11
- Qiao C, Li D, Guo Y, Liu C, Jiang T, Dai Q, Li D (2021b) Evaluation and development of deep neural networks for image super-resolution in optical microscopy. *Nat Methods* 18: 194–202
- Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. Paper presented at: International Conference on Medical image computing and computer-assisted intervention (Springer)
- Rust MJ, Bates M, Zhuang X (2006) Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods* 3: 793–795
- Shroff H, Galbraith CG, Galbraith JA, Betzig E (2008) Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. *Nat Methods* 5: 417–423
- Szegedy C, Ioffe S, Vanhoucke V, Alemi AA (2017) Inception-v4, inception-resnet and the impact of residual connections on learning. Paper presented at: Thirty-first AAAI conference on artificial intelligence
- Wang H, Rivenson Y, Jin Y, Wei Z, Gao R, Gunaydin H, Bentolila LA, Kural C, Ozcan A (2019) Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nat Methods* 16: 103–110
- Weigert M, Schmidt U, Boothe T, Muller A, Dibrov A, Jain A, Wilhelm B, Schmidt D, Broaddus C, Culley S, Rocha-Martins M, Segovia-Miranda F, Norden C, Henriques R, Zerial M, Solimena M, Rink J, Tomancak P, Royer L, Jug F, Myers EW (2018) Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat Methods* 15: 1090–1097
- Willig KI, Rizzoli SO, Westphal V, Jahn R, Hell SW (2006) STED microscopy reveals that synaptotagmin remains clustered after synaptic vesicle exocytosis. *Nature* 440: 935–939
- Zhang Y, Li K, Li K, Wang L, Zhong B, Fu Y (2018) Image super-resolution using very deep residual channel attention networks. Paper presented at: Proceedings of the European Conference on Computer Vision (ECCV)