Check for updates

small methods
www.small-methods.com

This Microtubule Does Not Exist: Super-Resolution Microscopy Image Generation by a Diffusion Model

*Alon Saguy, Tav Nahimov, Maia Lehrman, Estibaliz Gómez-de-Mariscal, Iván Hidalgo-Cenalmor, Onit Alalouf, Ashwin Balakrishnan, Mike Heilemann, Ricardo Henriques, and Yoav Shechtman**

Generative models, such as diffusion models, have made significant advancements in recent years, enabling the synthesis of high-quality realistic data across various domains. Here, the adaptation and training of a diffusion model on super-resolution microscopy images are explored. It is shown that the generated images resemble experimental images, and that the generation process does not exhibit a large degree of memorization from existing images in the training set. To demonstrate the usefulness of the generative model for data augmentation, the performance of a deep learning-based single-image super-resolution (SISR) method trained using generated high-resolution data is compared against training using experimental images alone, or images generated by mathematical modeling. Using a few experimental images, the reconstruction quality and the spatial resolution of the reconstructed images are improved, showcasing the potential of diffusion model image generation for overcoming the limitations accompanying the collection and annotation of microscopy images. Finally, the pipeline is made publicly available, runnable online, and user-friendly to enable researchers to generate their own synthetic microscopy data. This work demonstrates the potential contribution of generative diffusion models for microscopy tasks and paves the way for their future application in this field.

1. Introduction

Deep learning algorithms have been extensively used in the past decade to solve various microscopy challenges.[1–7] These algorithms outperform traditional computer vision methods in terms of reconstruction quality, analysis time, and more; however, deep learning solutions are hungry for data. Nowadays, to train a model, one should typically acquire and annotate thousands and sometimes even millions of images, a highly time and resource consuming process. An alternative approach is to produce synthetic data by developing mathematical models that approximate the structure of the biological specimen.[1,3,7–9] However, tuning the model parameters is a cumbersome and fundamentally imperfect process that leads to non-realistic features in the synthetic images due to parameter estimation errors and model inaccuracies.

A. Saguy, T. Nahimov, M. Lehrman, O. Alalouf, Y. Shechtman
Department of Biomedical Engineering
Technion – Israel Institute of Technology
Haifa 3200001, Israel
E-mail: yoavsh@bm.technion.ac.il

E. Gómez-de-Mariscal, I. Hidalgo-Cenalmor, R. Henriques
Optical cell biology group
Instituto Gulbenkian de Ciência
Oeiras 2780-156, Portugal

E. Gómez-de-Mariscal, I. Hidalgo-Cenalmor, R. Henriques
Optical cell biology group
Gulbenkian Institute of Molecular Medicine
Oeiras 2780-156, Portugal

A. Balakrishnan, M. Heilemann
Single Molecule Biophyiscs
Institute of Physical and Theoretical Chemistry
Goethe-University Frankfurt
60438 Frankfurt, Germany

R. Henriques
AI-driven Optical Biology
Instituto de Tecnologia Química e Biológica António Xavier
Universidade Nova de Lisboa
Oeiras 2780-157, Portugal

R. Henriques
UCL Laboratory for Molecular Cell Biology
University College London
London WC1E 6BT, UK

Y. Shechtman
Department of Mechanical Engineering
University of Texas at Austin
Austin, TX 78712-1591, USA

Y. Shechtman
Faculty of Electrical and Computer Engineering
Technion - Israel Institute of Technology
Haifa 3200001, Israel

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/smtd.202400672

© 2024 The Author(s). Small Methods published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

DOI: 10.1002/smtd.202400672

Recently, the field of generative models has seen a significant surge in terms of both development and application.[10–13] Generative models have moved far beyond their initial application in producing artificial images and are now being used to create synthetic datasets that can effectively mimic real-world data in diverse domains.[14] Two major contributors to this advancement have been Denoising Diffusion Probabilistic Models (DDPM)[10] and Denoising Diffusion Implicit Models (DDIM).[13] DDPM and DDIM offer a dynamic approach for the generation of synthetic data, relying on stochastic processes to create new images that still capture the inherent statistical behavior of the training dataset.

The capacity of diffusion models to accurately create realistic visual data is profoundly impacting many computer vision applications,[15] including microscopic imaging. Notably, acquiring high-quality large training datasets for microscopy is significantly harder than acquiring natural images, because of the complex experimental setups and the extensive sample preparations. Indeed, several studies already incorporate diffusion models to microscopy to reconstruct 3D biomolecule structures in Cryo-EM images,[16] predict 3D cellular structures from 2D images,[17] or design drug molecules,[18] among others.

Here, we propose the application of generative diffusion models in the field of super-resolution microscopy. First, we show the ability of diffusion models to generate realistic, high-quality, super-resolution microscopy images of microtubules and mitochondria. Then, we assess the capacity of the models to learn the intricate nature of the data domain by validating that the network rarely memorizes images from the training data. Next, we utilize the generated dataset to train a single-image super-resolution (SISR) deep learning model and show superior reconstruction quality compared to the same model trained on model-based simulated data or even on experimental data. The diffusion model approach proposed here is publicly available[19] on the ZeroCostDL4Mic platform,[20] enabling non-expert researchers to benefit from it.

2. Results

We base our work on a previously reported[21] diffusion model which we adapt to super-resolution microscopy. We trained two diffusion models on different biological samples, microtubules, and mitochondria. For the microtubule dataset, we used 7 images sourced from a publicly available database (ShareLoc.xyz),[22,23] where each image size is 1340 × 1340 *pixels*^{2} (corresponding to 53.6 × 53.6 μm^2). Since the data in ShareLoc is stored as a list of localizations per frame, we converted the localization lists of our training data to 2D localization histograms with bin size equal to the super-resolution pixel size (40 × 40 nm^2). Furthermore, we split each image into patches using sliding window of 256 × 256 pixels with 128 pixel overlap in each dimension, and transformed them using random horizontal flips and rotations of 90, 180, and 270 degrees to augment the training data. The augmentation step yielded a total of 2 000 training patches (250 unique patches + augmentations) for the microtubule training set.

The mitochondria dataset is composed of 10 stimulated emission depletion (STED) microscopy images. We split the mitochondria images of size 4096 × 4096 *pixels*^{2} (corresponding to 122.68 × 122.68 μm^2) into patches of 256 × 256 *pixels*^{2}. Since many patches in the mitochondria dataset do not contain any structure, we used manually generated masks to filter out empty patches from our database. Finally, we obtained 1150 training patches for the mitochondria data and augmented them to get ≈8000 training patches.

The images generated by our DDPM qualitatively resemble the training data used for the training, as can be clearly seen in the examples in **Figure 1**. For an additional quantitative similarity comparison see dataset similarity quantification section in the supporting information. To validate that our model does not memorize images, namely, copy existing images from the training set and generate them as network outputs, we calculated the maximal value of the normalized cross-correlation between every generated image (a total of 50 images) and all augmented patches used for training.

The cross-correlation calculation considered rotated, scaled, and translated versions of the training images, and same-size patches were used. The maximal value we got was 0.345 (0.485) for the synthetic microtubule (mitochondria) images. The mean value was 0.336 (0.211) for the microtubules (mitochondria) images. For benchmarking, we repeated this process with 50 experimental images from a different dataset, which yielded a mean value of 0.372 (0.186) and a max value of 0.483 (0.412) for the microtubule (mitochondria) data. The cross-correlation values are similar for experimental microtubule images from another dataset (imaged in similar conditions) and for the microtubule images that were generated by our diffusion model, showing the expected variability between different and independent datasets. In the case of mitochondria images, the cross-correlation values were slightly higher than those obtained when comparing with images from a different experimental technique, possibly implying a minor bias in the generated data toward the training samples.

To visualize the most similar images in order to rule-out memorization, we overlaid each generated image with the corresponding training image that obtained the highest cross-correlation score. **Figure 2** shows four such examples; the novelty of our generated data with respect to the training data is apparent.

Next, we tested the applicability of our generated data to improve the performance of a deep learning model for single-image super-resolution. Specifically, we used our generated data to train the Content-Aware Restoration (CARE)[1] model, aiming to transform a low-resolution image to a high-resolution image based on prior knowledge of image-statistics. Notably, obtaining single-image-based super-resolution algorithmically is yet an unsolved problem in microscopy, with results strongly dependent on the prior information provided, and is no match to experiment-based super-resolution microscopy methods (SMLM, STED, SIM, etc.).[24–27] Nevertheless, we use this task to demonstrate the potential of diffusion model-based data generation in virtual super-resolution microscopy imaging.

We trained CARE models for both biological samples using 1) randomly oriented sinusoidal synthetic microtubule filaments (see section 2.5.1 in the supplementary file of CARE manuscript); 2) images of microtubules or mitochondria generated by our diffusion model; 3) experimental images of microtubules or mitochondria used to train our diffusion model. During the training stage, we obtained the ground truth image in super-resolution either by simulation or from images reconstructed by one of the

 © 2024 The Author(s). Small Methods published by Wiley-VCH GmbH

23669608, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/smtd.202400672 by Cochrane Portugal, Wiley Online Library on [03/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

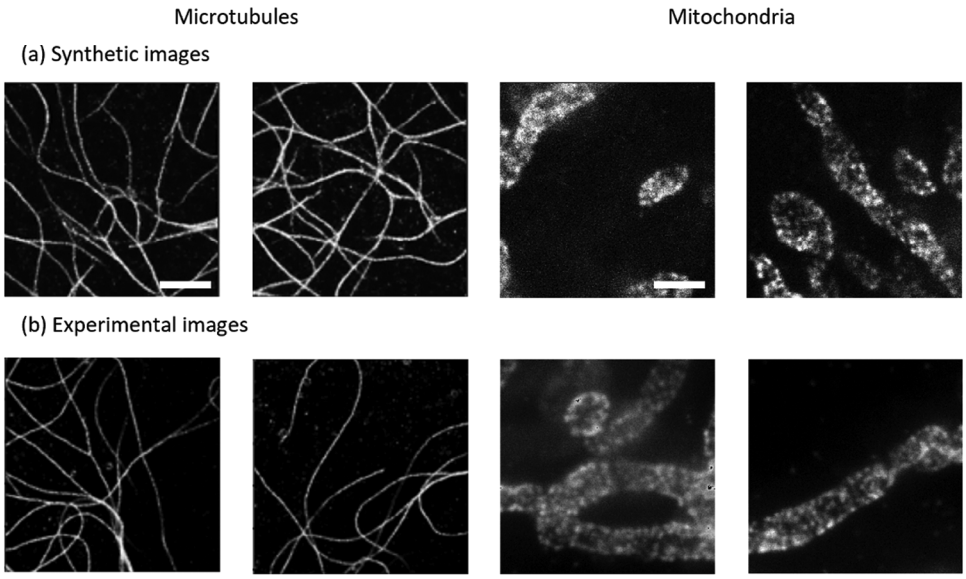

Figure 1. Qualitative comparison of experimental microscopy data versus data generated using our generative diffusion model. a) Example synthetic images of microtubules (alpha-tubulin – Alexa647) and mitochondria (TOM 22 – Alexa647) generated by our diffusion model. b) Example experimental super-resolution images, used as training data. Scale bars = 2.5 µm.

existing super-resolution methods (either SMLM or STED); next, we obtained low-resolution images by forward passing the high-resolution images through a model of our optical system (see methods section for more details). Ultimately, we used these low-resolution – high-resolution image pairs to train CARE. We tested CARE on 10 microtubule low-resolution – high-resolution image pairs of size 1024 × 1024 $pixels^2$, that were not used for training.

Visually, the CARE model trained on microtubules generated by the diffusion model yielded a better reconstruction in comparison the CARE model trained on microtubules generated by a mathematical model (**Figure 3**). To quantify the improvement, we have analyzed the spatial resolution we obtained in both reconstructions using the Fourier Ring Correlation (FRC) plug-in for ImageJ.[28] In brief, FRC is a similarity measure that seeks the maximal spatial frequency in which the reconstructed image and the ground truth image are correlated up to a predefined threshold. The similarity is quantified by the normalized cross-correlation between the Fourier transforms of both images inside a torus with increasing radius. A high cross-correlation value within the torus indicates high similarity between the images, in the corresponding spatial frequency band.

The mean spatial resolution of the reconstructed images, as quantified by FRC using a 1/7 threshold[28] when training on microtubule images generated by our diffusion model, was 115 nm with a standard deviation of 16 nm, while the mean spatial resolution obtained when training on synthetic microtubules generated via a mathematical model was 140 nm with a standard deviation of 21 nm.

Finally, we report the mean peak signal-to-noise ratio (PSNR), normalized root mean squared error (NRMSE), and multi-scale structural similarity index measure (MS-SSIM)[29] similarity metrics between 30 reconstructed images and the corresponding ground truth images (see **Table 1**). The CARE model trained on diffusion model outperformed the CARE model trained on mathematical model based synthetic data on all three quantitative measures. These results demonstrate the advantage of

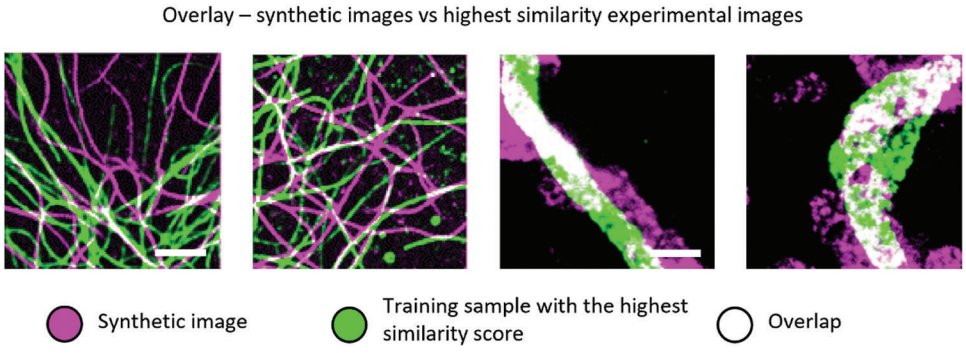

Figure 2. Overlay between each reconstructed image and the training image with highest resemblance (maximal cross-correlation score). Purple marks generated data, green marks the closest training sample, and white marks overlap between the two images. Our diffusion models do not exhibit a large degree of memorization of structures from the training images. Scale bars = 2.5 µm.

 © 2024 The Author(s). Small Methods published by Wiley-VCH GmbH

23669608, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/smtd.202400672 by Cochrane Portugal, Wiley Online Library on [03/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

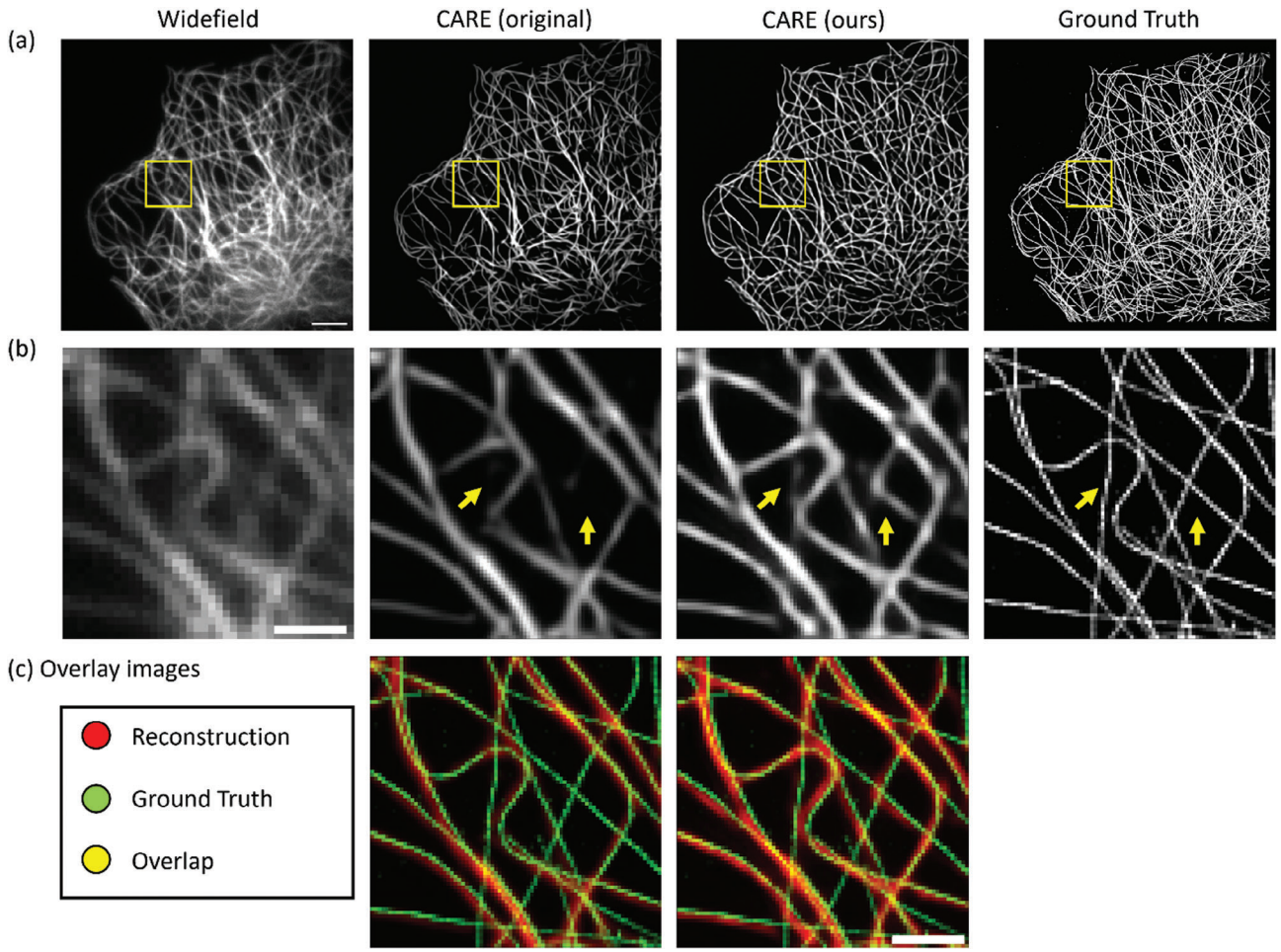

Figure 3. Performance of CARE trained on synthetic microtubule images generated by a mathematical model versus training on microtubules generated by our diffusion model. a) Left to right: widefield image, CARE reconstruction when trained on mathematical simulations, CARE reconstruction when trained on our synthetic data, and ground truth. Scale bar = 5 µm. b) Regions of interest (marked by yellow squares in (a)), yellow arrows mark areas in which CARE trained on our data outperformed the previous method. c) Left: overlay between CARE trained on mathematical simulations (red) and the ground truth (green). Right: overlay between CARE trained on our diffusion model-based synthetic data (red) and the ground truth (green). Scale bar = 1 µm.

using synthetic simulated data generated by our diffusion model in comparison to the mathematical model of microtubules.

Notably, microtubule images can be simulated with relatively high fidelity by a variety of well-established mathematical models.[30] However, for an arbitrary type of biological specimen, it is not easy to obtain a simple mathematical model describing its morphology. Therefore, the most remarkable feature of diffusion model-based data generation is the ability to generate synthetic data from non-mathematically- describable biological specimens.

We demonstrate this ability by training our diffusion model on STED images of mitochondria (**Figure 4**). Unlike for microtubules, there is no available mathematical model to generate mitochondria images. Therefore, we compare CARE's reconstruction versus a model trained on the same experimental data used for training the diffusion model. Moreover, we also explored other scenarios where we combined the generated images with the experimental data at hand, or generated a much larger number of unique samples using the diffusion model than the number of experimental samples.

We compared the performance on a test set of 30 mitochondria images of size 1024 × 1024 *pixels*^{2} using four CARE models: 1) CARE trained on 1150 experimental patches (before augmentation); 2) CARE trained on 1150 generated training patches (before augmentation); 3) CARE trained on both the experimental patches and the generated patches; 4) CARE trained on 6000 generated patches (before augmentation). The testing set is

Table 1. Quantitative comparison of the reconstructions obtained by CARE models trained on different microtubule datasets. We report the mean and standard deviation of PSNR, NRMSE, and MS-SSIM over the microtubule test set composed of $n = 10$ images.

CARE training data	Mean PSNR	Mean NRMSE	Mean MS-SSIM
Mathematical model	15.859 ± 0.2	1.839 ± 0.18	$0.9985 \pm 16 \cdot 10^{-5}$
Diffusion model	18.734 ± 1.41	1.3352 ± 0.24	$0.9992 \pm 29 \cdot 10^{-5}$

 © 2024 The Author(s). Small Methods published by Wiley-VCH GmbH

23669608, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/smtd.202400672 by Cochrane Portugal, Wiley Online Library on [03/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

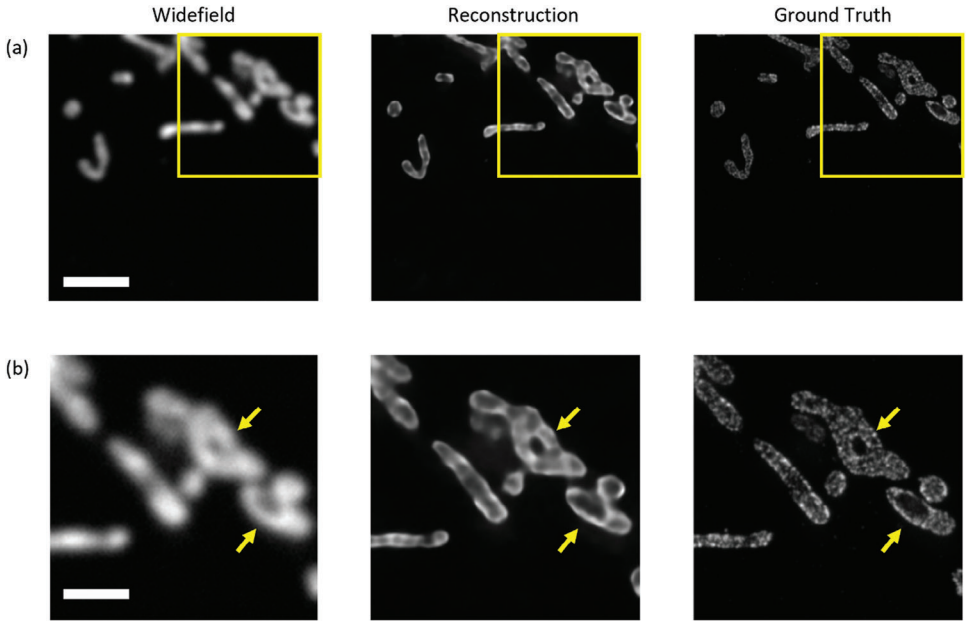

Figure 4. Performance of CARE trained on mitochondria generated by our diffusion model. a) Left to right: widefield image, CARE reconstruction when trained on our synthetic data, and ground truth. Scale bar = 8 µm. b) Region of interest; yellow arrow marks a subtle feature not visible in widefield imaging, which is made visible in our reconstruction. Scale bar = 2 µm.

composed of 10 STED split to images of size 1024 × 1024 $pixels^2$, corresponding to 30.67 × 30.67 μm^2.

The quantitative comparison, shown in **Figure 5** and **Table 2**, is based on PSNR, NRMSE, MS-SSIM, and FRC evaluation metrics. According to the results, when training a model using a small number of samples, the quantitative results are similar, for example mean PSNR of 27.26 for the small synthetic dataset compared to 27.11 for the experimental dataset. However, diffusion models could serve as additional data augmentation for the experimental data, namely, one could generate as many new synthetic images as one desires. Indeed, we demonstrate an improvement in all evaluation metrics when using a much larger number (6000 before augmentation) of synthetic images for the training. Lastly, we also tested a combination of the small synthetic dataset and the experimental dataset.

3. Discussion

In this work, we demonstrate the potential of diffusion models to generate large super-resolution microscopy datasets by relying on a relatively small number of super-resolution images. Given only 7 (10) microtubule (mitochondria) images we manage to generate realistic images that look different from the original training data. Importantly, existing work in this field[31,32]

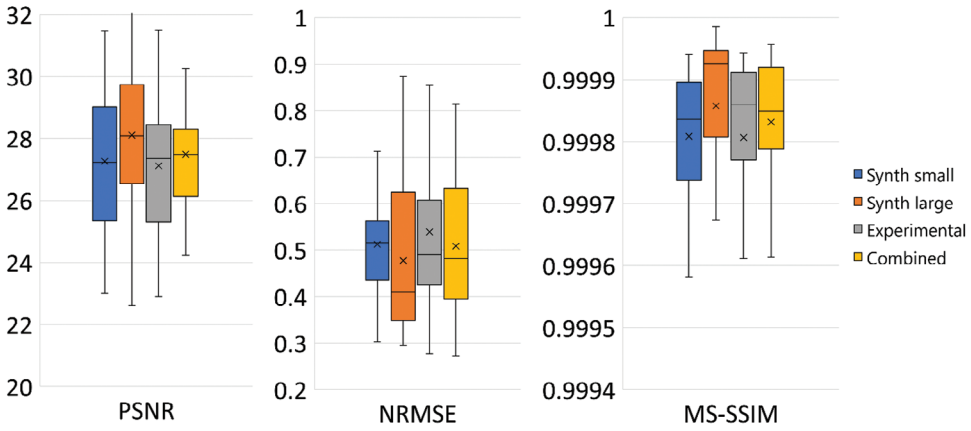

Figure 5. Quantitative comparison of CARE models for mitochondria over $n = 30$ samples. Boxes include data points inside the [25th,75th] percentile range, horizontal line marks the median and x marks the mean value. Synth small (blue) marks CARE trained on 1150 synthetic patches (before augmentation) generated by our diffusion model. Synth large (orange) marks CARE trained on 6 000 synthetic patches (before augmentation). Experimental (gray) marks CARE trained on 1150 experimental patches (before augmentation). Combined (yellow) marks CARE trained on the combination of 1150 synthetic patches and 1150 experimental patches (before augmentation).

 © 2024 The Author(s). Small Methods published by Wiley-VCH GmbH

23669608, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/smtd.202400672 by Cochrane Portugal, Wiley Online Library on [03/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2. Quantitative comparison of the reconstructions obtained by CARE models trained on different mitochondria datasets. We report the mean PSNR, NRMSE, MS-SSIM, and FRC over the mitochondria test set composed of 10 images.

CARE training data	Mean PSNR	Mean NRMSE	Mean MS-SSIM	FRC [nm]
Small synthetic dataset	27.26 ± 2.19	0.51 ± 0.11	$0.9981 \pm 1.2 \cdot 10^{-3}$	117 ± 10
Experimental dataset	27.11 ± 2.15	0.53 ± 0.19	$0.9981 \pm 1.5 \cdot 10^{-3}$	114 ± 15
Large synthetic dataset	28.10 ± 2.67	0.47 ± 0.15	$0.9998 \pm 1.5 \cdot 10^{-3}$	100 ± 25
Combined dataset	27.48 ± 2.00	0.50 ± 0.14	$0.9998 \pm 1.2 \cdot 10^{-3}$	109 ± 14

has shown that training diffusion models on smaller dataset increase the possibility the model will memorize the training set. Here, we report training on as little as 7 images (split into 250 patches before augmentations) without memorization of large structural features from the training data. Two possible explanations are the fact that we could generate thousands of patches for training out of those images, or the relatively high redundancy of information in microscopy images compared to natural images.

Next, we trained a single-image super-resolution deep learning model, namely CARE, to convert low resolution microscopy images into high-resolution images. Our results show that combining synthetic and experimental images in the model training improves model performance. Additionally, when we trained CARE model on more synthetic images than the number of experimentally acquired samples, we still managed to improve the performance, even beyond the CARE model trained on the combined dataset. Nevertheless, recent work[33,34] states that deep-learning models that are trained on purely synthetic data, created by generative AI models, might collapse to a relatively narrow distribution of observations due to over-representation of certain structures in the generative model training data. An interesting open question for future investigation is – given that the diffusion model can generate an arbitrarily large number of different images, at which point does adding new generated images not contribute anymore to performance? The answer will likely be sample and task dependent.

Creating synthetic images of biological data that are highly realistic and representative of the original data has important implications, especially for downstream tasks that do not require complicated annotation, or any annotation at all. For example, diffusion models enable the generation of super-resolution datasets that could be transformed to low-resolution observations by forward passing through an optical model of the imaging system; then, one may perform supervised model training without the need for extensive experimental data acquisition, which is often a limiting factor. The contribution of our method is particularly relevant for the general case where no simple mathematical model is available for synthetic image generation.

Although this work demonstrates the potential of a generative diffusion model in the task of single-image super-resolution, the applicability of such a technique for microscopy is naturally much broader. Numerous potential applications exist, including denoising, multi-image super-resolution, cross-modality imaging, live-cell dynamic imaging, and more. On the other side, quantitative evaluation of biological image data generation in the lack of annotated images is still an open question in the field that requires further work and consensus.

Finally, we share an easy-to-use notebook via the ZeroCostDL4Mic[20] platform to enable researchers to replicate our pipeline and harness diffusion model capabilities. We also distribute the pretrained models that allow the generation of data similar to the data presented in this work. Of note, training diffusion models is time consuming due to the large number of stochastic operations involved in the learning process.

In light of the encouraging results obtained from this study, future research should continue to focus on further optimizing and evaluating diffusion models for generating more types of synthetic microscopy data and on finding the applications where these capabilities are most impactful. Furthermore, due to the capacity of diffusion models to create virtual representations of nanoscale cellular structure, they can potentially predict prospective multi-structural spatial relationships that will guide observations and discovery in the field of microscopy. The emergence of generative models for microscopy represents an exciting phase for bio-medical research and holds promising potential for advancements in the near future.

4. Experimental Section

Optical Model for Low-Resolution Image Generation: To train CARE on low-resolution – high-resolution image pairs, high-resolution data were used and passed it through a model of the optical system to obtain low-resolution images. In this work, a simple model was used to simulate a 2D low-resolution image based on a 2D high-resolution image, described below. Let the imaged structure be depicted by $S(x, y)$ and let $H(x, y)$, the point spread function (PSF) of the optical system, be modeled as a 2D Gaussian:

$$H(x, y) = A \cdot e^{-\frac{(x-x_0)^2}{2\sigma_x^2} - \frac{(y-y_0)^2}{2\sigma_y^2}} \tag{1}$$

where A is the amplitude of the PSF, x_0, y_0 are the position of the emitter, and $\sigma_x = \sigma_y = \sigma$ represents the PSF width.

The low-resolution image formed at the camera plane is described by the convolution of the imaged structure with the system's PSF equation:

$$I(x, y) = P(S(x, y) * H(x, y)) + G(x, y) \tag{2}$$

where * indicates a convolution operator, $P(x, y)$ indicates a Poisson distribution of the emitted number of photons, and $G(x, y)$ indicates a Gaussian noise simulating the camera read noise.

Ultimately, to obtain a low-resolution image $\tilde{I}(x, y)$, the high-resolution image $I(x, y)$ is down sampled using 2D average pooling layer with window size = 4:

$$\tilde{I}(x, y) = AvgPooling2D\{I(x, y)\} \tag{3}$$

 © 2024 The Author(s). Small Methods published by Wiley-VCH GmbH

23669608, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/smtd.202400672 by Cochrane Portugal, Wiley Online Library on [03/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Diffusion Model Architecture and Training Details: The network architecture presented by Nichol, et al was adopted.[21] A single residual network (ResNet) block was used and the input and output layers of the model were changed to fit monochromatic data. To decrease the network size, the channel multiplication between different layers of the ResNet, namely was also changed, instead of (1, 1, 2, 2, 4, 4) multiplication (1, 1, 2, 2, 2, 2) multiplication was used, where the initial channel number is 64. Additionally, the number of diffusion steps was changed to 2000, set the batch size to 10, the learning rate to $1e^{-5}$, and employed a cosine noise schedule. To train the network, 7 publicly available (ShareLoc)[22] super-resolution localization lists of microtubule experiments, and 10 of mitochondria were used STED images; then, a super-resolved image was generated from each localization list scaled by a factor of 4 in comparison to the diffraction limited data, yielding pixel sizes of 40 and 30 nm for the microtubule and mitochondria images, respectively.

Finally, the generative diffusion model was trained over 80000/20000 steps for 8 (2) h for the microtubule (mitochondria) datasets on a single NVIDIA 32GB Titan RTX GPU. Ultimately, generation of a single super-resolution image depends on the image size, e.g. 15 s for images of size 256×256 pixels^2.

CARE Training Details: Super-resolution training data was obtained based on 1) the mathematically simulated data presented in the CARE paper; 2) the data generated by the trained diffusion model; or 3) the experimental data from ShareLoc (for microtubules) or captured with STED microscopy (for mitochondria). To generate the low-resolution data needed for training CARE network, a similar scheme as described in the CARE paper was followed by convolving the super-resolution data with a Gaussian microscope PSF model and adding Perlin noise, shot noise and Gaussian noise. Importantly, It was sured that image generated by the two methods described above shared properties such as signal-to-noise ratio, sample size, etc. Finally, the CARE network was trained on 2 000 synthetic low-resolution-high-resolution image pairs for the microtubule reconstruction and 8 000 for the mitochondria reconstruction. To maintain a fair comparison between CARE trained on the data versus CARE trained on the mathematically generated microtubules, the same training set size was used in both cases.

Cell Culture: U2-OS cells for mitochondrial immunostaining were cultured in DMEM/F12 media supplemented with 10% fetal bovine serum (FBS) (Corning, USA), 1% (v/v) penicillin-streptomycin (Thermo Fisher Scientific, Germany) and 5% (v/v) Glutamax (Thermo Fisher Scientific, Germany). The cells were incubated at 37°*C* and 5% CO_2 and were passaged every 2–3 days or when they were 80% confluent.

Immunostaining of Mitochondria: U2-OS cells were seeded onto fibronectin (Sigma-Aldrich, Germany) coated 8-well chambered coverglass (Sarstedt, Germany) at an amount of $1.5 \cdot 10^4$ cells per well and were incubated overnight at 37°*C* and 5% CO_2. The cells were fixed with 4% (v/v) formaldehyde (FA) (Sigma–Aldrich, Germany) and 0.1% (v/v) glutaraldehyde (Sigma–Aldrich, Germany) in pre-warmed 1x phosphate buffer saline (PBS) at 37°*C* for 20 min. The cells were washed with 1x PBS once and treated with sodium borohydride (a pinch) dissolved in 1 mL 1xPBS for 7 min (per coverglass). The sample was then washed thrice with 1x PBS. The sample was incubated for 10 min in immunofluorescence (IF) buffer (10% (v/v) FBS (Corning, USA) and 0.1% (w/v) Saponin (Sigma–Aldrich, Germany)). After this, the sample was incubated with primary antibody (TOM20, 1:500 (Rabbit TOM20 (sc-11415, Santa Cruz, USA))) dissolved in IF buffer for 1.5–2 h at room temperature (RT) with shaking. This was followed by washing thrice with 1x PBS and then the secondary antibody (Goat anti-Rabbit Abberior Star 635P at 1:1000 dilution (2.0012–007-2, Abberior Instruments)) dissolved in IF buffer and incubated for 1.5–2 h at RT with shaking. The sample was then washed thrice with 1x PBS and post fixed with 4%(v/v) FA (Sigma–Aldrich, Germany) for 10 min at RT followed by washing thrice with 1x PBS. The sample was then stored at 4°*C* in PBS for long-term and equilibrated to RT for an hour before starting measurements.

STED Imaging: STED imaging was performed using an Abberior expert line microscope (Abberior Instruments, Germany) equipped with a UPLXAPO 60x NA 1.42 oil immersion objective (Olympus Deutschland GmbH, Germany). An excitation laser of 640 nm (7.7 µ*W* at the back focal plane) and a depletion laser of 775 nm (136.5 *mW* at the back focal plane) were used for STED imaging, both pulsed at 40 MHz. Fluorescence was collected in the spectral range of 650 –760 nm using an avalanche photo diode (APD) with time gating enabled (0.75–8.75 ns). Pixel size was set to 15 nm with a line integration of 5, pixel dwell time of 5 µs and a pinhole of 0.81 airy units.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This research was supported in part by funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 802567-ERC-Five-Dimensional Localization Microscopy for Sub-Cellular Dynamics. Y.S. is supported by the Zuckerman Foundation and by the Donald D. Harrington Fellowship. E.G.M., I.H.C., and R.H. are supported by the Gulbenkian Foundation (Fundação Calouste Gulbenkian), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 101001332 to R.H.), the European Union through the Horizon Europe program (AI4LIFE project with grant agreement 101057970-AI4LIFE, and RT-SuperES project with grant agreement 101099654-RT-SuperES to R.H.) the European Molecular Biology Organization (EMBO) Installation Grant (EMBO-2020-IG-4734 to R.H.) and Postdoctoral Fellowship (EMBO ALTF 174–2022 to, E.G.,M.) and the Chan Zuckerberg Initiative via a Visual Proteomics Grant (vpi-0000000044 to R.H.) and an Essential Open Source Software for Science (EOSS6-0000000260 to R.H.). R.H. also acknowledges the support of LS4FUTURE Associated Laboratory (LA/P/0087/2020). M.H. and A.B. acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG; grants SFB 1177, INST 161/1020-1). Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are openly available in Microtubule stained in alexa 647 treated with 1 um of Nocodazole at https://doi.org/10.5281/zenodo.6948554, reference number 6948554.

Keywords

deep learning, generative AI, single molecule localization microscopy, super-resolution microscopy

Received: May 8, 2024
Revised: September 7, 2024
Published online:

[1] M. Weigert, U. Schmidt, T. Boothe, A. Müller, A. Dibrov, A. Jain, B. Wilhelm, D. Schmidt, C. Broaddus, S. Culley, M. Rocha-Martins, F. Segovia-Miranda, C. Norden, R. Henriques, M. Zerial, M. Solimena, J. Rink, P. Tomancak, L. Royer, F. Jug, E. W. Myers, *Nat. Methods.* **2018**, *15*, 1090.

 © 2024 The Author(s). Small Methods published by Wiley-VCH GmbH

23669608, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/smtd.202400672 by Cochrane Portugal, Wiley Online Library on [03/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[2] E. Nehme, L. E. Weiss, T. Michaeli, Y. Shechtman, *Optica* **2018**, *5*, 458.
[3] A. Saguy, O. Alalouf, N. Opatovski, S. Jang, M. Heilemann, Y. Shechtman, *Nat. Methods.* **2023**, *20*, 1939.
[4] Y. Nogin, T. Detinis Zur, S. Margalit, I. Barzilai, O. Alalouf, Y. Ebenstein, Y. Shechtman, *Bioinformatics* **2023**, *39*, btad137.
[5] E. Nehme, D. Freedman, R. Gordon, B. Ferdman, L. E. Weiss, O. Alalouf, T. Naor, R. Orange, T. Michaeli, Y. Shechtman, *Nat. Methods.* **2020**, *17*, 734.
[6] W. Ouyang, A. Aristov, M. Lelek, X. Hao, C. Zimmer, *Nat. Biotechnol.* **2018**, *36*, 460.
[7] A. Saguy, F. Jünger, A. Peleg, B. Ferdman, E. Nehme, A. Rohrbach, Y. Shechtman, *Opt. Express.* **2021**, *29*, 23877.
[8] K. W. Dunn, C. Fu, D. J. Ho, S. Lee, S. Han, P. Salama, E. J. Delp, *Sci. Rep.* **2019**, *9*, 18295.
[9] W. Xie, J. A. Noble, A. Zisserman, *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, Taylor and Francis, Milton Park, Oxfordshire UK **2016**, 283.
[10] J. Ho, A. Jain, P. Abbeel, *Adv. Neural. Inform. Proc. Sys.* **2020**, *33*.
[11] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu, M. Chen, arXiv preprint, arXiv:2204.06125, **2022**.
[12] S. Gu, D. Chen, J. Bao, F. Wen, B. Zhang, D. Chen, L. Yuan, B. Guo, In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, Piscataway, NJ **2022**, 10696.
[13] J. Song, C. Meng, S. Ermon, arXiv preprint, arXiv:2010.02502, **2020**.
[14] Z. Guo, J. Liu, Y. Wang, M. Chen, D. Wang, D. Xu, J. Cheng, *Nat. Rev. Bioeng.* **2024**, *2*, 136.
[15] T. Yang, Y. Ying, *ACM Computing. Surveys.* **2023**, *56*, 105.
[16] K. Kreis, T. Dockhorn, Z. Li, E. Zhong, arXiv preprint, arXiv:2211.14169, **2022**.
[17] D. J. E. Waibel, E. Röell, B. Rieck, R. Giryes, C. Marr, In 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), IEEE, Piscataway, NJ **2023**, 1.
[18] I. Igashov, H. Stärk, C. Vignac, A. Schneuing, V. G. Satorras, P. Frossard, M. Welling, M. Bronstein, B. Correia, *Nat. Machine. Intel.* **2024**, *6*, 417.
[19] Diffusion model for SMLM: https://colab.research.google.com/github/HenriquesLab/ZeroCostDL4Mic/blob/master/Colab_notebooks/Diffusion_Model_SMLM_ZeroCostDL4Mic.ipynb (accessed: February 2024).
[20] L. von Chamier, R. F. Laine, J. Jukkala, C. Spahn, D. Krentzel, E. Nehme, M. Lerche, S. Hernández-Pérez, P. K. Mattila, E. Karinou, S. Holden, A. C. Solak, A. Krull, T.-O. Buchholz, M. L. Jones, L. A. Royer, C. Leterrier, Y. Shechtman, F. Jug, M. Heilemann, G. Jacquemet, R. Henriques, *Nat. Commun.* **2021**, *12*, 2276.
[21] A. Q. Nichol, P. Dhariwal, in *Proceedings of Machine Learning Research (PMLR)*, **2021**, *139*, pp. 8162–8171.
[22] W. Ouyang, J. Bai, M. K. Singh, C. Leterrier, P. Barthelemy, S. F. H. Barnett, T. Klein, M. Sauer, P. Kanchanawong, N. Bourg, M. M. Cohen, B. Lelandais, C. Zimmer, *Nat. Methods.* **2022**, *19*, 1331.
[23] Manish S. I. N. G. H., *zenodo* **2022**.
[24] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, H. F. Hess, *Science* **2006**, *313*, 1642.
[25] M. J. Rust, M. Bates, X. Zhuang, *Nat. Methods.* **2006**, *3*, 793.
[26] S. W. Hell, J. Wichmann, *Opt. Lett.* **1994**, *19*, 780.
[27] M. G. L. Gustafsson, *J. Microsc.* **2000**, *198*, 82.
[28] R. P. J. Nieuwenhuizen, K. A. Lidke, M. Bates, D. L. Puig, D. Grünwald, S. Stallinga, B. Rieger, *Nat. Methods.* **2013**, *10*, 557.
[29] Z. Wang, E. P. Simoncelli, A. C. Bovik, In The Thrity-Seventh Asilomar Conference on Signals Systems & Computers, IEEE, Piscataway, NJ **2003**, *2*, 1398.
[30] A. Shariff, R. F. Murphy, G. K. Rohde, *Cytometry, Part A* **2010**, *77*, 457.
[31] G. Somepalli, V. Singla, M. Goldblum, J. Geiping, T. Goldstein, *Adv. Neural. Inform. Proc. Sys.* **2023**, *36*, 47783.
[32] G. Somepalli, V. Singla, M. Goldblum, J. Geiping, T. Goldstein, In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, Piscataway, NJ **2023**, 6048.
[33] M. Gerstgrasser, R. Schaeffer, A. Dey, R. Rafailov, H. Sleight, J. Hughes, T. Korbak, R. Agrawal, D. Pai, A. Gromov, D. A. Roberts, D. Yang, D. L. Donoho, S. Koyejo, arXiv preprint, arXiv:2404.01413, **2024**.
[34] I. Shumailov, Z. Shumaylov, Y. Zhao, Y. Gal, N. Papernot, R. Anderson, arXiv preprint, arXiv:2305.17493, **2023**.

 © 2024 The Author(s). Small Methods published by Wiley-VCH GmbH

23669608, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/smtd.202400672 by Cochrane Portugal, Wiley Online Library on [03/02/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License