From Stars to Molecules: AI Guided Device-Agnostic Super-Resolution Imaging

Dominik Vašinka, Filip Juráň, Jaromír Běhal, and Miroslav Ježek*
Department of Optics, Faculty of Science, Palacký University, 17. listopadu 12, 77900 Olomouc, Czechia

Super-resolution imaging has revolutionized the study of systems ranging from molecular structures to distant galaxies. However, existing deep-learning-based methods require extensive calibration and retraining for each imaging setup, limiting their practical deployment. We introduce a device-agnostic deep-learning framework for super-resolution imaging of point-like emitters that eliminates the need for calibration data or explicit knowledge of optical system parameters. Our model is trained on a diverse, numerically simulated dataset encompassing a broad range of imaging conditions, enabling robust generalization across different optical setups. Once trained, it reconstructs super-resolved images directly from a single resolution-limited camera frame with superior accuracy and computational efficiency compared to conventional methods. We experimentally validate our approach using a custom-built microscopy setup with ground truth emitter positions and demonstrate its versatility on astronomical and single-molecule localization microscopy datasets, achieving unprecedented resolution without prior information. Our findings establish a pathway toward universal, calibration-free super-resolution imaging, expanding its applicability across scientific disciplines.

I. INTRODUCTION

Every optical system is affected by the fundamental constraint on its resolution, known as the diffraction limit. This limitation restricts the ability to resolve structures significantly smaller than the wavelength of light divided by the numerical aperture of the optical system. Advances in super-resolution techniques have garnered significant interest across various fields by enhancing resolution beyond this inherent limit [1, 2]. In biomedical imaging, these techniques improve the quality of medical scans and enable detailed visualizations of cellular structures [3–5]. In material science, they allow the study of nanoscale properties of materials [6]. In astronomy, super-resolution techniques enable detailed object observation from space telescopes and probes, as well as resolving individual stars in densely packed star clusters [7, 8]. Many more applications rely on imaging objects composed of point-like or single-emitter sources. Localization microscopy provides detailed visualization of cellular structures at the molecular level [9–12]. In quantum physics, super-resolution aids in characterizing the optical properties and positioning of quantum dots [13] and enables precise site-resolved imaging of cold atoms in optical lattices [14, 15]. Whether the goal is to improve image resolution, localize emitting sources, or perform emitter tracking and counting, super-resolution techniques unlock new possibilities, providing unprecedented insight across diverse fields.

Super-resolution imaging can be achieved through various techniques. Interpolation-based approaches enhance resolution by inferring missing details from neighboring pixel information [16]. Tomographic methods reconstruct super-resolved images through inversion procedures of multiple perspectives [17, 18]. Statistical approaches leverage the stochastic nature of emitting sources [9, 19]. Numerous additional methods, including those based on deterministic algorithms [20], sparse representation [21], and deep learning [22, 23], have been developed. Regardless of the specific framework, super-resolution techniques can generally be classified into two categories: reconstruction and parameter estimation. Reconstruction aims to directly restore the whole super-resolved image [20, 24, 25], while estimation focuses on extracting key parameters and features, such as emitter positions [19, 26, 27]. In the following discussion, we will focus exclusively on the former, more universal approach: reconstruction techniques. As they do not require sparse or blinking sources, these techniques can cover a significantly broader range of applications.

Achieving accurate reconstruction of an object based on a single intensity image requires substantial knowledge about the imaging system under consideration. Central to this knowledge is an accurate estimation of the point spread function, which characterizes the response of the system to a point source of light. The reconstruction quality is inherently linked to the accuracy of the point spread function estimate; any inaccuracies will lead to poor reconstruction results [28]. However, its precise estimation entails additional calibration measurements involving an isolated emitting point source. This condition can prove very demanding, especially in systems with high concentration of emitters or low signal-to-noise ratio, which often necessitates preparing an additional finely tuned specimen for calibration purposes. Additionally, imaging systems commonly exhibit variations in the shape of the point spread function across the image. Consequently, reconstructing larger areas may require separate calibration for each segment [29].

Omitting the need for calibration data, some approaches can be adapted to perform a blind reconstruction, operating without explicit prior knowledge of the point spread function [30, 31]. Instead, these methods impose additional assumptions on the imaging system,

* Correspondence email address: jezek@optics.upol.cz

typically in the form of a predefined point spread function shape. During reconstruction, parameters of the point spread function, such as its width or other defining characteristics, are iteratively estimated directly from the input image. As a result, these techniques introduce imperfections by simultaneously reconstructing the observed object and the point spread function of the imaging system. Altogether, blind reconstruction remains a non-convex optimization problem, posing a challenging task in real-life scenarios and often yielding inferior results compared to non-blind algorithms.

Similar device-specific constraints arise with reconstruction using deep learning techniques. While these methods often outperform classical algorithms in both reconstruction and parameter estimation tasks [22, 24, 32–35], their calibration presents an even more significant obstacle. Unlike traditional approaches, these techniques operate on a data-driven paradigm, learning the reconstruction mapping by extracting relevant information from observed samples. However, due to the device-specific nature of a typical training process, the model can only be applied to a particular point spread function profile and additional factors such as emitter power, background noise distribution, and concentration of emitters, each requiring prior estimation from the calibration datasets [36]. Modifying these parameters necessitates remeasuring the calibration set and retraining the model, making it a resource-expensive procedure.

This work addresses the challenges of super-resolution techniques requiring calibration measurements for accurate image reconstruction. We present a novel approach to developing deep learning models, illustrating that their device-specificity can be overcome. The proposed device-agnostic approach, which extensively benefits from numerically simulating training datasets, produces a model capable of accurate reconstructions without calibration data or prior knowledge of imaging system parameters. Additionally, the trained model exhibits inherent robustness to non-uniformity and instability of imaging components. We validate the resolution-enhancement ability of our approach using experimentally acquired images and benchmark its performance against the well-established Richardson-Lucy deconvolution method. Furthermore, we highlight the device-agnostic versatility of our model by applying it to publicly available astronomy and microscopy datasets, achieving ultra high-resolution reconstructions without calibration or parameter estimation.

II. RESULTS

We developed a device-agnostic modeling network (DAMN) designed to reconstruct an intensity image of point-like emitting sources, as illustrated in Fig. 1. Unlike traditional approaches, the DAMN model operates using a single intensity frame without requiring additional information about the optical parameters of the imaging system. Without the need for calibration measurements or retraining, it can directly process images from diverse applications. Our model leverages the convolutional architecture of a neural network (refer to the Materials and Methods section), enabling it to reconstruct frames of varying shapes and sizes. Moreover, the reconstruction process is performed by a single-pass propagation through the neural network, resulting in rapid reconstruction of super-resolved images.

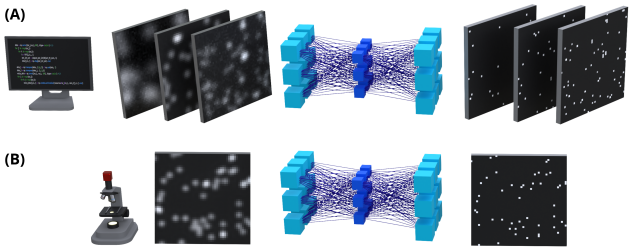

FIG. 1. Schematic representation of the DAMN approach and its application to intensity images of point-like emitting sources. (A) The model is trained using numerically simulated data pairs comprising resolution-limited noisy images alongside their super-resolved counterparts. Each training sample represents a unique combination of underlying optical parameters, such as the width of the point spread function and the signal-to-noise ratio. (B) Following training, this model is applied to enhance the resolution of experimental images acquired using a real-life imaging system.

We train the DAMN model using pairs of targeted ideal objects and their corresponding resolution-limited images generated through numerical simulation. The ensemble of data pairs covers various optical parameters, such as emitter power and concentration, background noise distribution, and even different shapes and widths of the point spread function. Their ranges were selected to cover the majority of realistic situations (additional details in Materials and Methods section). The extent of the synthetic training data is unprecedented in the astronomy or emitter microscopy fields. Subsequently, we utilized incremental learning techniques [37] for the model to acquire knowledge about all possible parameter combinations. As a result of the optimization and training process, this single model accurately reconstructs super-resolved images independently of the underlying parameter values. Consequently, without the need for retraining or calibration data, the model can be applied to diverse imaging systems and is resistant to their parameter changes, including inhomogeneity across a field of view and instability, such as drifts and fluctuations in time.

To characterize the effectiveness of the DAMN model, we conduct a comparative analysis with the Richardson-Lucy deconvolution based on Bayesian estimation with a uniform prior [38, 39]. This widely used iterative algorithm implements image reconstruction by repeatedly utilizing a known point spread function. We evaluate the

results using a separate test set of image pairs not included in the training data to assess the generalization ability for new samples. We quantify the performance of both methods using the mean absolute error, calculated between the reconstructed image and its target object, averaged over the whole test dataset. In contrast to the Richardson-Lucy algorithm, which requires (and was supplied with) prior knowledge of the imaging system, our DAMN model operates independently of any device-specific information. Despite this distinction, the device-agnostic model (without the prior information) outperforms the classical Richardson-Lucy deconvolution (with complete information on the point spread function) in terms of reconstruction quality, as discussed in the following text. Additionally, due to the iterative nature of the Richardson-Lucy algorithm, it requires three orders of magnitude higher computational time on average, further emphasize the advantages of the DAMN approach, see the Materials and Methods section.

Panels (A) - (C) in Fig. 2 illustrate the dependence of mean absolute error on the emitter power, the width of the point spread function, and the concentration of emitters, respectively. The results of these log-log graphs are characterized using an average value over the test dataset with a corresponding 90% confidence interval. Panel (A) contains dual horizontal axes representing the varying emitter power using both the signal-to-noise ratio (SNR) and the peak-to-noise ratio (PNR). For additional details regarding the explored parameters and evaluation of both approaches, see the Materials and Methods section. The DAMN model, represented by the red color, exhibits significantly better performance than the green-colored Richardson-Lucy algorithm. Similar results can also be observed in panels (B) and (C). Panel (B) depicts the error dependence on the point spread function width (consult the Materials and Methods section). The remaining panel (C) characterizes the dependence on the emitter concentration, referring to the collective number of emitters in a single image. As seen across all three panels, our device-agnostic approach provides up to two orders of magnitude more accurate reconstructions than the device-specific Richardson-Lucy deconvolution. The depicted errors were evaluated across the whole 50×50 pixel image and can be, alternatively, expressed as per-pixel errors by dividing the values by the number of pixels.

The panel (D) in Fig. 2 visualizes the error dependence on the point spread function shape, gradually transitioning from an Airy pattern to a Gaussian profile of the same width. As seen from this semi-logarithmic graph, the performance of both methods remains approximately constant during this transition. Such behavior is expected from the Richardson-Lucy deconvolution, as we always provide the correct point spread function profile. On the other hand, the DAMN model was trained using solely the exact Gaussian and Airy point spread functions constituting half of the dataset each. Despite this simplification in the training process, the graph unambiguously

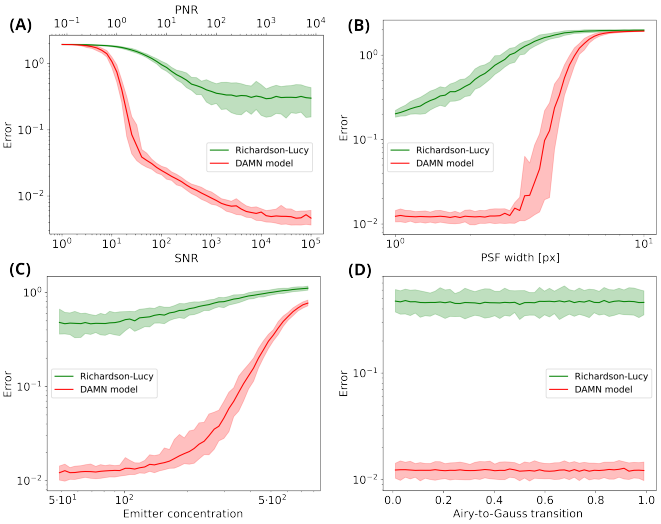

FIG. 2. The dependence of the mean absolute error on the signal-to-noise ratio (SNR), the width of a Gaussian point spread function (PSF), the number of emitters in the image (concentration), and the continuous transition between an Airy and a Gaussian PSF, respectively. The resulting averages of the green Richardson-Lucy algorithm and the red DAMN model are accompanied by their 90% confidence intervals over the test set. Panel (A) is provided with a secondary horizontal axis recalculating the SNR values to the peak-to-noise ratio (PNR). Across all panels, the DAMN model consistently outperforms the Richardson-Lucy deconvolution by up to two orders of magnitude. The optical parameters not investigated in a given panel have the following values: SNR = 500, the average noise intensity = 10, the concentration = 50, and the Gaussian PSF with $\sigma = 2$ px.

demonstrates the adaptive ability of the DAMN approach to generalize to previously unseen point spread function shapes. This adaptability is beneficial during training, as it significantly reduces the data and simulation complexity required for training device-agnostic models. Altogether, the presented panels display numerous benefits of the DAMN approach, proving it superior to the state-of-the-art algorithm.

Demonstrating the validity of our approach beyond the simulated data, we applied it to reconstruct experimentally acquired images. Explicitly for this, we developed an advanced optical microscopy setup shown in Fig. 3, which provides complete control over the spatial distribution of emitting point-like sources. An incoherent light illuminates a digital micromirror device (DMD), onto which a mask representing the targeted ground truth image is imposed. By configuring the micromirrors, we can prepare a mask containing an arbitrary number of sources positioned at desired locations. This mask is then re-imaged by a high-demagnification and high-resolution preparation system, creating point-like emitters and thus forming a sample object. Subsequently, a low-resolution imaging microscope projects the emitters onto a camera. The resulting intensity profile captured by the camera serves as the resolution-limited input for reconstruction

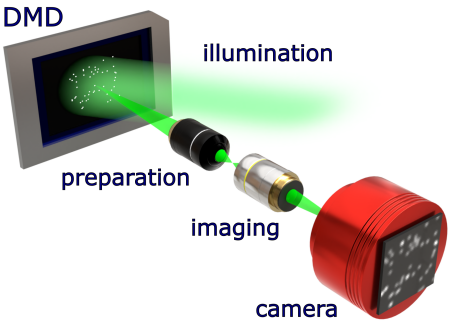

FIG. 3. Schematic illustration of the optical setup used to collect experimental data pairs. A ground-truth mask is imposed on the digital micromirror device (DMD) by configuring its mirrors. An incoherent illumination light reflected by these mirrors impinges a high-resolution preparation system. The DMD-imposed mask is re-imaged into the front sample plane of the preparation system, creating point-like emitters with the intended spatial distribution. The imaging part of the setup, comprised of a low-resolution microscope objective, images the sample-plane emitters onto a camera. The resulting camera-captured intensity image and the DMD-imposed mask represent the experimental data pairs.

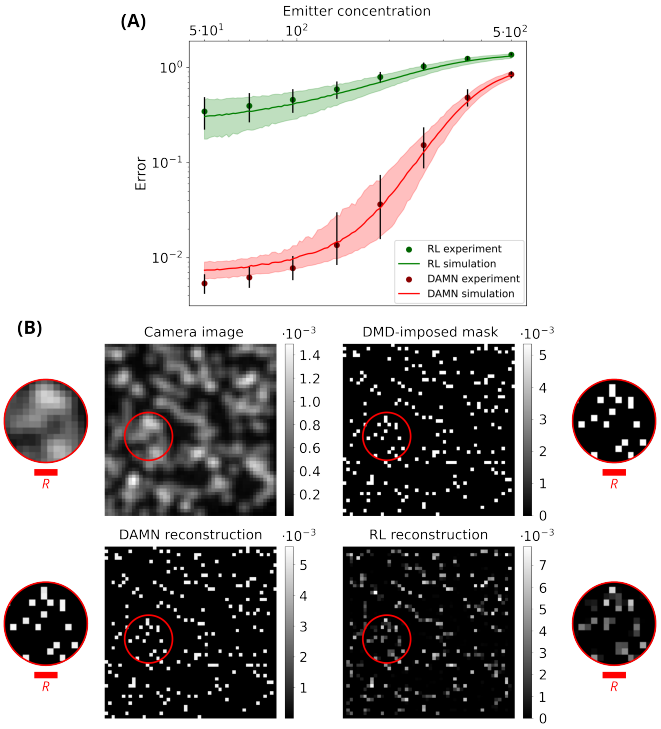

FIG. 4. (A) The red and green dots represent the mean absolute error between the DMD-imposed masks and the super-resolved images reconstructed by each method. These errors were evaluated across various emitter concentrations. The accompanying continuous lines depict error values derived from simulated data using optical parameters estimated for our imaging system. (B) A typical camera image containing nearly 200 emitters, alongside its corresponding DMD-imposed mask and each method reconstructions. The circled areas contain a magnified region for easier visual comparison. It is evident that the DAMN model significantly outperforms the Richardson-Lucy (RL) algorithm even in regions where mutual emitter distance is well-below the Rayleigh resolution limit $R = 3.9$ px (inset scalebars).

methods. Together with the ground truth mask, this setup provides experimental test data pairs, which allow for exact metric quantification of each reconstruction method performance. For details, see the Materials and Methods section.

Using this setup, we investigated the performance of both methods while varying the concentration of emitters. Panel (A) in Fig. 4 illustrates the mean absolute error between reconstructed images and their corresponding target objects using the experimental data, shown as dot markers. For comparison, the continuous lines represent the results obtained using simulated data with corresponding optical parameters. The DAMN model outperforms the Richardson-Lucy deconvolution by orders of magnitude, even when applied to real-world images acquired from experimental measurements. This improvement is achieved despite operating significantly below the resolution limit of our optical system. To further emphasize these findings, panel (B) in Fig. 4 shows a representative experimental image containing nearly 200 emitters and its reconstruction by each method. As observed, the DAMN model offers a near-perfect reconstruction of the original image without the need for any calibration. The Richardson-Lucy deconvolution exhibits severe deviations and artifacts despite knowing the point spread function of the imaging setup. This experimental demonstration highlights the benefits of our device-agnostic approach, which significantly outperforms the established state-of-the-art algorithm requiring prior knowledge of the point spread function.

We have also explored the benefits of incorporating upsampling layers into the DAMN convolutional architecture, further extending its capabilities. The modified network reconstructs images with an eightfold increase in dimensions, leading to an even more prominent improvement in the resolution. We applied this model to an astronomy image of a dense star cluster from observations of the southern spiral galaxy NGC 300 in the Sculptor constellation by the European Southern Observatory [40], see Fig. 5 (A). Panel (B) depicts a magnified segment of the galaxy with marked regions of interest, while panel (C) presents their high-resolution reconstruction produced by the DAMN model. Despite the stars in all three regions being below the resolution limit of the imaging system, the DAMN reconstruction clearly reveals them. Notably, a detailed examination of region III shows two distinct intensity peaks. While the original data resemble a single structure, our reconstruction indicates that it might be two proximal stars hidden deep below the resolution limit. Confirming this hypothesis would require further analysis using additional data. While the DAMN model can process the entire

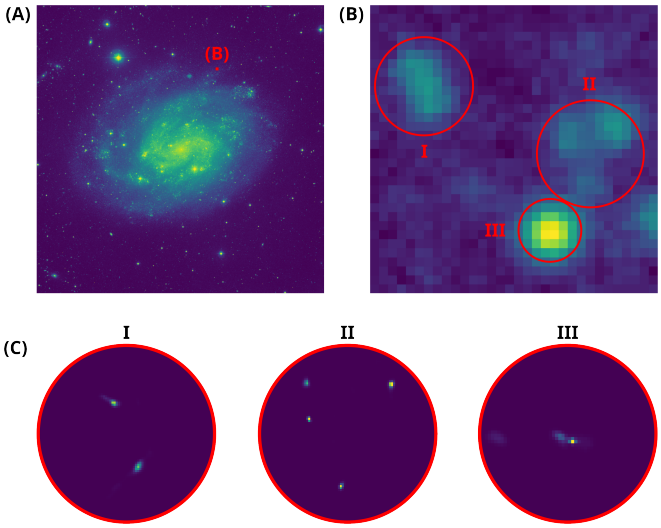

FIG. 5. Demonstration of the DAMN model super-resolving capabilities on an astronomy image of a dense star cluster. (A) Intensity image of the spiral galaxy NGC 300 acquired by European Southern Observatory. (B) Magnified region of the galaxy, highlighting key areas of interest. (C) High-resolution reconstructions of these regions generated by the DAMN model, revealing details of individual stars. These reconstructions do not use any calibration data or prior information on the employed imaging system. The region III reconstruction indicates the presence of a secondary source completely beyond the optical resolution of the telescope.

7600×7600 pixel image, saturation artifacts in the central part of the original data (Fig. 5 (A)) led us to demonstrate the reconstruction of smaller outer regions of the galaxy. Moreover, as shown in (C), the resolution enhancement is so significant that proper visualization requires considerable zoom into the reconstructed areas.

Additionally, we applied the same DAMN model, without retraining, to a publicly available dataset [41, 42] acquired for the single-molecule localization microscopy challenge. Fig. 6 depicts the performance evaluation using these 500 tubulin images of 128×128 pixels with high molecule concentration (A). Panel (B) shows a reference image provided by the SOSplugin [43], a least-square localization method assuming a Gaussian point spread function, which participated in the challenge and is available as a plugin for ImageJ. Panel (C) depicts a tubulin reconstruction from Deep-STORM [24], a state-of-the-art deep learning model for image reconstruction in localization microscopy. Unlike our approach, Deep-STORM operates in a device-specific regime, requiring extensive optical system information, such as camera specification, point-spread-function model, approximate SNR, and expected emitter density. In comparison, the DAMN model directly processes the dataset without calibration or prior knowledge of the imaging system and provides the super-resolved tubulin image, see panel (D). The highlighted areas show a close-up comparison of all three approaches. Our DAMN model accurately captures the underlying tubulin structure, achieving resolution significantly surpassing that of the SOSplugin localization method. Unlike the localization method, the DAMN model clearly resolves the two vertically oriented, overlapping microtubules and reconstructs the horizontal one without gaps. Moreover, a close inspection reveals that our model provides a sharper image with a cleaner and thinner microtubule structure compared to the device-specific Deep-STORM. Panel (E) visualizes the projection of the microtubule profile over the yellow segment, demonstrating the resolution improvement. This result ultimately underlines the reconstruction ability of device-agnostic models in real-world applications. Although the presented model provides super-resolved reconstructions, the same paradigm of device-agnostic deep learning can be applied to localization tasks, i.e., directly predicting the positions and intensities of the emitting sources. Such a model could be highly advantageous for single-molecule localization microscopy of densely populated samples.

III. CONCLUSION

We demonstrated a device-agnostic approach to super-resolution imaging of point-like emitting sources that utilizes deep learning techniques. The presented method solves the long-standing problems of measuring calibration data and estimating parameters accompanying the training of a deep learning model. Using only numerically simulated data, we develop a model capable of reconstructing intensity images using a single frame without explicit knowledge of the imaging system. This model can be applied to images of arbitrary shape and size originating from various optical systems, and the calibration-free nature grants it resistance against non-uniformity and drifts of the optical parameters. We compared this device-agnostic model with the well-established Richardson-Lucy deconvolution algorithm. The analysis results undeniably show the superior performance of our approach, outperforming the Richardson-Lucy algorithm by orders of magnitude in terms of both reconstruction accuracy and speed. Moreover, we designed an advanced optical setup for acquiring experimental images of emitting sources together with their precise ground truth references. Providing full control over the spatial distribution of emitters, this setup enables exact performance quantification of any reconstruction method. Using these experimental data pairs, we verified the superior performance of our approach.

To further demonstrate the universality of our approach, we applied the model to astronomy and microscopy data, achieving significant resolution improvements in both domains. We reconstructed a high-resolution tubulin image from single-molecule localization microscopy data. While in astronomy, we enhanced the resolution of a densely packed star cluster image. Notably, all DAMN reconstructions were performed without any prior knowledge of the optical setup, data preprocess-

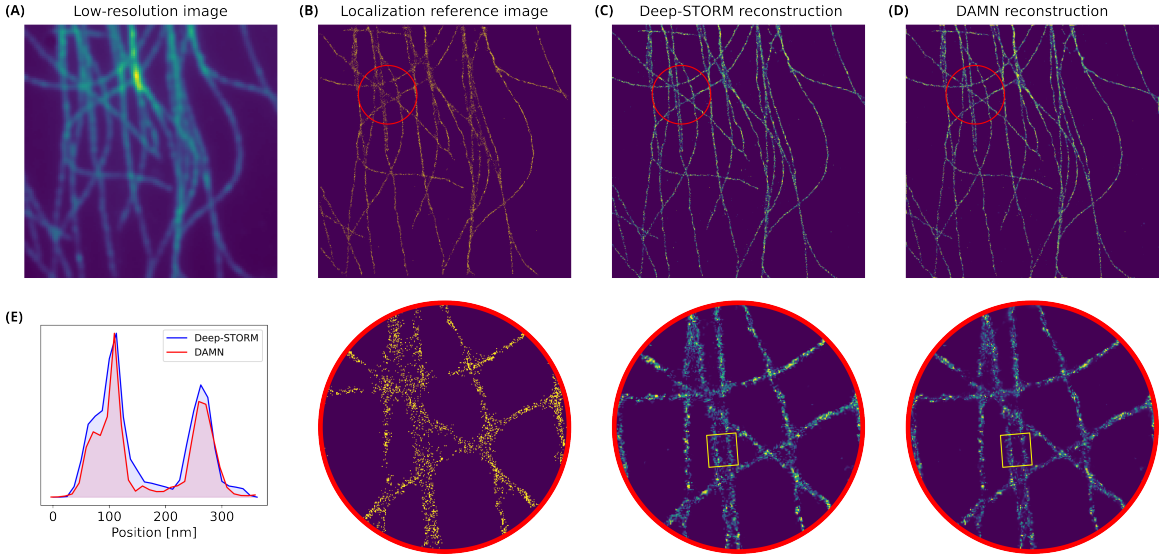

FIG. 6. Demonstration of the DAMN model on a high-concentration tubulin dataset from the single-molecule localization microscopy challenge. (A) A low-resolution image integrated from the set of 500 measured images, containing numerous emitting molecules. (B) A reference binary map of localized molecule positions derived from the dataset. (C) A reconstruction generated by Deep-STORM, a device-specific deep learning method calibrated using detailed information on the imaging setup. (D) A super-resolved image produced from the calibration-free, device-agnostic model without prior information. Panels (B-D) include inset images with a magnified region highlighting the resolution details. (E) Projection of the microtubule profile over the yellow rectangle segment.

ing, or parameter estimation. Despite this, the obtained results surpass the state-of-the-art methods. With sufficient computation resources, even further improvements can be achieved by incorporating more complex data simulations and larger network architectures. By taking this first step, our work lays a foundation for universal image reconstruction, entirely independent of optical settings, thus contributing to the advancement of image processing in various fields.

DATA AVAILABILITY

The code and data supporting the results of this study are publicly available on GitHub [44].

ACKNOWLEDGMENTS

We acknowledge the massive use of cluster computing resources provided by the Department of Optics, Palacký University Olomouc. We thank J. Provazník for maintaining the cluster and providing support.

AUTHOR CONTRIBUTIONS

M.J. conceived the idea of a device-agnostic approach to image super-resolution and supervised the project. D.V. developed deep-learning models and performed numerical simulations and data processing. F.J. and J.B. developed the optical experiment, performed data acquisition, and participated in data processing. J.B. supervised the experimental part of the project. D.V. wrote the manuscript, and all authors were involved in revising the manuscript.

FUNDING

This work was supported by the Czech Science Foundation (project 21-18545S), and the Ministry of Education, Youth, and Sports of the Czech Republic (project OP JAC CZ.02.01.01/00/23_021/0008790). DV acknowledges the support by Palacký University (project IGA-PrF-2024-008).

DISCLOSURES

The authors declare no conflict of interest.

IV. MATERIALS AND METHODS

A. Data Simulation

A resolution-limited image can be characterized by its underlying optical parameters. In our case, these parameters include the shape and width of the point spread function, the background noise intensity, the emitter

power representing the number of emitted photons, and the number of emitters present within a 50×50 pixels field of view of an image, which we term emitter concentration. Consequently, all possible resolution-limited images form a high-dimensional space. To implement the device-agnostic framework with a data-driven deep learning algorithm, the model must observe data samples covering the space during training. Therefore, the approach highly benefits from using simulated data pairs, as collecting a sufficient amount of experimentally acquired samples would be exceptionally time-consuming. Additionally, data simulation allows using incremental learning techniques (see Deep Learning Model), which are ideal for applications with large datasets. The following ranges of optical parameters were used for the data simulation: the emitter power $\in \left[1, 10^5\right]$, the average noise $\in [1, 100]$, the emitter concentration $\in [5, 500]$, and the point spread function width $\in \left[10^{-0.25}, 10^{1.25}\right]$ px $\approx [0.5, 17.75]$ px. For the first time in the emitter visualization community, data with such a broad scale of imaging parameters have been generated and used to train an artificial neural network.

The generation process of a single simulated data sample follows these steps. First, we generate the emitter concentration, followed by assigning each emitter its power and pixel position in the image. Next, we perform a convolution with the point spread function to simulate the effects of finite resolution. Subsequently, shot noise is added to each pixel of the image. With knowledge of the average noise intensity and emitter power, we can calculate the signal-to-noise ratio of the generated sample as $\mathrm{SNR} = \frac{\text{emitter power}}{\text{average noise intensity}}$. Alternatively, we can calculate the peak-to-noise ratio, $\mathrm{PNR} = \frac{\text{peak emitter intensity}}{\text{average noise intensity}}$, using the maximum value of the convolved emitter intensity. It is worth noting that the simulation process can be further enhanced by incorporating additional features and parameters, such as point spread function asymmetry and aberrations [45, 46]. For this showcase study, we opted to simplify the simulation to reduce the computational demand during training. Despite this simplification, the DAMN model achieves excellent results on experimental data and demonstrates significant generalization ability beyond the expected regime, as shown in Fig. 2 (D).

To perform the convolution, we generate the convolution matrix of the point spread function using two distinct parameters - shape and width. The shape parameter distinguishes between an Airy pattern A, simulating low numerical aperture scenarios, and a two-dimensional Gaussian pattern G, typically used for cases with higher numerical aperture values [47]. The convolution matrix is normalized to a unit sum to conserve the emitter power. The pattern functions can be expressed as

$$A(r,\sigma) \sim \left[\frac{\mathrm{J}_1(\frac{2r}{\sigma})}{\frac{r}{\sigma}}\right]^2,$$

$$G(r,\sigma) \sim \exp\left(-\frac{r^2}{\sigma^2}\right),$$

where r is the distance from the center, and J_1 is the first-order Bessel function of the first kind. The variable σ is the width parameter, which dictates the full width at half maximum of the point spread function as $2\sigma\sqrt{\ln 2}$.

B. Deep Learning Model

The DAMN model is a deep convolutional neural network comprising 35 hidden layers with 71 channels each and a final output layer with a single channel. Information flows through each hidden layer via trainable 7×7 filters that apply local convolutional operations, allowing the network to detect localized patterns. These transformations are followed by a LeakyReLU activation function with a 0.05 negative slope, which introduces a computationally efficient non-linearity without causing the dying ReLU problem [48]. Since we normalize the input samples to a unit sum before processing, the output layer utilizes a softmax activation function [49] to preserve this condition and ensure non-negative values. To prevent overfitting, each hidden layer is paired with dropout regularization at a 0.01 rate, randomly setting a fraction of input units to zero during training [50]. This regularization encourages the network to learn robust features that do not rely on any specific neurons. As the DAMN model consists solely of convolutional layers, it can process images of arbitrary shapes and sizes without retraining. Additional resolution improvement is achieved by incorporating upsampling layers into the convolutional architecture. These layers increase the image dimension by repeating its rows and columns. The modified model applied in Fig. 5 and 6 contains three upsampling layers that double the input dimensions, resulting in their eightfold increase. This architecture makes the network highly flexible and effective for image-processing applications.

The presented model architecture results from extensive optimization utilizing a mesh adaptive direct search using the Nomad library [51], as well as numerous additional manual adjustments. Altogether, our model contains over 8 million trainable parameters and was trained incrementally on nearly 3 million simulated images, with a 25% validation set split. Incremental learning techniques gradually present the training data to the network. These techniques involve training the model using randomly generated data, which are continuously replaced with newly generated samples throughout the training process [37]. The level of network complexity and dataset size is unprecedented in the field of emitter reconstruction

and localization. Nevertheless, the model remains relatively modest compared to modern large language models, suggesting further potential to integrate additional model features by increasing complexity.

The training of the DAMN model follows the backpropagation algorithm, which computes the gradient of the loss function with respect to each weight in the network using the chain rule. We use a mean squared error (MSE) loss, calculated between the predicted I_1 and target I_2 images and averaged over a batch of data samples. Given a set of two-dimensional image pairs, we calculate the loss as

$$\mathrm{MSE} = \frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L}\left[I_1\left(j,k,l\right) - I_2\left(j,k,l\right)\right]^2,$$

where J is the batch size, K and L are the image dimensions, and $I(j,k,l)$ is the intensity of the j-th image in the batch at the pixel position (k,l). The gradients obtained from this loss are used to iteratively update the weights with each batch of 128 training samples. To improve convergence towards the minimum loss, we implement the Adam optimizer [52], which incorporates adaptive moment estimation. Adam updates the weights using both the first and second moments of the gradient, making the descent more efficient. A portion of the training data is reserved as a validation set to actively monitor the convergence using the mean absolute error (MAE) metric

$$\mathrm{MAE} = \frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L}\left|I_1\left(j,k,l\right) - I_2\left(j,k,l\right)\right|.$$

The metric provides an additional measure of model performance that is, unlike a loss function, not directly optimized during training.

C. Richardson-Lucy Algorithm

The Richardson-Lucy algorithm is an iterative method used to restore an image blurred by a known point spread function [38, 39]. This sophisticated and highly flexible algorithm is derived from the Bayesian probability theory with a flat prior. It maximizes the posterior probability that the observed image results from the estimated image convolved with the given point spread function. Assuming a multinomial distribution of detection events, the Richardson-Lucy is equivalent to the expectation-maximization algorithm for likelihood maximization in positive linear inverse problems [53, 54]. The quality of the reconstructed image relies heavily on accurate knowledge of the point spread function; discrepancies can lead to artifacts and inaccurate reconstructions. Since we simulate the testing data, the Richardson-Lucy algorithm can use the precise point spread function, making it an excellent method for providing a state-of-the-art baseline for evaluating the DAMN model.

In addition to the point spread function P, the iterative algorithm requires an initial guess for the reconstruction, $I^{(0)}$. A common practice is to set this guess as the observed resolution-limited image I_{blurred} or to start with a uniform image. We explored both approaches and chose to set $I^{(0)} = I_{\mathrm{blurred}}$, as there was a negligible difference in accuracy and computational speed. Then, an iterative procedure updates the estimation at the $k+1$ step

$$I^{(k+1)} = I^{(k)} \cdot \left(\frac{I_{\mathrm{blurred}}}{I^{(k)} \circledast P} \circledast P^*\right),$$

where $\circledast$ is convolution, and P^* denotes the flipped point spread function. Both the initial guess $I^{(0)}$ and the point spread function P need to be normalized to a unit sum. We implement two stopping criteria for the algorithm: either the mean absolute error between successive iterations becomes smaller than 10^{-10} per pixel, or the number of iterations exceeds 10^6. The first criterion is typically met in the majority of cases.

Furthermore, we evaluated the computation time required by the Richardson-Lucy algorithm. As the time depends on the underlying optical parameters of an image, we have simulated an extensive test dataset containing approximately 300,000 randomly generated samples. Using an AMD EPYC 7443P 24-core processor, the Richardson-Lucy needs almost five days to reconstruct this dataset. In comparison, the DAMN model utilizing the same CPU resources reconstructs the data in a little over two hours. Moreover, the DAMN model with a GPU-enabled operation requires only approximately eight minutes for the same task when using an NVIDIA RTX A5500 GPU.

D. Experiment

The optical setup, which provides resolution-limited experimental images with their corresponding target objects, consists of three segments, see Fig. 7. The first segment prepares the illumination of the digital micromirror device (DMD). An incoherent broadband light with a central wavelength of 555 nm, produced by a laser-induced fluorescence light source (Crytur MonaLIGHT B01), is coupled to a 4 mm diameter light guide. The emerging light is collimated by a microscope objective $\mathrm{MO_s}$ (OLYMPUS 20x/0.4). An iris diaphragm D is positioned proximal to the objective. This diaphragm is re-imaged onto the DMD (Texas Instruments DLP LightCrafter 6500) by a telescope consisting of two achromatic doublets, $\mathrm{L_{s1}}$ and $\mathrm{L_{s2}}$, with focal lengths of 50 mm and 150 mm, respectively. The telescope improves the illumination homogeneity, and re-imaging the diaphragm reduces reflections from the passive parts of the DMD.

In the second setup segment, point-like emitters with the targeted spatial distribution are prepared. By addressing the DMD, we impose a ground truth mask with each pixel comprising a 5×5 micromirror grid with a

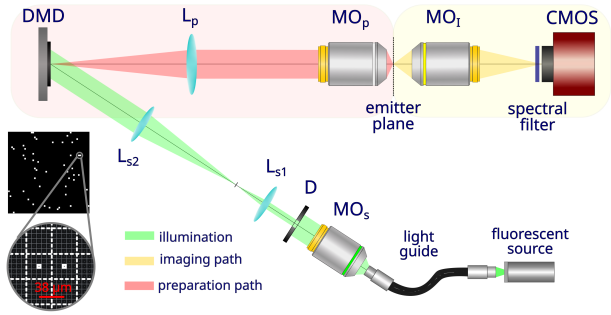

FIG. 7. A detailed visualization of the optical setup. In the illumination path, we generate an incoherent light collimated by a microscopic objective MO_s. The subsequent iris diaphragm D and two achromatic doublets, L_{s1} and L_{s2}, improve the illumination homogeneity and reduce undesired reflections from the passive DMD parts. In the preparation path, the DMD-induced ground truth mask is re-imaged by an achromatic doublet L_p and a high-resolution microscope objective MO_p, creating point-like emitters in the emitter plane. Lastly, this plane is projected by a low-resolution microscope objective MO_I onto a camera with a mounted spectral filter.

7.6 μm micromirror pitch. As we address only the central micromirror of the grid, the closest distance between two emitting sources is 38 μm in the DMD plane, see Fig. 7. This mask is re-imaged into the front focal plane of MO_p by a preparation telescope consisting of a 200 mm focal-length achromatic doublet L_p and a high-resolution microscope objective MO_p (OLYMPUS 100x/0.9). As the point-like emitters, of the approximate 310 nm full width at half maximum and 340 nm minimal distance, are created in this plane, we term it the emitter plane.

Unlike the first two segments, which prepare the emitters and provide complete control over their spatial distributions, the final segment represents an imaging setup with optical parameters unknown to the DAMN model. A low-resolution microscope objective MO_I (OLYMPUS 10x/0.25) projects the emitter plane onto a CMOS camera (ZWO-ASI 178MM, 2.4 μm square pixels) with a mounted spectral filter (central wavelength of 532 nm with 10 nm spectral width). The MO_I resolution is approximately 1.13 μm, regarding the full width at half maximum in the emitter plane. In comparison, the 340 nm minimal emitter distance is 3.3× below this resolution limit. The emitter plane is imaged with an approximate 230× magnification to match the major part of the camera chip. The exposure time for each measurement is set to 5.7 seconds to utilize the majority of the 14-bit dynamic range of the camera while avoiding its saturation. The captured image is down-scaled by a factor of 32 to match the 50×50 ground-truth mask size to enable metric evaluations. Using this imaging setup, we collected the experimental dataset of resolution-limited camera images and their corresponding ground truth masks for various emitter concentrations, see Fig. 4 (A). Additionally, we measured a high-SNR calibration set for precise estimation of the point spread function, which we provided only to the Richardson-Lucy algorithm. Besides the concentration ranging from 50 to 500 emitters, the optical parameters in the collected dataset exhibit the following estimated values: SNR = 2300, average noise intensity = 10, and an Airy-shaped kernel with $\sigma = 2.05$ px, equivalent to the 3.4 px full width at half maximum. The Rayleigh resolution limit of our imaging system, given by the radius R of the first dark ring in the Airy pattern, is approximately 3.9 px, see Fig. 4 (B).

[1] Lepcha, D. C., Goyal, B., Dogra, A. & Goyal, V. Image super-resolution: A comprehensive review, recent trends, challenges and applications. *Information Fusion* **91**, 230–260 (2023).

[2] Astratov, V. N. *et al.* Roadmap on label-free super-resolution imaging. *Laser and Photonics Reviews* **17**, 12 (2023).

[3] Greenspan, H. Super-Resolution in Medical Imaging. *The Computer Journal* **52**, 43–63 (2008).

[4] Sahl, S. J., Hell, S. W. & Jakobs, S. Fluorescence nanoscopy in cell biology. *Nature Reviews Molecular Cell Biology* **18**, 685–701 (2017).

[5] Schermelleh, L. *et al.* Super-resolution microscopy demystified. *Nature Cell Biology* **21**, 72–84 (2019).

[6] Pujals, S., Feiner-Gracia, N., Delcanale, P., Voets, I. & Albertazzi, L. Super-resolution microscopy as a powerful tool to study complex synthetic materials. *Nature Reviews Chemistry* **3**, 68–84 (2019).

[7] Starck, J. L., Pantin, E. & Murtagh, F. Deconvolution in astronomy: A review. *Publications of the Astronomical Society of the Pacific* **114**, 1051–1069 (2002).

[8] Bertero, M., Boccacci, P., Desiderà, G. & Vicidomini, G. Image deblurring with Poisson data: from cells to galaxies. *Inverse Problems* **25**, 123006 (2009).

[9] Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods* **3**, 793–796 (2006).

[10] Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006).

[11] Hess, S. T., Girirajan, T. P. & Mason, M. D. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal* **91**, 4258–4272 (2006).

[12] Lelek, M. *et al.* Single-molecule localization microscopy. *Nature Reviews Methods Primers* **1**, 39 (2021).

[13] Lidke, K. A., Rieger, B., Jovin, T. M. & Heintzmann, R. Superresolution by localization of quantum dots using blinking statistics. *Optics Express* **13**, 7052–7062 (2005).

[14] McDonald, M., Trisnadi, J., Yao, K.-X. & Chin, C. Superresolution microscopy of cold atoms in an optical lattice. *Physical Review X* **9**, 021001 (2019).

[15] Impertro, A. *et al.* An unsupervised deep learning algorithm for single-site reconstruction in quantum gas microscopes. *Communications Physics* **6**, 166 (2023).

[16] Thévenaz, P., Blu, T. & Unser, M. *Image Interpolation and Resampling*, 393–420 (Elsevier, 2000).

[17] Luo, W., Greenbaum, A., Zhang, Y. & Ozcan, A. Synthetic aperture-based on-chip microscopy. *Light: Science & Applications* **4**, e261–e261 (2015).

[18] Bianco, V. *et al.* Label-free intracellular multi-specificity in yeast cells by phase-contrast tomographic flow cytometry. *Small Methods* **7**, 2300447 (2023).

[19] Balzarotti, F. *et al.* Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science* **355**, 606–612 (2017).

[20] Bertero, M., Boccacci, P. & De MoI, C. *Introduction to Inverse Problems in Imaging* (CRC Press, 2021). URL `http://dx.doi.org/10.1201/9781003032755`.

[21] Yang, J., Wright, J., Huang, T. S. & Ma, Y. Image super-resolution via sparse representation. *IEEE Transactions on Image Processing* **19**, 2861–2873 (2010).

[22] Rivenson, Y. *et al.* Deep learning microscopy. *Optica* **4**, 1437 (2017).

[23] Moen, E. *et al.* Deep learning for cellular image analysis. *Nature Methods* **16**, 1233–1246 (2019).

[24] Nehme, E., Weiss, L. E., Michaeli, T. & Shechtman, Y. Deep-storm: super-resolution single-molecule microscopy by deep learning. *Optica* **5**, 458 (2018).

[25] Meiniel, W., Olivo-Marin, J.-C. & Angelini, E. D. Denoising of microscopy images: A review of the state-of-the-art, and a new sparsity-based method. *IEEE Transactions on Image Processing* **27**, 3842–3856 (2018).

[26] Schmidt, R. *et al.* Minflux nanometer-scale 3D imaging and microsecond-range tracking on a common fluorescence microscope. *Nature Communications* **12**, 1478 (2021).

[27] Li, Y. *et al.* Global fitting for high-accuracy multi-channel single-molecule localization. *Nature Communications* **13**, 3133 (2022).

[28] Booth, M., Andrade, D., Burke, D., Patton, B. & Zurauskas, M. Aberrations and adaptive optics in super-resolution microscopy. *Microscopy* **64**, 251–261 (2015).

[29] Xu, F. *et al.* Three-dimensional nanoscopy of whole cells and tissues with in situ point spread function retrieval. *Nature Methods* **17**, 531–540 (2020).

[30] Kundur, D. & Hatzinakos, D. Blind image deconvolution. *IEEE Signal Processing Magazine* **13**, 43–64 (1996).

[31] Campisi, P. & Egiazarian, K. *Blind Image Deconvolution* (CRC Press, 2007).

[32] Jin, K. H., McCann, M. T., Froustey, E. & Unser, M. Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing* **26**, 4509–4522 (2017).

[33] Speiser, A. *et al.* Deep learning enables fast and dense single-molecule localization with high accuracy. *Nature Methods* **18**, 1082–1090 (2021).

[34] Li, Y. *et al.* Incorporating the image formation process into deep learning improves network performance. *Nature Methods* **19**, 1427–1437 (2022).

[35] Fu, S. *et al.* Field-dependent deep learning enables high-throughput whole-cell 3D super-resolution imaging. *Nature Methods* **20**, 459–468 (2023).

[36] Belthangady, C. & Royer, L. A. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nature Methods* **16**, 1215–1225 (2019).

[37] Van de Ven, G. M., Tuytelaars, T. & Tolias, A. S. Three types of incremental learning. *Nature Machine Intelligence* **4**, 1185–1197 (2022).

[38] Richardson, W. H. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America* **62**, 55 (1972).

[39] Lucy, L. B. An iterative technique for the rectification of observed distributions. *The Astronomical Journal* **79**, 745 (1974).

[40] European Southern Observatory. Wide Field Imager view of the southern spiral NGC 300. `https://www.eso.org/public/images/eso1037a/` (2010).

[41] Olivier, N., Keller, D. & Manley, S. Super-resolution microscopy hub. `https://srm.epfl.ch/srm/dataset/challenge-2D-real/Real_High_Density/index.html` (2013).

[42] Sage, D. *et al.* Quantitative evaluation of software packages for single-molecule localization microscopy. *Nature Methods* **12**, 717–724 (2015).

[43] Reuter, M. *et al.* BRCA2 diffuses as oligomeric clusters with RAD51 and changes mobility after DNA damage in live cells. *Journal of Cell Biology* **207**, 599–613 (2014).

[44] Vašinka, D. Github repository. `https://github.com/VasinkaD/DAMN` (2024).

[45] Li, J., Xue, F. & Blu, T. Fast and accurate three-dimensional point spread function computation for fluorescence microscopy. *Journal of the Optical Society of America A* **34**, 1029 (2017).

[46] Mikš, A. & Novák, J. Point spread function of an optical system with defocus and spherical aberration—analytical formulas. *Applied Optics* **58**, 5823 (2019).

[47] Stallinga, S. & Rieger, B. Accuracy of the Gaussian point spread function model in 2D localization microscopy. *Optics Express* **18**, 24461 (2010).

[48] Maas, A. L., Hannun, A. Y. & Ng, A. Y. Rectifier nonlinearities improve neural network acoustic models. *Proceedings of the 30th International Conference on Machine Learning* **28**, 6 (2013).

[49] Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016). `http://www.deeplearningbook.org`.

[50] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* **15**, 1929–1958 (2014).

[51] Audet, C., Le Digabel, S., Rochon Montplaisir, V. & Tribes, C. Algorithm 1027: NOMAD version 4: Nonlinear optimization with the MADS algorithm. *ACM Transactions on Mathematical Software* **48**, 35:1–35:22 (2022).

[52] Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *Proceedings of the 3rd International Conference on Learning Representations* (2015). arXiv:1412.6980.

[53] Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* **39**, 1–22 (1977).

[54] Vardi, Y. & Lee, D. From image deblurring to optimal investments: Maximum likelihood solutions for positive linear inverse problems. *Journal of the Royal Statistical Society: Series B (Methodological)* **55**, 569–612 (1993).