Seeing the Invisible: Breaking the Diffraction Limit with Geometry-aware Deep Learning

Benquan Wang[1,2], Eng Aik Chan[1,2], Jin-Kyu So[1,2], Yuhan Peng[2], Ruyi An[3], Yewen Li[3]
Zexiang Shen[1,2], Bo An[3], Giorgio Adamo[1,2], Nikolay I. Zheludev[4]

[1] *Centre for Disruptive Photonic Technologies, TPI, Nanyang Technological University, Singapore 637371, Singapore*
[2] *School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore*
[3] *College of Computing and Data Science, Nanyang Technological University, Singapore 639798, Singapore*
[4] *Optoelectronics Research Centre, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom*

1. Introduction

The opportunity to glimpse the wonders of the tiny world with one's eyes has fascinated researchers for millennia. However, due to the physical phenomenon of diffraction, the optical resolution is restricted to approximately half the wavelength (λ) of light, which impedes the observation of subwavelength objects, typically smaller than 200 nm. This constrains its application in numerous scientific and industrial fields that aim to observe objects beyond the diffraction limit, such as native state coronavirus inspection. Fortunately, deep learning methods have shown remarkable potential in uncovering underlying patterns within data, promising to overcome the diffraction limit by revealing the mapping pattern between diffraction images and their corresponding ground truth images.

In this work, we demonstrate a new universal technique for super-resolution imaging of arbitrary objects without a prior training assumption on the sizes and shapes of the imaging objects, achieved by learning fundamental geometry elements via deep learning analysis.

We report an image resolution of λ/11 which is 5.5 times beyond the classical diffraction limit. Our method is label-free and does not require photoactivation, or bleaching, opening transformative opportunities such as biology imaging, precision manufacturing, semiconductor inspection, and advanced materials characterization.

2. Methodology

2.1 Related work

To surpass the diffraction limit, traditional optical methods, such as scanning near-field optical microscopy[1] offer high resolution but require invasive near-field probes and cannot image internal structures. Meanwhile, fluorescent-based methods like [2, 3] stimulated emission depletion microscopy and single-molecule localization microscopy offer resolutions of tens of nanometers but necessitate invasive fluorescent labeling[4]. These limitations have spurred interest in AI-enhanced solutions.

2.2 Methodology

We first make a big binary sample consisting of random arrangements of lines, circle segments shaped apertures. The white binary region is defined as apertures where illuminated light is fully transmitted, and the black region is defined as regions with no light transmission. The sizes of the apertures vary from λ/10 to λ/2 in widths and diameters. We illuminate the sample with a diffraction-limited beam spot of diameter, D = 1.22λ/NA, and the diffraction pattern collected by an NA=0.9 imaging system at a distance H = 2λ from the sample (Fig. 1).

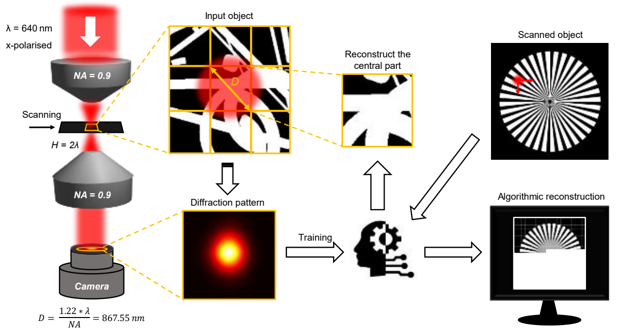

Fig. 1: Working principle: A 2D binary object consists of geometry shape components (line, circle, segment with random size and place), is illuminated by a tightly focused laser beam with a wavelength of 640nm and progressively scans along two orthogonal directions X and Y by fixed steps. A set of diffraction patterns is then recorded at the distance of 2λ from the object along the z-axis with a numerical aperture, NA=0.9 by fully scanning the imaging target. An encoder-decoder convolutional neural network is trained on diffraction patterns from all the scanning small areas of the full imaging target, surrounded by random arbitrary shape objects. The trained neural network aims to reconstruct the image of the central part covered by the diffraction limit spot (without the surroundings) from the unseen diffraction pattern, and the full imaging target is reconstructed progressively image by the image with the neural network results.

2.3 Sample fabrication and experimental setup

We fabricate test samples using a high-precision dual-beam FIB system on a 130nm Au film deposited on glass coverslips. Two representative samples were produced: a calligraphic "Light" pattern featuring complex curved shapes and a Siemens star with radial periodic lines for resolution benchmarking. Sample imaging was conducted using a custom-built microscopy system with a 640nm coherent light source, achieving an effective pixel resolution of 41.7nm. The setup employs a high-NA objective mounted on a piezoelectric stage, with diffraction signals detected by a high-sensitivity sCMOS camera in the far-field. Environmental stability is maintained through comprehensive vibration isolation and acoustic shielding to minimize mechanical drift and noise during high-precision measurements.

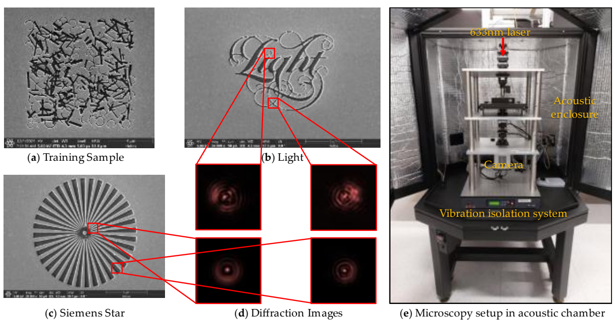

Fig. 2. ***a~d:*** Fabricated samples and the diffraction images; ***e:*** High-precision microscopy for data acquisition.

2.3 Image translation by deep learning

With the collected data, we now define our task as an image-to-image translation problem characterized by a mapping function

$$\boldsymbol{\mathcal{F}_{\phi}}: \boldsymbol{R}^{\boldsymbol{H_1 \times W_1 \times 1}} \rightarrow \boldsymbol{R}^{\boldsymbol{H \times W \times 1}} \quad (1)$$

parameterized by a neural network ϕ . This function $\mathcal{F}$ is designed to transform input diffraction images into outputs that approximate the corresponding ground truth object localization images.

To optimize our image translation mapping function $\mathcal{F}$, we define a fundamental loss function for training:

$$\mathcal{L}_{\mathcal{F}} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{l}(\mathcal{F}(x_i), y_i) \quad (2)$$

As the ground truth y_i uses binary values to indicate object presence at each pixel, we employ the Binary Cross-Entropy (BCE) loss function:

$$\mathrm{l}(\mathcal{F}(x_i), y_i) = -\frac{1}{H \times W}\sum_{h=1}^{H}\sum_{w=1}^{W}\left[y_i^{(h,w)} \log \mathcal{F}(x_i)^{(h,w)} + \left(1 - y_i^{(h,w)}\right)\log\left(1 - \mathcal{F}(x_i)^{(h,w)}\right)\right]$$

where $\mathcal{F}(\mathrm{x_i})^{(\mathrm{h,w})}$ *and* $y_i^{(h,w)}$ denote the predicted and ground truth values at pixel location (h, w), respectively. The neural network parameterizing $\mathcal{F}$ could be optimized using gradient-based techniques, such as SGD and Adam, to minimize $\mathcal{L}_{\mathcal{F}}$ over the training dataset. Once trained, this model is capable of predicting object localization images from new diffraction images.

3. Results

3.1 Imaging ability of arbitrary shape

As a demonstration of imaging arbitrary shape and size objects, we image a "Light" sample with a size of 14λ in length and 7λ in height. The sample is made up of non-uniform calligraphic strokes with a span of aperture width, that ranges from ~ λ to deeply subwavelength λ/64. The deep subwavelength features could be easily distinguished in the image formed by our method. We compare the image reconstructed with a confocal microscope, with equivalent illumination beam spot diameter, captured with the same scanning step and reconstructed from the averaged values from 16 repeated captures. The image captured with a confocal microscope has notably decreased overall contrast and decreased resolution in smaller features.

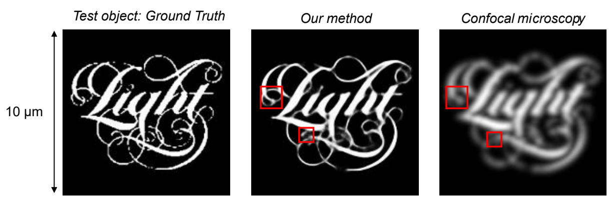

Fig. 3: Test results of "Light" sample demonstrating complex curved features and arbitrary shapes.

3.2 Resolution evaluation

To evaluate the resolution of our method, we image a sample of Siemens star. The Siemens star, composed of 72 spokes radiating from the center, measures 13.6λ in diameter. We evaluate the resolvable threshold as the half-pitch of the stokes gets narrower toward the center of the Siemens star with the visibility metric. The visibility of the image is defined according to the Michelson contrast formula = $(I_{max} - I_{min})/(I_{max} + I_{min})$, where I_{max} is the maximum intensity and I_{min} is the minimum intensity between the pitches of the Siemens stokes. We define the resolvable threshold with the Rayleigh criterion, at which the difference between the maximum intensity and minimum intensity is 20%, which corresponds to a visibility of 0.11. The resolvable threshold for our method is λ/11, which is 5.5 times beyond the classical optical resolution limit.

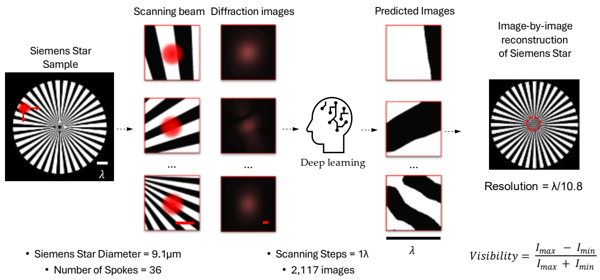

Figure 4: Imaging resolution on Siemens Star

4. Conclusion

We report a far-field label-free optical imaging technique enabled by artificial intelligence, which could work on arbitrary shapes with a resolution of λ/11. This 5.5-fold improvement beyond the classical diffraction limit was achieved without making prior assumptions on target geometry or dimensions. The neural network architecture successfully learned fundamental geometric elements from diffraction patterns, establishing a robust mapping that generalizes across diverse structural configurations. More importantly, our method is scalable to samples with different materials and for demonstration, offering numerous applications including non-invasive biological specimen imaging, high-precision semiconductor inspection, and advanced materials characterization.

References

1. Keilmann, F. and R. Hillenbrand, *Near-field microscopy by elastic light scattering from a tip.* Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 2004. **362**(1817): p. 787-805.
2. Vicidomini, G., P. Bianchini, and A. Diaspro, *STED super-resolved microscopy.* Nature Methods, 2018. **15**(3): p. 173-182.
3. Rust, M.J., M. Bates, and X. Zhuang, *Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM).* Nature Methods, 2006. **3**(10): p. 793-796.
4. Nehme, E., et al., *Deep-STORM: super-resolution single-molecule microscopy by deep learning.* Optica, 2018. **5**(4): p. 458-464.

Acknowledgments

This work was supported by the Singapore National Research Foundation (Grant No. NRF-CRP23-2019- 0006), the Singapore Ministry of Education (Grant No. MOE2016-T3-1-006)