New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG IOP Institute of Physics

Optical superresolution assisted by multi-mode fiber and neural network

To cite this article: Tom Kuusela and Lauri Kinnunen 2025 *New J. Phys.* **27** 114103

IOP Publishing New J. Phys. 27 (2025) 114103 https://doi.org/10.1088/1367-2630/ae1f34

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

PAPER

Optical superresolution assisted by multi-mode fiber and neural network

Tom Kuusela* and **Lauri Kinnunen**

Department of Physics and Astronomy, University of Turku, 20014 Turku, Finland
* Author to whom any correspondence should be addressed.

E-mail: tom.kuusela@utu.fi

Keywords: diffraction limit, superresolution, neural network

OPEN ACCESS

RECEIVED
26 June 2025

REVISED
14 October 2025

ACCEPTED FOR PUBLICATION
13 November 2025

PUBLISHED
25 November 2025

Abstract

We demonstrate a novel approach for surpassing the diffraction limit in passive optical imaging using a standard step-index multi-mode fiber (MMF) combined with a simple neural network. Unlike previous techniques based on spatial mode demultiplexing and multi-plane light converters, our method relies on the complex speckle pattern generated by mode interference in the MMF. This speckle pattern is highly sensitive to small changes in the input field and is analyzed using a perceptron-type neural network trained to extract parameters such as the separation and intensity ratio of two incoherent point sources. Our experimental results show that the system can resolve beam separations well beyond the classical diffraction limit. The method is flexible and cost-effective, enabling high-resolution and multi-parameter measurements using standard optical components. This work opens new possibilities for passive super-resolution imaging in diverse applications where structured illumination or active modulation is not feasible.

1. Introduction

The angular or spatial resolution of optical imaging systems is a fundamentally limiting factor both in microscopy and in astronomical observations. Although the performance of optical devices has steadily improved, the fundamental physical limits of optical instruments arising from light diffraction were already identified in the late 19th century by Abbe [1] and Rayleigh [2]. According to the Rayleigh criterion, two point sources can be resolved by direct imaging (DI) only if they are separated on the image plane by at least spot size of the point-spread function (PSF) of the imaging system. The diffraction limit can be surpassed in microscopy using certain active imaging techniques that allow control over the properties of the sources or employ specially structured illumination, but these methods are limited to specific applications and are not feasible for general-purpose passive imaging [3–5].

Although camera technology has continued to improve, for example in terms of noise characteristics, no significant advancement has been available for overcoming the diffraction limit in passive imaging, where the spatial intensity distribution is measured—often referred to as DI. The Fisher information, regarded as the standard measure of precision in imaging studies, vanishes in DI as the distance between two sources approaches zero. However, quantum metrology has shown that it is possible to attain maximal Fisher information regardless of the separation between the sources [6]. This can be achieved in practice by applying the so-called spatial mode demultiplexing (SPADE) approach [7, 8].

In practical experiments, spatial mode multiplexing has been implemented using a commercially available multi-plane light converter (MPLC). It decomposes an input field on the HG mode basis and each mode is coupled to a single-mode fiber. By measuring light intensity (or photon counts) from these output fibers, each mode can be determined simultaneously. Typical experimental configurations use 5–10 HG modes [9–12].

In this work, we present a novel approach to surpassing the diffraction limit. Instead of using an MPLC, the image-plane field is coupled into a standard step-index multi-mode fiber (MMF), which supports thousands of modes (not Hermite–Gaussian modes, but the so-called linear LP modes). In

practical situations, many modes with various initial phases are excited in the fiber, and all of them propagate at different velocities. At any point on the fiber output facet, these modes may interfere constructively or destructively, ultimately forming a highly complex intensity speckle pattern. The speckle pattern is not random but is fully determined by the input field—though in a highly complicated manner—and due to interference, it is extremely sensitive to even small changes in the input field. Finally, the speckle pattern is recorded with a digital camera.

Because the speckle pattern is so complex, we do not attempt to model it analytically or measure individual modes separately, as in SPADE. Instead, we interpret the resulting images using the simplest type of neural network—a perceptron—trained to recognize different input fields. Thus, our approach is fundamentally different from the previous ones. The combination of the MMF and the attached neural network can also be understood as an application of a special case of classical reservoir computing, namely extreme machine learning: the fiber acts as a complex nonlinear transformation, and the neural network serves as a thin readout layer that produces the desired output through training [13].

2. Theoretical background

In the following, we briefly describe the theoretical framework of superresolution. A more detailed treatment can be found in [6, 7].

Consider two incoherent point sources of equal brightness, separated by a distance s along the image axis. The system point spread function (PSF) is $h(x)$, normalized such that $\int h(x)\,\mathrm{d}x = 1$. With N detected photons, the image-plane intensity of two partially overlapping spot is given by

$$I(x;s) = \frac{N}{2}\left[h(x-s/2) + h(x+s/2)\right]. \tag{1}$$

We wish to estimate the separation parameter s. From general estimation theory, the Fisher information quantifies how much information about the unknown parameter s is carried by the measurement outcomes [14]. For DI (photon positions in the image plane), the Fisher information is

$$\mathcal{J}_{\mathrm{direct}}(s) = \int \frac{1}{I(x;s)}\left(\frac{\partial I(x;s)}{\partial s}\right)^2 \mathrm{d}x. \tag{2}$$

Furthermore, the Cramér–Rao bound (CRB) sets the lower limit for the variance of any unbiased estimator [15]

$$\mathrm{Var}(\hat{s}) \geqslant \frac{1}{\mathcal{J}_{\mathrm{direct}}(s)}. \tag{3}$$

For small s, a Taylor expansion of the PSF yields

$$h(x \pm s/2) = h(x) \pm \frac{s}{2}h'(x) + \frac{s^2}{8}h''(x) + \mathcal{O}\left(s^3\right), \tag{4}$$

$$I(x;s) = N\left[h(x) + \frac{s^2}{8}h''(x) + O\left(s^4\right)\right], \tag{5}$$

$$\frac{\partial I}{\partial s} = N\left[\frac{s}{4}h''(x) + \mathcal{O}\left(s^3\right)\right]. \tag{6}$$

Therefore,

$$\mathcal{J}_{\mathrm{direct}}(s) \approx N\frac{s^2}{16}\int \frac{h''(x)^2}{h(x)}\mathrm{d}x \propto Ns^2 \quad (s \to 0), \tag{7}$$

and consequently, the CRB diverges as

$$\mathrm{Var}(\hat{s}) \gtrsim \frac{\text{const.}}{Ns^2}. \tag{8}$$

For example, for a Gaussian PSF $h(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right)$ we obtain $\mathrm{Var}(\hat{s}) \gtrsim \frac{8\sigma^4}{Ns^2}$. This loss of information at small separations is known as *Rayleigh's curse*: when only the image-plane intensity is measured, the information about the separation between the sources inevitably vanishes as the spots become closer.

From an information-theoretic perspective, image-plane measurements do not utilize all the information available in the optical field. Instead of measuring in the image plane, one may project onto an

orthogonal spatial-mode basis (e.g. Hermite–Gaussian modes) and count the photons in each mode. Formally, if the state of a single photon from a displaced PSF is $|\psi_\pm\rangle$, the mixed states is

$$\rho(s) = \frac{1}{2}\left(|\psi_+\rangle\langle\psi_+| + |\psi_-\rangle\langle\psi_-|\right). \tag{9}$$

Projecting onto a set of orthogonal modes $\{|\phi_n\rangle\}$, the detection probabilities are

$$p_n(s) = \frac{1}{2}\left(|\langle\phi_n|\psi_+\rangle|^2 + |\langle\phi_n|\psi_-\rangle|^2\right). \tag{10}$$

In this formalism, the Fisher information is expressed as

$$\mathcal{J}_{\text{modes}}(s) = \sum_n \frac{1}{p_n(s)}\left(\frac{\partial p_n(s)}{\partial s}\right)^2. \tag{11}$$

For small s, the Hermite–Gaussian modes are linearly sensitive to displacement, and the Fisher information does not vanish as $s \to 0$. For a Gaussian PSF of width σ, the quantum Fisher information per photon is

$$\mathcal{J}_{\text{Q}} = \frac{1}{4\sigma^2}. \tag{12}$$

This bound can be attained by spatial-mode demultiplexing (HG projection). Hence, with N photons, the Fisher information and the corresponding Cramér–Rao bound are

$$\mathcal{J}_{\text{modes}} \approx \frac{N}{4\sigma^2}, \qquad \text{Var}(\hat{s}) \geqslant \frac{4\sigma^2}{N}. \tag{13}$$

Crucially, this bound is independent of s, in stark contrast to the DI case. Therefore, spatial-mode multiplexing completely avoids Rayleigh's curse.

If the PSF of the optical system is Gaussian, the optimal spatial mode basis for demultiplexing is the Hermite–Gaussian (HG) mode set. By resolving these modes–i.e. performing spatial-mode demultiplexing–the separation between the sources can be accurately estimated. In previous studies, this has typically been implemented using a MPLC. In the present work, an new alternative implementation is employed: an optical MMF is used to decompose the incoming field into linearly polarized (LP) modes (see appendix). Since both HG and LP modes constitute orthogonal bases, the LP modes can be represented as superpositions of HG modes. Consequently, the superresolution method described above can be equivalently realized using LP modes, providing a practical route to implement spatial-mode multiplexing in fiber-based systems.

3. Experimental setup and data analysis

3.1. Optical design

The experimental setup is shown in figure 1. The two point-like sources are simulated using two laser beams that are combined at a beam splitter. The beams have a Gaussian shape, and their waist w_0 is approximately 500 μm (the corresponding FWHM $\approx 830\,\mu$m). According to the Rayleigh criterion, these point-sources are hard to resolve when their transverse separation is comparable or smaller than FWHM. The beam splitter is mounted on a horizontally movable stage (adjustable by a micrometer and a piezo element), which allows the distance between the two beams to be adjusted in a controlled manner along the horizontal axis. In addition, the second beam can be moved vertically using a two-mirror periscope, where the position of the second mirror is adjustable by a micrometer in the vertical direction. The resolution of both the horizontal and vertical displacements is one micrometer.

Since the speckle pattern emerging from the fiber is sensitive to the input polarization, a fixed polarizer (PO3) is placed in front of the fiber. The polarizers PO1 and PO2 allow independent adjustment of the intensity of each beam. A step-index type MMF (Thorlabs M43LO2, the core diameter 105 μm and NA 0.22) is used, as it provides the clearest speckle patterns, which improves the system's sensitivity to distinguish small distances between the beams. It can be shown that the number of optical modes in the fiber is $M \approx (4/\pi^2)V^2$, where $V = 2\pi aNA/\lambda$, a is the core radius and λ is the wavelength [16]. We can estimate that, at the laser wavelength of 633 nm, this fiber supports about 5300 different modes.

No lens is used at the output of the fiber; instead, the speckle pattern is directly projected onto the surface of the camera chip. The distance between the end of the fiber and the camera chip is 10 mm.

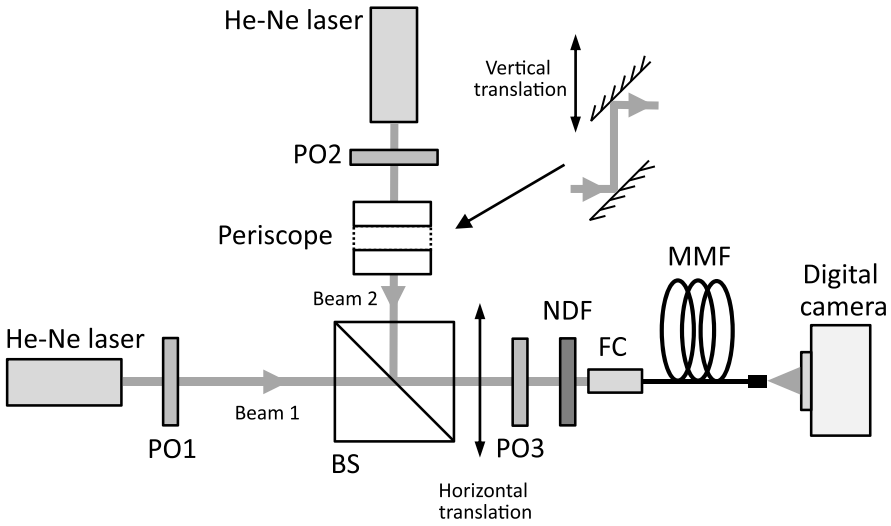

Figure 1. Experimental setup of the superresolution measurement system. The light beams from two He-Ne lasers (multi-mode, non-polarized) are combined at a beam splitter (BS), which is mounted on a horizontally movable stage. Beam 2 passes through a periscope consisting of two mirrors before reaching the beam splitter. The second mirror of the periscope can be moved vertically. The combined beams are guided through a polarizer (PO3), a neutral density filter (NDF), and a fiber collimator (FC) into a multi-mode optical fiber (MMF), the other end of which is connected to a digital camera capturing the speckle patterns.

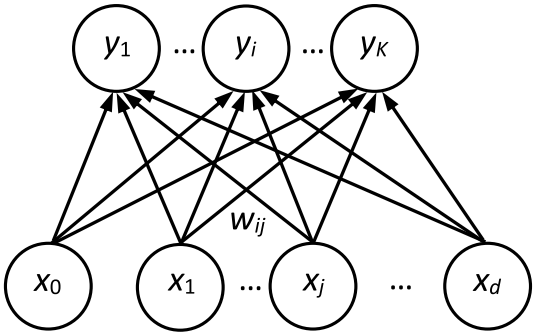

Figure 2. K parallel perceptrons with the inputs x_j, $j = 1, \ldots, d$, and outputs y_i, $i = 1, \ldots, K$.

The digital camera is equipped with a CMOS sensor of 1280 × 1024 pixels, with an active area of 6.784 mm × 5.427 mm. The pixel intensity measurements are 8-bit. Before any data analysis, we crop the images to the size of 1024 × 1024 pixels. In some comparative experiments, we have also employed the DI method, in which the two-beam image was captured directly by a camera without any fiber, while the image analysis was carried out in the same manner as in the fiber-based case

3.2. Image analysis by a neural network

Data analysis was carried out using the simplest type of neural network, namely the perceptron, which has multiple inputs and a single output [17]. Although more advanced neural networks such as convolutional neural networks (CNNs) are also widely used in image analysis, their multilayered structure requires considerably more effort in network design and in the selection of optimal parameters. CNNs are also more prone to overfitting. Since speckle patterns are structurally so simple, a parameter-free perceptron is sufficient for their modeling and does not require any specialized expertise.

In this work, we have used several perceptrons in parallel (see figure 2). It has inputs x_j, $j = 1, \ldots, d$ and outputs y_i, $i = 1, \ldots, K$. The K outputs are

$$y_i = \sum_{j=1}^{d} w_{ij} x_j + w_{i0}, \tag{14}$$

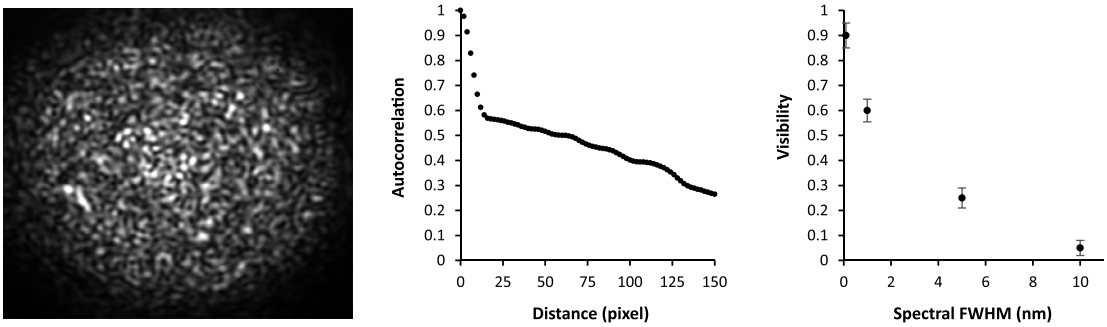

Figure 3. Left panel: speckle pattern measured at the end of the multi-mode optical fiber; middle panel: autocorrelation of the pixel data as a function of the pixel distance; right panel: visibility of speckles as a function of spectral FWHM of the light source.

where w_{ij} is the $K \times (d+1)$ weight matrix. To train the perceptron, i.e. to find the optimal weights, we minimize the error function

$$E(w) = \frac{1}{2}\sum_t \left(r^{(t)} - y^{(t)}(w)\right)^2, \tag{15}$$

where $r^{(t)}$ are the target values indexed by t. The target values used in our system are the horizontal and vertical distance of the beams and their intensity ratio. The optimal weights are obtained using the gradient *descent method*, in which each iteration step is taken in the direction of the negative gradient of E. In other words, the update of the weight w_{ij} is given by

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} = \eta \sum_t \left(r_i^{(t)} - y_i^{(t)}\right) x_j, \tag{16}$$

where η is the learning factor.

We used fixed value $\eta = 10^{-4}$ in our analysis. Before the training procedure, the weights w are initialized to a small random number. The number of iterations ranges from 10000 to 70000, and iterations are terminated when the relative error function $E(w)/E(w_{\text{initial}})$ is less than 10^{-4}. The neural network was trained with 10–455 images depending on the experiment, and in most cases subsequently tested with five times as many images. The average maximum deviation from the target values was then calculated and reported for each experiment. The training phase is practically very fast, as it takes only a few seconds on a standard desktop computer.

3.3. Data pre-processing

In practice, it is not necessary to use all over one million camera pixels to train the neural network to distinguish between the two beams. A visual inspection of the speckle pattern reveals that the speckles have a certain average size (see figure 3, left panel). This property can be quantified by calculating the autocorrelation function from the pixel data:

$$C(\delta) = \frac{\langle X_{ij} X_{i+\delta, j+\delta} \rangle}{\langle X_{ij} \rangle^2}, \tag{17}$$

where X_{ij} is the pixel data, δ is the pixel distance, and $\langle . \rangle$ denotes averaging over all pixels. When examining the autocorrelation as a function of distance (see figure 3, right panel), we observe that it initially drops rapidly over a distance of 8–10 pixels but changes very little beyond that. We interpret this rapid drop as corresponding to the average size of the speckle features.

The correlation length of a speckle pattern is determined by the numerical aperture (NA) of the fiber. In principle, the characteristic speckle angular grain size scales as λ/NA, where λ is the wavelength [18, 19]. Although the pattern is complex, it possesses a well-defined statistical grain size that is nearly uniform throughout the entire speckle field. This arises from the fact that the interference of all guided modes can be statistically modeled as a sum of random vectors, resulting in what is known as a fully developed speckle field with a universal distribution. The modal structure of the fiber sets the scale: since the fiber supports only a finite number of modes and their angular spread is limited by the NA, the speckle grain size cannot be smaller than this fundamental resolution limit.

In our analysis, the images were averaged over 8×8 pixel rectangles. The amount of data was further reduced by cropping a 32×32 pixel region from the center of the image. As a result, we obtain a 1024-pixel image data set, which serves as the input vector x_j for the neural network.

The visibility of speckle as a function of the source spectral FWHM is shown in figure 5(right panel). The results were obtained using a He–Ne laser, a semiconductor laser, and LED source combined with interference filters. For He–Ne lasers, the visibility is close to unity, but it decreases rather rapidly as the spectral width increases. In practice, speckle formation is hardly observable when the spectral width exceeds about 5 nm, in which case the optical fiber provides no benefit.

The pixel intensities are 8-bit, meaning they can take on values between 0 and 255. The intensities are scaled by first subtracting 128 and then dividing the result by 128. The initial value of all weights w_{ij} is set to a random number between -0.01 and $+0.01$.

4. Results

In our first test, the second beam was shifted only horizontally in steps of 10 μm over a range from zero to 500 μm, resulting in a training dataset consisting of 50 images. The intensity of both beams was kept constant and equal. Figure 4 shows a typical result when the neural network trained in this way was used to model corresponding test data. The result shows that the predicted distances closely match the true distance values d, with an average maximum deviation of only about ± 1 μm. A shift of 500 μm is on the order of the beam waist, and therefore this range is not particularly challenging. In the following results, we focus on a smaller displacement range of 0–90 μm corresponding to the ratios $d/w_0 < 0.2$.

For comparison, figure 4 also shows the results obtained without the MMF, i.e. using the DI method together with a neural network (open dots). The images were preprocessed in the same way as when using the fiber. In this case, the maximum deviation is even ± 40 μm, and therefore it is not practically possible to achieve more accurate distance measurements with this approach. In the case of DI, the neural network also converges very poorly when there are many images. If the required accuracy is not higher than this and the distance range is not much smaller than the spot size, the DI method combined with a neural network is a viable alternative. In the following presentation of results, however, we focus only on the use of the MMF.

If we train the neural network using images measured at 10 μm intervals over the distance range 0–90 μm, the network models the distances excellently, as shown in figure 5(left panel, solid dots). However, this is true only when the actual distances match exactly (or very closely) the corresponding values in the training set, meaning that the model is not practically useful as such. On the other hand, if the training data consists of images measured at 1 μm intervals, the resulting model performs satisfactorily (figure 5, right panel). In that case, the average maximum deviation is about ± 1 μm, which is clearly larger than the training data accuracy of 0.1 μm.

A model trained with larger distance steps, which is very easy to train, may still be useful if the required accuracy is only 10 μm. Figure 6(left panel) shows the predicted distances of such a model tested with 1 μm resolution data. Although the actual and predicted distances do not match very well, and the correspondence is not even monotonic, the model still predicts 98% of the values correctly into 10 μm-wide bins (and the remaining 2% fall into one of the neighboring bins). Depending on the application, this may be a sufficient level of performance. A significantly smaller training set is a clear advantage when modeling must also account for additional variables, as this will always increase the total data size substantially.

The predictability of distance shown in figure 6(left panel) can be utilized when higher distance resolution is desired than what is achievable with a single model. A separate neural network can be trained for each 10 μm distance interval, with improved accuracy within that range. Once the correct bin is identified, the corresponding higher-resolution model is selected. The results obtained using this approach are shown in figure 6(right panel): the model error in the terms of the maximal deviation is only ± 0.3 μm. It should be noted that this strategy does not increase the size of the training set at all.

In reality, we cannot assume that both light sources have the same intensity. For this reason, the model must include intensity, or the intensity ratio, as an additional variable. To address this, we created a training dataset consisting of 455 separate measurements, where the beam separation was varied with a 1 μm resolution over the range 0–90 μm for five different intensity ratios I_1/I_2 (1.0, 0.8, 0.6, 0.4 and 0.2). In each case, the total intensity $I_1 + I_2$ was kept constant.

Figure 7 shows the modeling results for distance (left panel) and intensity ratio (right panel) as a function of the training data index. From the results, it is evident that accurate distance prediction is no

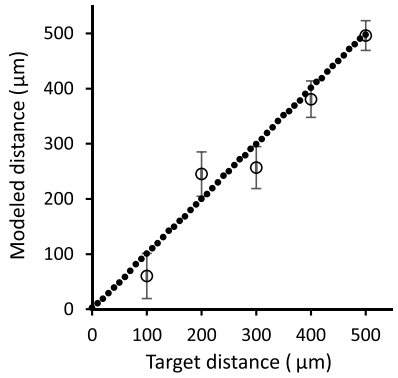

Figure 4. The modeled output distance of the neural network as a function of the target distance d of the beams with direct imaging and the neural network (open dots) and a multi-mode fiber and neural network (solid dots). The maximum distance of 500 μm corresponds to the ratio $d/w_0 = 1$.

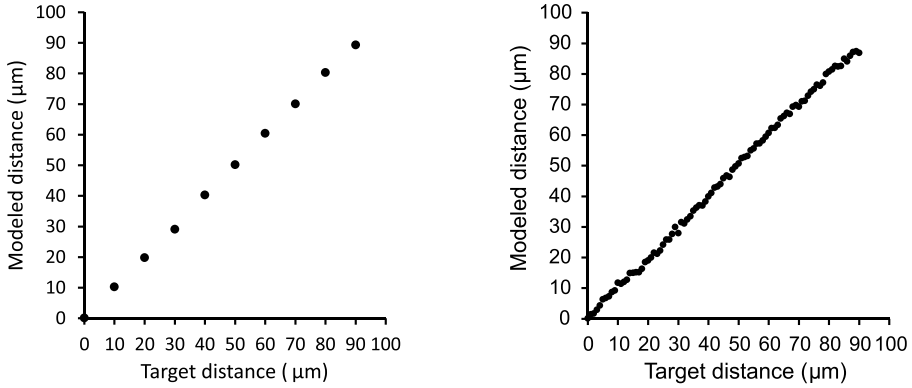

Figure 5. The modeled output distance of the neural network as a function of the target distance d of the beams. Left panel: the training set consisting of 10 distances with the step of 10 μm; right panel: the training set consisting of 100 distances with the step of 1 μm.

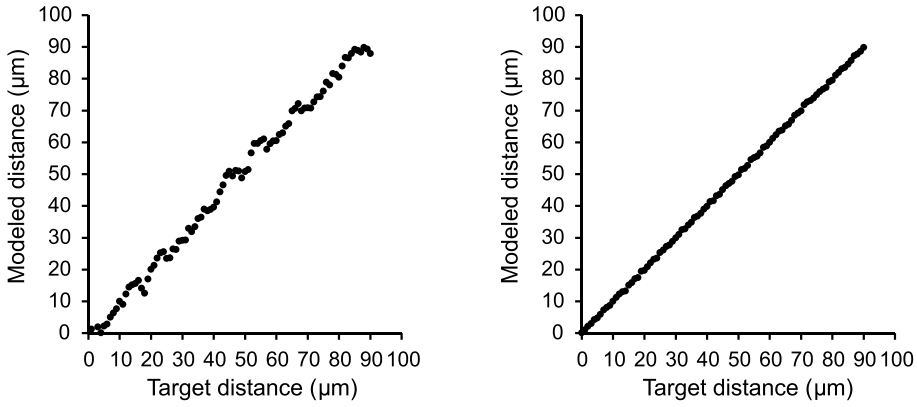

Figure 6. Left panel: the training set as in the figure 5(left panel) but the test set with 1 μm steps; right panel: the combination of 10 separately trained neural networks, each over the range of 10 μm.

longer feasible. In contrast, the intensity ratios can be predicted quite reliably, and training the neural network for this purpose is straightforward. We found that interpolation of the intensity ratio between the values in the training set also works well.

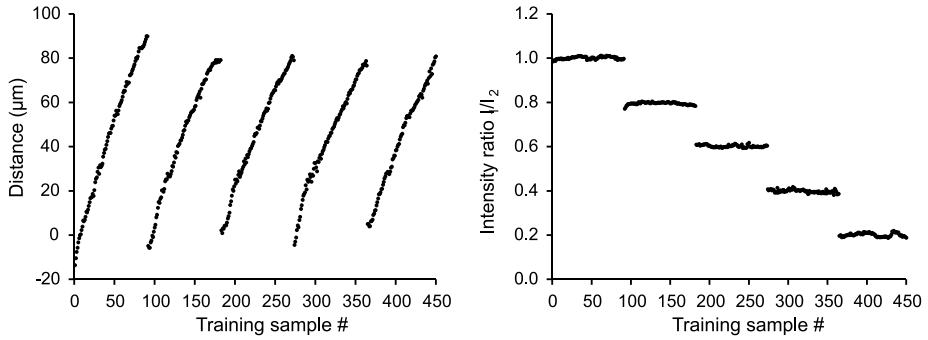

Figure 7. Modeled beam distance (left panel) and intensity ratio I_1/I_2 as a function of the training sample index.

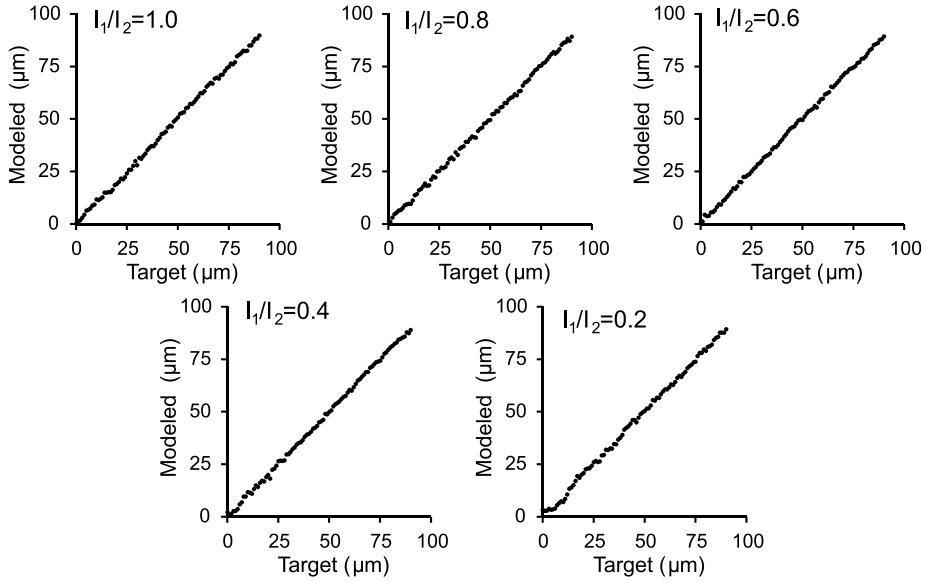

Figure 8. The separately modeled output distance as a function of the target distance d of the beams with five different intensity ratios I_1/I_2.

Here, we can apply the same approach as before: once the beam intensity ratio is identified and assigned to the correct bin, we use a separate neural network trained specifically for that data. The results are shown in figure 8. For each intensity ratio, the distances can be predicted accurately within the maximal deviation of $\pm 1\,\mu$m—except for the smallest ratio, $I_1/I_2 = 0.2$, at beam separations below $10\,\mu$m. It is natural that when the intensity ratio is very small, modeling the separation of two beams becomes more challenging. Even in this case, prediction accuracy can be further improved by using segmented distance modeling, as described earlier.

Finally, we have tested the performance of our method in a situation where the beams differ not only in the horizontal direction but in both horizontal and vertical directions. Figure 9 shows the modeled distances in two dimensions. The training dataset consists of 100 images in which one of the beams has been shifted in steps of 10 μm over a total range of 90 μm in both directions. The figure shows that the modeling (open dots) performs reasonably well. As a test set, we used cases where the second beam was also shifted in two directions by 10 μm steps, but in such a way that the beam positions fall halfway between the training data positions. The predictions match quite well (see figure 9, black dots) with the maximal deviation of ± 2 μm, and it is therefore possible to use the trained model at least for classifying the distance into two-dimensional bins, or even for interpolation within the bins. Our observation is that the two-dimensional training data actually performs even better in this respect than the one-dimensional case, presumably due to the neural network's ability to more effectively utilize the information contained in the neighboring points surrounding each data point.

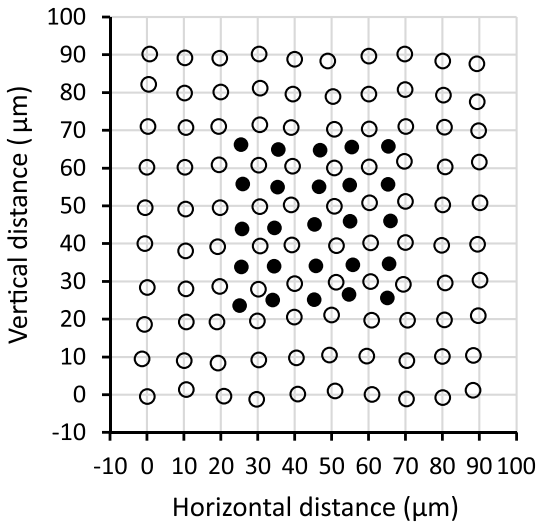

Figure 9. The horizontal and vertical distances of the beams modeled by the trained neural network (open dots). The black dots represent some examples of the model predictions of the test data.

5. Discussion

We have shown above that a combination of a MMF and the simplest perceptron neural network enables easy measurement of the distance between two point-like sources with a resolution that surpasses the diffraction limit by at least three orders of magnitude–in practice, well below 1 μm for a beam of the waist $w_0 = 500\,\mu$m. In fact, we have successfully tested even smaller distances, but in those cases the limiting factor is merely the stability of the setup rather than the method itself.

So far, the best outcomes have been achieved using SPADE and MPLC approaches, which currently represent the most promising methods for super-resolution imaging, as reported in reference 12. In that study, the resolution in the high-flux regime for a 1 mm beam size was as small as 20 nm, although the estimated separation deviated on average by approximately 1 μm. The resolution of our proposed method does not reach this level, but its accuracy may be better than 1 μm. On the other hand, it is typical in such superresolution experiments that the main sources of deviations are not the method itself but various errors in calibration, non-identical beam shapes, and electrical noise, as is also partly the case in our approach.

Our approach is highly flexible: measurements can be performed at any wavelength range supported by the chosen MMF (and the digital camera), whereas currently available MPLC solutions mainly operate at telecommunications wavelengths. The method also naturally supports multi-parameter measurements, such as both lateral beam separation and intensity ratio—features that have received a little attention or testing in the SPADE framework. Furthermore, our method is applicable to any form of point-source function, as long as it is represented in the training data.

The method we propose has a few critical characteristics. First, the speckle pattern is highly sensitive to the optical wavelength and its spectral distribution. If the spectrum is broad, the pattern becomes blurred, making it more difficult to train the neural network to achieve high resolution. We have conducted experiments with multi-mode semiconductor lasers with the spectral FWHM of about 1 nm, and those were able to reach performance close to the results presented here. In real-world applications, it will likely be necessary to filter the input beams using a narrow-band interference filter, because if the source spectral FWHM exceeds 5 nm, the multimode fiber no longer provides any particular advantage and the resolution is essentially the same as in DI. On the other hand, most experimental demonstrations of superresolution imaging reported in the literature also rely on very narrow-band telecom laser sources. If the source is inherently nearly monochromatic, the neural network could even be trained to recognize the optical wavelength [20].

Second, the polarization of the light coupled into the fiber must be fixed. In practice, this is not a severe limitation, as it can always be ensured using a polarizer and/or half-wave plates. Alternatively, the neural network can also be trained to recognize the polarization state of the source if desired.

Third, the MMF is sensitive to all environmental changes, especially mechanical disturbances. In our experiments, we used a commercially available 2 m fiber, which is unnecessarily long and thus vulnerable

to bending and external perturbations. In practice, the fiber length could be reduced to just a few centimeters, making it relatively easy to stabilize against environmental noise.

Training the system always requires generating high-quality training data, which can be laborious, especially when modeling several parameters. However, as we have shown, it is not always necessary to train a single model to determine all parameters. Instead, the influence of different parameters can be distributed among several neural networks, and even the different value ranges of a single parameter can be handled by separate networks. This significantly reduces the required amount of training data. Based on our preliminary tests, for a realistic three-parameter system (beam separation in two dimensions and intensity ratio), a few hundred training images would likely be sufficient. As our neural network is the simplest one, training task is not demanding. It should also be noted that we have not, in any way, directly utilized the obvious two-dimensional structural features of the speckle patterns when forming the one-dimensional input vector for the perceptron. It is reasonable to assume that convolutional neural networks, which are specifically designed for image analysis, could further improve the performance of our method but with the expense of more complex parametrization of the network.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Appendix. Multi-mode fiber

In a homogeneous transparent medium, an optical wave is described by the wave function $U = U(\mathbf{r}, t)$, where $\mathbf{r}$ denotes the position and t the time. This wave function satisfies the wave equation $\nabla^2 U - \frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} = 0$, where $c = c_0/n$, c_0 is the speed of light in vacuum, and n is the refracting index of the medium. Substituting the separable ansatz $U(\mathbf{r},\mathbf{t}) = U(\mathbf{r})\mathrm{e}^{\mathrm{i}\omega t}$, where $U(\mathbf{r})$ is the complex amplitude of the wave and ω the angular frequency, into the wave equation leads to the Helmholtz equation

$$\nabla^2 U(\mathbf{r}) + n^2 k_0^2 U(\mathbf{r}) = 0, \tag{A.1}$$

where $k_0 = \omega/c_0 = 2\pi/\lambda_0$, k is the wavenumber, and λ_0 is the wavelength.

An optical fiber is a cylindrical dielectric waveguide made of silica glass, consisting of a central core surrounded by an outer cladding. In general, the refractive index of the fiber is a function $n(r)$ of the radial coordinate r. In particular, for a step-index fiber one has $n(r) = n_1$ in the core ($r < a$) and $n(r) = n_2$ in the cladding ($r > a$), where n_i are constants. For the case of an optical fiber, it is natural to express equation (2) in cylindrical coordinates, yielding for each monochromatic field component

$$\frac{\partial^2 U}{\partial r^2} + \frac{1}{r}\frac{\partial U}{\partial r} + \frac{1}{r^2}\frac{\partial^2 U}{\partial \phi^2} + \frac{\partial^2 U}{\partial z^2} + n^2 k_0^2 U = 0, \tag{A.2}$$

where $U = U(r, \phi, z)$. Assuming light propagates in the z-direction with propagation constant β, and taking into account the circular symmetry of the waveguide (the field must be invariant under the transformation $\phi \to \phi + 2\pi$), we obtain solutions of the form $U(r, \phi, z) = u(r)\mathrm{e}^{-\mathrm{i}l\phi}\mathrm{e}^{-\mathrm{i}\beta z}$, where l is an integer. Substituting this into equation (2) gives the equation for the radial part

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \left(n^2(r)k_0^2 - \beta^2 - \frac{l^2}{r^2}\right)u = 0, \quad l = 0, \pm 1, \pm 2, \ldots \tag{A.3}$$

The solutions of this equation are Bessel functions for $r < a$ and modified Bessel functions for $r > a$. In the paraxial approximation one may assume that the longitudinal electric and magnetic field components are much weaker than the transverse components. Requiring the continuity of the field and its derivative at $r = a$ leads to the condition that for each l, the propagation constant β can only take discrete values. These solutions, called *LP (linear polarization) modes*, can be indexed as $LP_{lm}(r)$, and the electric field written in the form $E_{lm}(r, \phi, z) = LP_{lm}(r)\mathrm{e}^{-\mathrm{i}(l\phi + \beta_{lm} z)}$. Any input field at the fiber entrance can then be expanded into the fiber modes as a linear superposition $E_0(r, \phi, 0) = \sum_{l,m} a_{lm} LP_{lm}(r, \phi, 0)$ with complex coefficients α_{lm}. The modal solutions at the output face of the fiber are then readily calculated. In practice, typically hundreds of modes with different initial phases are excited in the fiber. Each mode propagates with a different velocity, and at the fiber output they interfere to produce a complex intensity distribution that is highly sensitive to the input field.

ORCID iDs

Tom Kuusela 0000-0002-1436-9138
Lauri Kinnunen 0009-0003-6186-3328

References

[1] Abbe E 1873 Beitrage zür theorie des mikroskops und der mikroskopischen wahrnehmung *Arch. für mikroskopische Anat.* **9** 413
[2] Lord Rayleigh F 1879 Investigations in optics, with special reference to the spectroscope *Philos. Mag. J. Sci.* **8** 261
[3] Moerner W E and Kador L 1989 Optical detection and spectroscopy of single molecules in solid *Phys. Rev. Lett.* **62** 2535
[4] Hell S W and Wichmann J 1994 Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy *Opt. Lett.* **19** 780
[5] Klar T A, Jacobs S, Dyba M, Egner A and Hell S W 2000 Fluorescence microscopy with diffraction resolution barrier broken by stimulated emssion *Proc. Natl Acad. Sci. USA* **97** 8206
[6] Tsang M, Nair R and Lu X-M 2016 Quantum theory of superresolution for two incoherent optical point sources *Phys. Rev. X* **6** 031033
[7] Nair R and Tsang M 2016 Far-field superresolution of thermal electromagnetic sources at the quantum limit *Phys. Rev. Lett.* **117** 190801
[8] Tsang M 2017 Subdiffraction incoherent optical imaging via spatial-mode demultiplexing *New J. Phys.* **19** 023054
[9] Boucher P, Fabre C, Labroille G and Treps N 2020 Spatial optical mode demultiplexing as a practical too for optimal transverse distance estimation *Optica* **7** 1621
[10] Santamaria L, Pallotti D, Siciliani de Cumis M, Dequal D and Lupo C 2023 Spatial-mode demultiplexing for enhanced intensity and distance measurement *Opt. Express* **31** 33930
[11] Greenwood A B, Oulton R and Gerse H 2023 On the impact of realistic point sources in spatial mode demultiplexing super resolution imaging *Quantum Sci. Technol.* **8** 015024
[12] Rouvière C, Barral D, Grateau A, Karuseichyk I, Sorelli G, Walschaers M and Treps N 2024 Ultra-sensitive separation estimation of optical sources *Optica* **11** 166
[13] Mujal P, Martinez-Peña R, Nokkala J, Garcia-Beni J, Giorgi G L, Soriano M C and Zambrini R 2021 Opportunities in quantum reservoir computing and extreme machine learning *Adv. Quantum Technol.* **9** 2100027
[14] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (Academic)
[15] Giovannetti V, Lloyd S and Maccone L 2011 Advances in quantum metrology *Nat. Photon.* **5** 222
[16] Saleh B E A and Teich M C 2007 *Fundamentals of Photonics* 2nd edn (Wiley)
[17] Alpaydim E 2020 *Introduction to Machine Learning* 4th edn (The MIT Press)
[18] Bossy E and Gigan S 2016 Photoacoustics with coherent light *Photoacoustics* **4** 22
[19] Wang Q 2013 Discussion on the fully developed speckle field *Optik* **124** 2948
[20] Redding B, Popoff S M and Cao H 2013 All-fiber spectrometer based on speckle pattern reconstruction *Opt. Express* **21** 6584