



OPEN

## Deep learning for blind structured illumination microscopy

Emmanouil Xypakis<sup>1,2</sup>✉, Giorgio Gosti<sup>1,6</sup>, Taira Giordani<sup>1,3</sup>, Raffaele Santagati<sup>4,5</sup>, Giancarlo Ruocco<sup>1</sup> & Marco Leonetti<sup>1,2,6</sup>

Blind-structured illumination microscopy (blind-SIM) enhances the optical resolution without the requirement of nonlinear effects or pre-defined illumination patterns. It is thus advantageous in experimental conditions where toxicity or biological fluctuations are an issue. In this work, we introduce a custom convolutional neural network architecture for blind-SIM: BS-CNN. We show that BS-CNN outperforms other blind-SIM deconvolution algorithms providing a resolution improvement of 2.17 together with a very high *Fidelity* (artifacts reduction). Furthermore, BS-CNN proves to be robust in cross-database variability: it is trained on synthetically augmented open-source data and evaluated on experiments. This approach paves the way to the employment of CNN-based deconvolution in all scenarios in which a statistical model for the illumination is available while the specific realizations are unknown or noisy.

In optics, the diffraction limit is the minimum distance at which two objects can be identified and it depends on wavelength and numerical aperture<sup>1</sup>. It is possible to surpass the diffraction limit and improve the image contrast by removing or reversing the effects of diffraction and other sources of noise<sup>2</sup> with deconvolution methods. The performance of these methods depends on both the knowledge of the experimental conditions and the imaged sample properties. For example, sparsity of molecular emitters induced by photoactivation or photobleaching enables the observer to achieve a more-than-ten times resolution enhancement<sup>3–5</sup> thanks to single molecule localization techniques. Another popular technique, Structured Illumination Microscopy (SIM)<sup>6</sup>, improves the resolution up to a factor 2 by exploiting multiple spatially periodic illumination patterns. However, this resolution enhancement rapidly degrades if patterns are instable or if knowledge of the spatial illumination profile is imprecise. Indeed, image reconstruction and deconvolution may be performed also with unknown illumination patterns<sup>7–11</sup>. These blind-SIM techniques are not subject to pattern evaluation errors, such as misalignment and repeatability issues, and are immune to aberrations on the illumination path<sup>12</sup>.

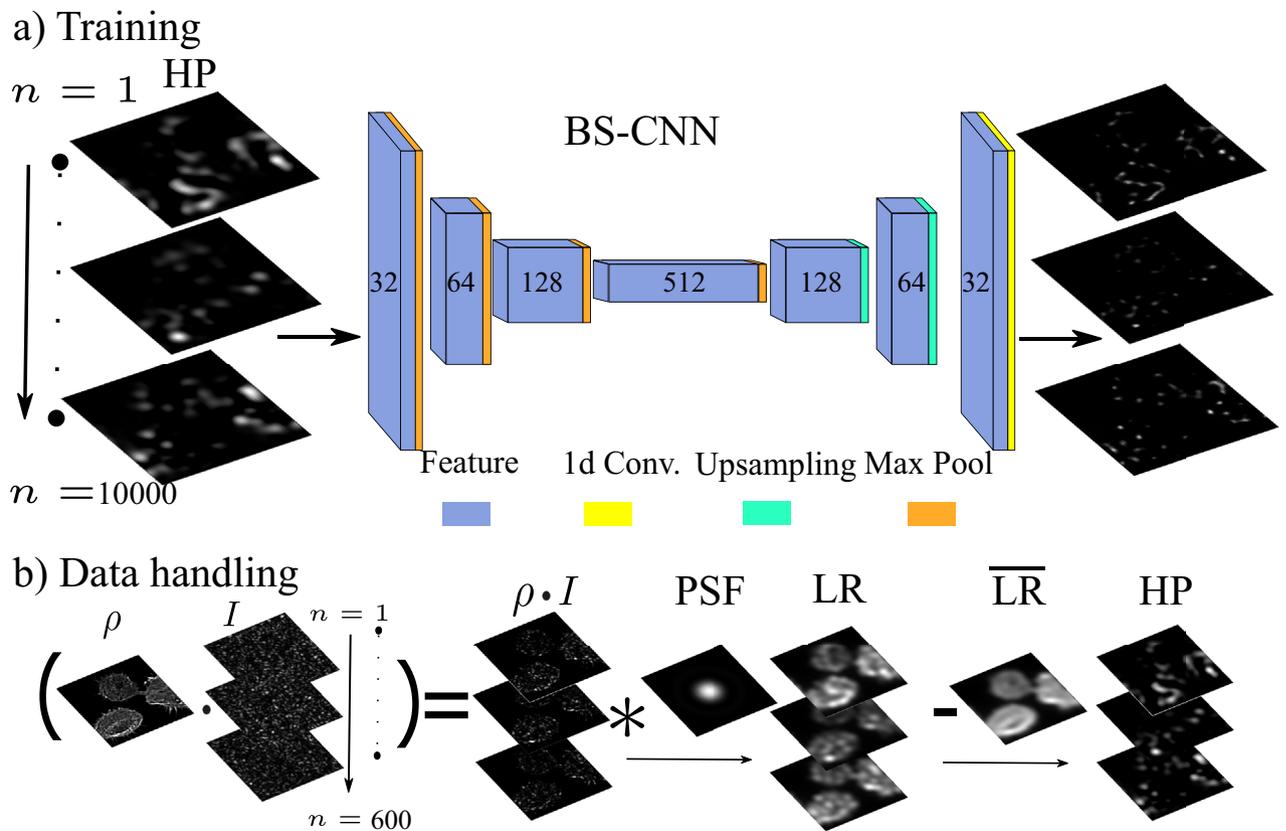
Blind-SIM deconvolution methods, such as Scattering Assisted Imaging (SAI)<sup>11</sup>, often requires complex and computationally expensive deconvolution algorithms. Luckily, the advent of Neural Networks (NN) provided new solutions for image analysis. NN are capable of reducing the computational time for deconvolution and improve image resolution<sup>13</sup>. Typically they exploit dictionary learning and the upsampling of sparse representations<sup>14,15</sup> that efficiently remove complex blurring<sup>16,17</sup>. In Microscopy, Deep Neural Networks can tackle the inverse imaging problem<sup>18</sup>, deconvolve images from different fluorescence microscopy modalities such as wide-field fluorescence or confocal microscopy<sup>19</sup>, and efficiently localize single molecules<sup>20,21</sup> or dimers<sup>22</sup>. In SIM, a NN enables a substantial reduction of the required frames number by using a multi-channel architecture<sup>23</sup>, and can provide an improved noise resilience<sup>24</sup>. To the best of our knowledge neural networks have never been applied to blind-SIM.

In this work, we design, test, and validate a Convolutional Neural Network for the reconstruction of blind SIM deconvolutions: the BS-CNN. This approach enables to deconvolve images in a few seconds and to surpass other state of the art blind-SIM deconvolution algorithms in performance such as<sup>7,11</sup>.

### Results

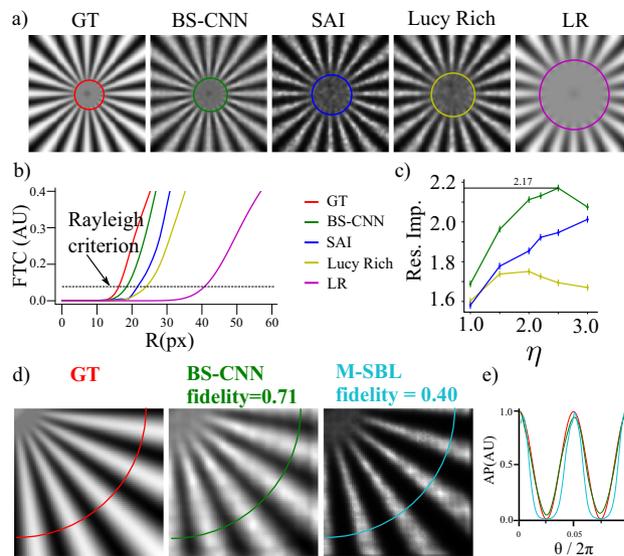
**CNN training.** BS-CNN deconvolves images by employing an encoder-decoder architecture that has been adapted and modified for blind SIM. Encoder-decoder network architectures are widely used in biology applications such as image deconvolution<sup>21</sup> or cell counting<sup>26</sup>. They are very easy to adapt and train if compared with the more sophisticated and potentially more powerful UNET architecture<sup>27,28</sup>. They are based on an information

<sup>1</sup>Center for Life Nano- and Neuro-Science, Istituto Italiano di Tecnologia, Viale Regina Elena 291, 00161 Rome, Italy. <sup>2</sup>D-TAILS srl, 00161 Rome, Italy. <sup>3</sup>Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185 Rome, Italy. <sup>4</sup>Quantum Engineering Technology Labs, University of Bristol, Bristol BS8 1FD, UK. <sup>5</sup>Quantum Lab, Boehringer-Ingelheim, Doktor-Boehringer-Gasse 5-11, 1120 Wien, Austria. <sup>6</sup>Soft and Living Matter Laboratory, Institute of Nanotechnology, Consiglio Nazionale delle Ricerche, 00185 Rome, Italy. ✉email: Emmanouil.Xypakis@iit.it



**Figure 1.** BS-CNN scheme: (a) We train the CNN with HP frames so that each illumination and each fluorescent density appears only once and minimizes the loss by comparing with the corresponding  $\rho \cdot I$  image. The BS-CNN architecture: for the encoder-decoder architecture we use two-dimensional convolutional layers of size  $3 \times 3$  followed by an element-wise ReLU non-linearity. The feature number is first increased (encoder) by a factor of two from 32 to 512 while the image size is decreased by a  $4 \times 4$  max pool layer. Then the opposite procedure is applied (decoder) up to the original image size. Finally, a one-dimensional convolutional layer is applied to produce a single-channel image. (b) Data handling: a ground truth image is selected from an open microscopy repository<sup>25</sup>; the ground truth image is illuminated 600 times by different speckle realizations  $I$ , producing 600 high-frequency images; we convolve with an Airy disk Point spread function (PSF) of resolution  $\eta$  times bigger than the speckle correlation size and obtain 600 low-resolution frames corresponding to the same GT. By subtracting the low-resolution mean  $\overline{\text{LR}}$  from each low-resolution frame LR and keeping only the high positive part we obtain the HP images.

theory framework, in which a first encoding step generates a high-dimensional dense feature representation of the input, and a second decoding step upsamples the information in the desired representation<sup>29</sup>. In the encoding stage (see Fig.1 panel a), BS-CNN encodes the input in a dense feature representation by applying two-dimensional  $3 \times 3$  convolutional layers of feature size 32, 64, 128 and 512 respectively, then followed by an element wise ReLU layer and a Max Pool layer. In the decoding stage, the 512 dense features are decoded to the original input size by three  $3 \times 3$  two-dimensional deconvolutional layers of feature size 128, 64 and 32 respectively, followed by an element wise ReLU layer and an upsampling layer. The aim of this CNN is to reconstruct, from low-resolution fluorescence images, a high-resolution fluorophores-density spatial distribution  $\rho$ . To perform the training, we designed an original protocol to augment images from a publicly available open-source data-set<sup>25</sup> with synthetically generated illumination patterns. More precisely, we started from fluorescence microscopy images of biological targets, which we used as high resolution images target ground truth (GT). They represent the spatial distribution of the fluorophore molecules  $\rho$ . We pre-processed images from the repository by randomly extracting 10000 patches, each with a field of view 11 times larger than the Point Spread Function (PSF) of the optical system. We used 7500 of these patches for training, 2000 for validation, and 500 for testing. Then we “illuminated” the patches with speckle patterns as required for the standard blind-SIM approach. Speckle patterns<sup>30</sup>, which are light structures generated on a laser beam reflected by a rough surface, occur in the presence of disordered phase patterns. We generated numerically synthetic speckle patterns  $I(\mathbf{r})$  employing a model based on summing plane waves with random phases, random amplitudes, and a cutoff spatial frequency that enables us to control the average speckle grain size of  $d_{sp}$  (see Supplementary Materials). In a typical blind-SIM experiment, a few hundred of different illuminations replicas ( $N = 600$ ) are acquired to extract a high-resolution image. High-resolution fluorescent signal data are obtained simply by multiplying  $I \cdot \rho$ . Note that, to ensure generality, each patch and illumination appears only once in the dataset. This syntacti-



**Figure 2.** Resolution measurement: (a) A Siemens star of density  $\rho = 1 + \cos 2\theta$  (field of view is 13 times  $d_{PSF}$ ). From left to right: the ground truth GT, the LR deconvolved with BS-CNN, with SAI, and with Lucy Richardson and the low resolution LR. All the deconvolution have been obtained with  $N = 1000$ ; The last frame is the Low-Resolution one (LR) obtained from the GT convolving it with the PSF of FWHM  $d_{PSF}$  ( $\eta = 2.5$ ). The radii of circles highlighted in the figure correspond radius at which the Siemens star rays are resolved. (b) The Fourier transform contrast FTC versus the radius  $R$ . The dashed line represents the Rayleigh criterion. (c) Resolution improvement of the BS-CNN, Lucy Richardson and SAI as a function of  $\eta$  with error-bars. The maximum resolution improvement is 2.17. (d) From left to right: quarter of the Siemens star for the GT, BS-CNN, and M-SBL (e) azimuthal profiles from (d).

cally augmented data-set contains sufficient variability to incorporate the physical properties of the illumination patterns in the deconvolution process. The low-resolution images are then generated from the high-resolution ones applying a blurring convolutional kernel, with the size of the collection point spread function  $d_{PSF}$ . This convolution operation mimics the information loss effect due to a collection objective with a limited numerical aperture (see Supplementary Materials). The ratio between  $d_{PSF}$  and  $d_{sp}$

$$\eta = \frac{d_{PSF}}{d_{sp}} \quad (1)$$

is the parameter driving the effectiveness of the deconvolution. In a standard experiment, performed with an objective for both illuminated and collected light, focusing and collection PSF are identical (apart a small factor due to wavelength differences), that is  $d_{PSF} = d_{sp}$ . However, if two different optics are employed,  $d_{PSF}$  and  $d_{sp}$  could be different (see Supplementary Materials). In fact the Scattering Assisted Imaging technique<sup>11</sup> takes advantage of a reduced  $d_{sp}$  to improve resolution. Thus we studied the general case in which  $d_{sp} < d_{PSF}$  that is  $\eta \geq 1$ .

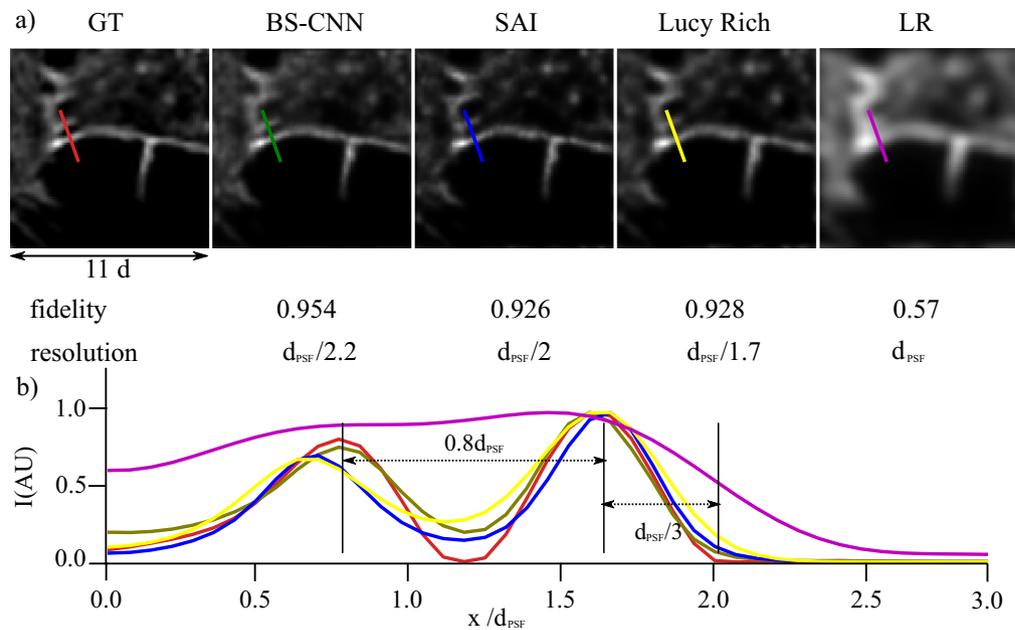
Before submitting data to the CNN for training the High Positive part of the Intensity of the signal HP is extracted from the low-resolution fluorescence signal

$$HP_n = (LR_n - LR)^+, \quad (2)$$

where  $()^+$  denotes that we keep only the matrix elements that are above zero and  $LR_n$  is a set of  $N$  Low Resolution images with an average  $LR$ . This data handling sparsifies the low-resolution images so that the algorithm makes use of the sparse nature of high intensity speckles (see<sup>11</sup> and Supplementary Materials). The model is trained using the adam<sup>31</sup> algorithm, with exponential decay rate for the 1st moment  $\beta_1 = 0.9$ , and exponential decay rate for the 2nd moment  $\beta_2 = 0.999$ . We complete training after 120 epochs with a learning rate  $lr = 0.001$  and we use as a loss the structural similarity index SSIM (see Supplementary Materials and<sup>32</sup>). We stop the training when the validation loss is not improving. We found that the best result is obtained when we use a batchsize of 32.

**Resolution measurement and fidelity.** In Figs. 2 and 3 we show the performance of the BS-CNN algorithm averaged over 1000 illuminations and compare it with the SAI algorithm, a custom local search deconvolution algorithm that exploits the properties of the illumination statistics (see Supplementary Materials and<sup>11</sup>) and with the Lucy Richardson algorithm, fixing  $\eta = 2.5$ .

In Fig. 2a we show the Siemens Star object: a standard to extract image resolution. It has been degraded to obtain the Low-Resolution image (image on the right, LR) and then super resolved with (from left to right) BS-CNN, SAI, and Lucy Richardson. The colored rings centered on the each star represent the radial value  $R$  where the Rayleigh criterion is satisfied. BS-CNN show the the best resolution. Resolution enhancement is indeed



**Figure 3.** Comparison of different algorithms. In panel (a) the images from left to right are the ground truth GT, BS-CNN, SAI and Lucy Richardson (Lucy Rich.) deconvolutions of the low-resolution LR. All deconvolutions are for  $N=1000$  and  $\eta=2.5$ . Below the images we label the *Fidelity* and the resolution expressed in terms of the PSF FWHM  $d_{PSF}$ . In panel (b) we show the line profile for each algorithm which has length three times  $d_{PSF}$ . The field of view for each image is  $11 d_{PSF}$ .

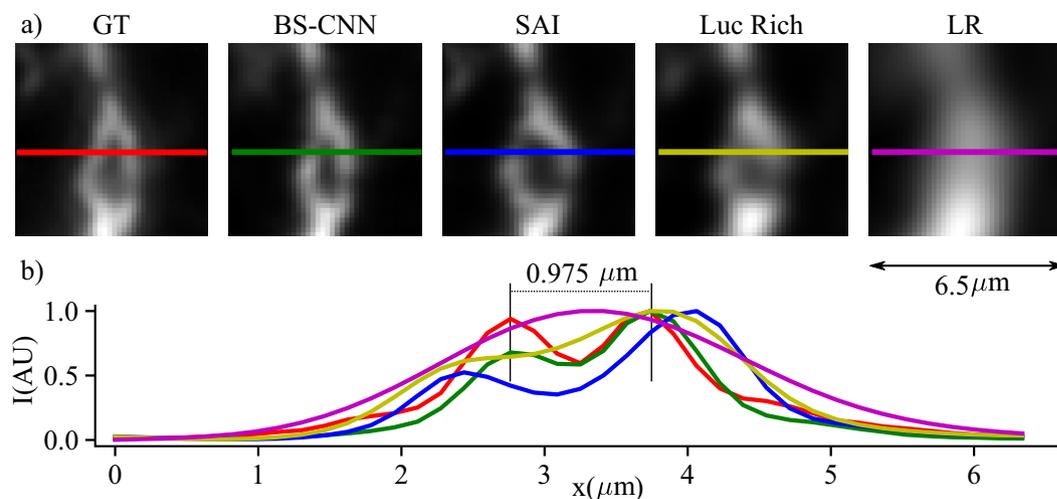
further quantified in Fig. 2b, where we show the Fourier transform contrast FTC (see Supplementary Materials and<sup>12</sup>) of the azimuthal profiles for different radii  $R$ , where the Rayleigh criterion is satisfied when  $FTC=0.08$ . The BS-CNN (Green Curve) provides the best performance over the investigated algorithms, since it achieves a higher FTC at all radii. In particular for  $\eta=2.5$  it achieves a Resolution Improvement (see Supplementary Materials) of 2.17, higher than the other algorithms. This is further confirmed by panel 2c where we report the resolution improvement versus the  $\eta$  parameter. The BS-CNN (green Curve) always outperforms other algorithms. Additionally BS-CNN produces images which are less affected by residual “granularity”, i.e. from speckle sized intensity fluctuations which are due to the limited number of illuminations. In panel 2d we test *Fidelity* (i.e. the capability to provide artifact free results) of BS-CNN. The *Fidelity* is obtained by computing the structural similarity index<sup>32</sup>, see Supplementary Materials) between deconvolved images and the ground truth. *Fidelity* is equal to 1 if the image resulting from deconvolution is identical to the GT while it is close to 0 for a poorly performing algorithm. BS-CNN provides a *Fidelity* of 0.71. This value has to be compared to other blind-SIM deconvolution techniques such as M-SBL<sup>7,33</sup> which, even providing slightly sharper images, deform the shape of the Siemens Star. The higher fidelity to the GT of BS-CNN is also confirmed by intensity profiles in Fig. 2 panel e providing a comparison of the profiles obtained with the different algorithms to the GT.

**Validation in experimental data.** In Fig. 3 we report results obtained applying BS-CNN to the analysis of live cells image, together with the degree of *Fidelity* of each technique. BS-SIM achieves a *Fidelity* of 95.4 per cent, in contrast to SAI and Lucy Richardson algorithm which achieve 92.6 and 92.8 per cent respectively. In particular, in the line profiles of the same figure (Fig. 3 panel b), we show two objects lying closer than the PSF resolution  $d_{PSF}$  so that they are not resolved in the low-resolution image. BS-CNN finds accurately both the position of the maxima and their relative intensity, while SAI and Lucy Richardson misplace the left maximum and retrieve it with a lower intensity.

Finally, we validate the BS-CNN approach on fully experimental data in Fig. 4, where speckles are generated employing a Digital Micromirror Device (DMD), with a random pattern controlling the illumination laser wavefront. Different illuminations are successively delivered to the sample by recasting the random pattern. To control the  $\eta$  parameter, the collection PSF (and in particular the  $d_{PSF}$  parameter) has been tuned with an optical iris placed between the collection objective and the eyepiece<sup>11</sup>. In Fig. 4 we again compare BS-CNN, SAI and Lucy Richardson for  $\eta = 2.5$  (maximum distance between distinguishable objects of  $2.37 \mu\text{m}$ ) and with  $N = 600$  illuminations. As we observe from the plot, BS-CNN can resolve objects up to a distance of  $0.975 \mu\text{m}$  which is 2.5 times smaller than the low-resolution object outperforming Lucy Richardson and SAI.

## Discussion

In summary we trained and studied a CNN able to deconvolve blind-SIM image stacks in a wide range of  $\eta$  values thus capable to be effective in different experimental configurations. BS-CNN has been trained employing images of biological samples (cell images) taken from a repository and “illuminated” by synthetic optical



**Figure 4.** Performance on experimental measurements: the output of the BS-CNN averaged over 600 low-resolution frames for mouse neurons prepared with Alexa Fluor 533 fluorophores along with the line profiles and comparison with the standard SAI and Lucy Richardson deconvolution. The field of view is  $6.50\ \mu\text{m}$ . The low resolution image has a resolution of  $2.37\ \mu\text{m}$ . BS-CNN achieves an image resolution of  $0.975\ \mu\text{m}$ .

patterns generated with a model for speckles, but also tested on real experimental data on biological samples. Even though the testing conditions are very different than those of the training, the BS-CNN confirms its effectiveness and robustness to data variability. We found that the BS-CNN outperforms other blind-SIM deconvolution algorithms and achieves an average resolution improvement with respect to the low-resolution images of 2.17, while it can reach up to a resolution improvement of 2.5 in specific spots. This improved performance is probably due to the neural networks generalization capability. The BS-CNN also produces images closer to the ground truth, as demonstrated quantitatively by the *Fidelity* parameter, indeed avoiding granularity or nonlinear effects. Moreover, we show an original approach for the CNN training, coupling a model of the illumination with experimental data. This strategy may be exported beyond microscopy, in all experimental configurations in which a modeling of the illumination is available, thus extending the data set available for training without degrading the final performance or the *Fidelity*.

## Methods

**Microscope details.** A continuous-wave (CW) laser emits light at the wavelength  $\lambda = 638\ \text{nm}$ . A first telescope enlarges the beam's size with two lenses respectively of focal lengths  $f_1 = 7\ \text{cm}$  and  $f_2 = 20\ \text{cm}$ . Then the light is modulated by a Digital Micromirror Device (DMD) by a random binary mask to generate different speckle illuminations. This plane is imaged through a  $4f$  lens system, made by a lens  $f = 20\ \text{cm}$ , and an objective (numerical aperture  $NA = 0.75$ ), to the sample. The objective collects the reflected light and the fluorescence that are then imaged by a third lens ( $f = 20\ \text{cm}$ ) to the camera (ORCA-Flash4.0 V3 Digital CMOS camera). A dichroic mirror and a long-pass spectral filter select the fluorescence signal. The camera's pixel size is  $6.5\ \mu\text{m}$ , while the overall magnification is 40. The spatial resolution of the microscope can be degraded by means of an iris placed in front of the third lens. The sample is composed by retina neurons prepared with Alexa Fluor 533 fluorophores. The fluorescence measurements have been performed by illuminating the sample with 600 speckles illuminations and an exposure time of 0.2 s.

In SAI each low resolution frame  $LR_n(\mathbf{r})$  is approximated with a synthetic frame  $g_n(\mathbf{r})$  composed by a sum of  $K$  Airy disks  $\tilde{h}(\mathbf{r})$  that have a size of  $d_{sp}$  and are located in random positions  $\mathbf{r}_{nk}$ .

**Scattering assisted imaging (SAI) deconvolution.** The task of the deconvolution algorithm is to find the fluorophore density  $\rho(\mathbf{r})$  given the low resolution set  $LR_n(\mathbf{r})$ . In SAI each  $LR_n(\mathbf{r})$  is approximated with a synthetic frame  $g_n(\mathbf{r})$  composed by a sum of  $K$  Airy disks  $\tilde{h}(\mathbf{r})$  that have a size of  $d_{sp}$  and are located in random positions  $\mathbf{r}_{nk}$ .

$$g_n(\mathbf{r}) = \sum_{k=1}^K P_{nk} \tilde{h}(\mathbf{r}; \mathbf{r}_{nk}). \quad (3)$$

The distribution of  $P_{nk}$  is a properly normalized multiplying factor respecting an exponential distribution. SAI finds the most probable configuration for the support  $\mathbf{r}_{nk}$  for each frame by a local search process that minimizes the mean absolute error between the  $HP_n$  and the convoluted frame of  $g_n(\mathbf{r})$

$$F = \sum_{\mathbf{r}} |(g_n(\mathbf{r}) * h(\mathbf{r})) - HP_n(\mathbf{r})|, \quad (4)$$

where  $*$  is the convolution operator. The final output of the algorithm  $G(\mathbf{r})$  is an average over the ensemble

$$G(\mathbf{r}) = \sum_{n=1}^N \frac{g_n(\mathbf{r})}{N}. \quad (5)$$

Received: 13 October 2021; Accepted: 11 May 2022

Published online: 21 May 2022

## References

1. Abbe, E. Beiträge zur theorie des mikroskops und der mikroskopischen Wahrnehmung. *Arch. Mikrosk. Anat.* **9**, 413–468. <https://doi.org/10.1007/BF02956173> (1873).
2. Sibarita, J.-B. Microscopy techniques. *Adv. Biochem. Eng./Biotechnol.* <https://doi.org/10.1007/b102215> (2005).
3. Hess, S. T., Girirajan, T. P. & Mason, M. D. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys. J.* **91**, 4258–4272. <https://doi.org/10.1529/biophysj.106.091116> (2006).
4. Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645. <https://doi.org/10.1126/science.1127344> (2006).
5. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–796. <https://doi.org/10.1038/nmeth929> (2006).
6. Gustafsson, M. G. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *J. Microsc.* **198**, 82–87. <https://doi.org/10.1046/j.1365-2818.2000.00710.x> (2000).
7. Min, J. *et al.* Fluorescent microscopy beyond diffraction limits using speckle illumination and joint support recovery. *Sci. Rep.* **3**, 2075. <https://doi.org/10.1038/srep02075> (2013).
8. Yeh, L.-H., Tian, L. & Waller, L. Structured illumination microscopy with unknown patterns and a statistical prior. *Biomed. Opt. Exp.* **8**, 695–711 (2017).
9. Idier, J. *et al.* On the superresolution capacity of imagers using unknown speckle illuminations. *IEEE Trans. Comput. Imaging* **4**, 87–98. <https://doi.org/10.1109/TCI.2017.2771729> (2018).
10. Mangeat, T. *et al.* Super-resolved live-cell imaging using random illumination microscopy. *Cell Rep. Methods*. <https://doi.org/10.1016/j.crmeth.2021.100009> (2021).
11. Leonetti, M., Grimaldi, A., Ghirga, S., Ruocco, G. & Antonacci, G. Scattering assisted imaging. *Sci. Rep.* **9**, 4591. <https://doi.org/10.1038/s41598-019-40997-6> (2019).
12. Mudry, E. *et al.* Structured illumination microscopy using unknown speckle patterns. *Nat. Photon.* **6**, 312–315. <https://doi.org/10.1038/nphoton.2012.83> (2012).
13. Barbastathis, G., Ozcan, A. & Situ, G. On the use of deep learning for computational imaging. *Optica* **6**, 921. <https://doi.org/10.1364/optica.6.000921> (2019).
14. Dong, C., Loy, C. C., He, K. & Tang, X. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**, 295–307. <https://doi.org/10.1109/TPAMI.2015.2439281> (2016).
15. Yang, W. *et al.* Deep learning for single image super-resolution: A brief review. *IEEE Trans. Multimed.* **21**, 3106–3121. <https://doi.org/10.1109/TMM.2019.2919431> (2019). 1808.03344.
16. Levin, A., Weiss, Y., Durand, F. & Freeman, W. T. Understanding and evaluating blind deconvolution algorithms. in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 1964–1971. <https://doi.org/10.1109/CVPR.2009.5206815> (IEEE, 2009).
17. Schuler, C. J., Burger, H. C., Harmeling, S. & Scholkopf, B. A Machine learning approach for non-blind image deconvolution. in *2013 IEEE Conference on Computer Vision and Pattern Recognition*. 1067–1074. <https://doi.org/10.1109/CVPR.2013.142> (IEEE, 2013).
18. Rivenson, Y. *et al.* Deep learning microscopy. *Optica* **4** (2017).
19. Wang, H. *et al.* Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nat. Methods* **16**, 103–110. <https://doi.org/10.1038/s41592-018-0239-0> (2019).
20. Ouyang, W., Aristov, A., Lelek, M., Hao, X. & Zimmer, C. Deep learning massively accelerates super-resolution localization microscopy. *Nat. Biotechnol.* **36**, 460–468. <https://doi.org/10.1038/nbt.4106> (2018).
21. Nehme, E., Weiss, L. E., Michaeli, T. & Shechtman, Y. Deep-STORM: Super-resolution single-molecule microscopy by deep learning. *Optica* **5**, 458. <https://doi.org/10.1364/OPTICA.5.000458> (2018).
22. Pu, T. *et al.* Unlabeled far-field deeply subwavelength topological microscopy (DSTM). *Adv. Sci.* **8**, 2002886. <https://doi.org/10.1002/advs.202002886> (2021).
23. Jin, L. *et al.* 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> (2020).
24. Christensen, C. N., Ward, E. N., Lio, P. & Kaminski, C. F. ML-SIM: A deep neural network for reconstruction of structured illumination microscopy images. *Biomed. Opt. Exp.* **12**, 2720 (2021).
25. Sero, J. E. & Bakal, C. <https://idr.openmicroscopy.org/webclient/?show=screen-1801note=idr0030-sero-yap/screen.A>.
26. Xie, W., Noble, J. A. & Zisserman, A. Microscopy cell counting and detection with fully convolutional regression networks. *Comput. Methods Biomech. Biomed. Eng. Imaging Visual.* **6**, 283–292. <https://doi.org/10.1080/21681163.2016.1149104> (2018).
27. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N. & Liang, J. 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> (2020).
28. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*. (Navab, N., Hornegger, J., Wells, W. M. & Frangi, A. F. eds.) . 234–241. (Springer, 2015).
29. Ye, J. C. & Sung, W. K. Understanding geometry of encoder–decoder CNNs. in *Proceedings of the 36th International Conference on Machine Learning*. Vol. 97. *Proceedings of Machine Learning Research* (Chaudhuri, K. & Salakhutdinov, R. eds.). 7064–7073. (PMLR, 2019).
30. Goodman, J. W. *Speckle Phenomena in Optics: Theory and Applications* (Roberts and Company Publishers, 2007).
31. Kingma, D. P. & Ba, J. *A Method for Stochastic Optimization*. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (Adam, 2014).
32. Wang, Z., Bovik, A., Sheikh, H. & Simoncelli, E. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612. <https://doi.org/10.1109/TIP.2003.819861> (2004).
33. Hojman, E. *et al.* Photoacoustic imaging beyond the acoustic diffraction-limit with dynamic speckle illumination and sparse joint support recovery. *Opt. Exp.* **25**, 4875–4886. <https://doi.org/10.1364/OE.25.004875> (2017).

## Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 713694. GR acknowledges the support by European Research Council Synergy grant ASTRA (No. 855923). ML thanks Project LOCALSCENT, Grant PROT. A0375-2020-36549, Call POR-FESR "Gruppi di Ricerca 2020".

## Author contributions

E.X. performed all calculations, E.X. and M.L. conceived this research, T.G. conducted the experiment, E.X., G.G., R.S., M.L. and G.R. analysed the results. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-12571-0>.

**Correspondence** and requests for materials should be addressed to E.X.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**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/>.

© The Author(s) 2022