2023 20th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). Mexico City, Mexico. October 25-27 , 2023.

2023 20th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE) | 979-8-3503-0676-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/CCE60043.2023.10332844

Res-U2Net: Augmenting 2D/3D Image Reconstruction through Untrained Deep Learning Models for Phase Retrieval Enhancement

C. Osorio Quero[1], J.A. Cisneros Martinez[2], and R. Ramos-Garcia[2]
[1] *Electronics Department*, [2] *Optics Department*
Instituto Nacional de Astrofísica, Óptica y Electrónica. (INAOE)
Luis Enrique Erro 1, 72840 Tonantzintla, Puebla, Mexico
email:{caoq,cisneros,rgarcia}@inaoep.mx

***Abstract*—In the field of computational imaging (CI), supervised training methods have long been the dominant approach for neural networks in optics. These methods heavily rely on large amounts of labeled data to adjust network weights and biases effectively. However, obtaining a substantial number of ground-truth images for training poses significant challenges in real-world scenarios like phase retrieval or 2D/3D imaging applications. To overcome this limitation, we propose an innovative approach that merges principles from physics with deep neural networks. Our objective is to reduce the dependency on extensive labeled data by incorporating a comprehensive physical model that accurately represents the image formation process. This unique approach allows us to achieve 3D imaging through phase retrieval, utilizing techniques such as Gerchberg-Saxton (GS) and Fourier-Rytov (FR), in combination with deep learning architectures to extract intricate information from the phase. Consequently, this information enables us to detect changes in an object's surface and generate a mesh representation of its 2D/3D structure. In our proposal, we introduce Res-U2Net, a novel untrained neural network designed to estimate the 3D structure of objects. By adopting a unified method for object analysis, this approach presents a new paradigm for neural network design, seamlessly integrating physical models. Furthermore, this framework can be extended to address a wide range of other computational imaging challenges**

***Index Terms*—2D/3D Phase retrieval, Deep Learning, Untrained, Res-U2Net, Gerchberg-Saxton (GS), Fourier-Rytov (FR),NR-IAQ BRISQUE,NIQE.**

I. Introduction

Recently, the field of computational imaging (CI) has made remarkable strides, thanks to the integration of deep learning (DL) techniques [1]. DL has shown great promise in addressing the challenging inverse problems encountered in CI applications [2]. Pioneering studies have demonstrated the efficacy of DL in various CI domains, such as optical tomography, 3D reconstruction, phase retrieval, computational ghost imaging, digital holography, imaging through scattering media, fluorescence lifetime imaging, unwrapping, and fringe analysis [3]–[10]. Typically, DL-based artificial neural networks in CI rely on extensive labeled datasets for training to optimize their weight and bias parameters [11]. This process enables the network to learn a universal function that maps data from the object space to the image space. However, this training can be time-consuming, while the reconstruction itself is often rapid [12]. Acquiring a diverse and substantial training dataset becomes challenging in real-world scenarios where novel objects or scenarios are encountered, limiting the network's generalization capabilities [13].

Recent advances in imaging applications have shown promise with unsupervised learning techniques, particularly utilizing untrained networks [14]. These approaches leverage the inherent structure of neural networks without requiring training data, leading to impressive results. Examples include the deep image prior [14] and deep decoder [15], which exploit the network structure as a prior for image statistics, even without prior training. This involves using a deep network with randomly initialized weights as an image generator, which is iteratively updated using a loss function that compares the generated image with the input data.

While these methods have been successful in simulated image denoising, deblurring, phase retrieval, and super-resolution tasks [13], challenges arise in computational imaging where acquired measurements do not directly resemble the reconstructed image. Instead, a forward model based on the underlying physics of the image formation governs the relationship between the scene and the measurements. For instance, in phase retrieval, the model constructs the phase of a sample using known amplitude distribution pairs at both the sample and measuring planes, employing methods such as Gerchberg-Saxton (GS) [16] and Fourier-Rytov (FR) [17].To address these challenges and enhance the capabilities of deep neural networks in computational imaging, a combination of the UNet architecture and physics-informed techniques can be employed [13], [18]. The UNet's encoder-decoder structure enables effective capture of both local and global image features. By incorporating physics-based constraints and priors into the network's design, a physics-enhanced deep neural network can more accurately model the image formation process, leading to improved quality and fidelity of reconstructed images.

In this study, we propose a method for 2D/3D image reconstruction using phase retrieval with untrained deep learning and mesh estimation over the phase image [19]. We explore various physics-based methods, including Gerchberg-Saxton (GS) and Fourier-Rytov (FR), implemented using the Untrained Deep Learning UNet [20], and we propose a new architecture type called Res-U2Net.

II. Related Works

A. Optimizing Nonlinear Phase Retrieval

In the initial phases of development, phase retrieval techniques primarily leaned on the iterative alternating minimization (AM) approach as detailed in [21]. These techniques involved a cyclic process of updating the estimated image $\tilde{x} \in \mathbb{C}^{NxN}$ by transitioning between the spatial and Fourier domains. Nevertheless, the AM-based algorithms came with certain limitations. They often faced issues of stagnation and demonstrated gradual convergence rates, sometimes demanding more than 1000 iterations to arrive at a solution [22].

B. Deep Learning Phase Retrieval

Extensive research explores deep learning methods for phase retrieval (PR), offering faster non-iterative solutions compared to traditional optimization. These methods excel in PR from a single Fourier intensity measurement, previously deemed impossible. Categorizations of deep learning-based PR are based on physics integration.

- The first category involves feedforward networks estimating target images from intensity measurements. While some work for simple images, efficacy with complex ones is uncertain. Generative adversarial networks and ResNet structures in this category struggle with intricate datasets [23].
- The second category enhances reconstructed image quality by incorporating physics. Methods include spectral initialization or cascaded networks with multiple multilayer perceptrons (MLPs), which can produce noisy images or need large networks [24].
- The third category employs UNet, a popular architecture, on iterative methods. It capitalizes on the iterative nature of PR and problem physics. UNet serves as a neural network mapping intensity to phase. Trained on intensity-phase pairs, it refines phase estimates iteratively for new data until convergence [13].

III. Physics model

The process of phase retrieval aims to reconstruct a desired object Eq. (1), denoted as x^*, using information about the measured intensities y and the characteristics of the imaging system represented by the non-linear operator A [25]. The objective is to recover an estimate of the signal, $\hat{x}$, from the measurements y. This problem can be formulated as a non-convex optimization task: $\hat{x} = \text{argmin}_x \, \|y - |Ax|\|_2^2$ [26].

$$y = |Ax^*|^2 \tag{1}$$

The matrix A is closely associated with the physical model of the imaging system. To address this challenge, the Fourier transform can be employed by defining A as the 2D Fourier Transform matrix. Several methods utilize the Fourier Transform for phase retrieval, including the Gerchberg-Saxton (GS) and Fourier-Rytov (FR) methods. However, it should be noted that accurately reconstructing the original signal is difficult due to the lack of phase information. The absence of phase information results in an inverse problem with infinitely many potential solutions that can produce the same amplitudes in the Fourier domain.

A. Gerchberg-Saxton (GS)

The Gerchberg-Saxton (GS) algorithm [16] is an iterative method for phase retrieval from magnitude data. Operating in the Fourier domain, it begins with random phases assigned to the Fourier transform, then iterates by updating phases while keeping magnitudes fixed. This process alternates between Fourier and spatial domains until convergence. The GS algorithm assumes phase can be determined from magnitude data, under constraints including non-negativity, often achieved by using absolute values or squared module of the Fourier transform.

B. Fourier-Rytov (FR)

The Fourier-Rytov (FR) algorithm [27] is a method employed for phase retrieval, which involves the recovery of phase information from intensity measurements of a wavefield with complex values. By relying on the principle that the Fourier transform of an object's autocorrelation function holds the desired phase information. Through an iterative process, the algorithm estimates the phase of the wavefield by utilizing its measured intensity and the Fourier transform of its autocorrelation.

C. Ill-Posed Problem of Phase Retrieval

In the realm of image reconstruction and similar inverse problems, an ill-posed scenario emerges when minor alterations in measured data lead to substantial uncertainties in the estimated solution [28]. Such situations arise when crucial information is absent, as seen in image reconstruction lacking phase details. This absence of phase information renders the problem ill-posed due to countless potential solutions that match given amplitude measurements. Traditional methods and iterative algorithms like Gerchberg-Saxton and Fourier-Rytov face challenges in achieving a definitive and accurate solution due to this ill-posed nature, exacerbated by measurement noise. To mitigate this, leveraging UNet-based techniques is proposed [18]. UNet is a deep learning architecture that can learn complex mappings between input and output data improved phase estimation and convergence for algorithms like GS and FR. Integrating classical algorithms with UNet offers a promising approach to address the difficulties posed by ill-posed problems, enhancing reconstruction robustness.

IV. 2D/3D Phase Retrieval

In the initial stage of phase retrieval (refer to Fig. 1), an untrained approach is employed. This involves the application of a physics model H, as defined in Section III-A (GS/FR), to evaluate the input image and generate a diffraction model I. Subsequently, this diffraction model serves as input for an untrained neural network in the proposed model. We assess the performance of two distinct neural networks: UNet and Res-U2Net, which have been proposed for this purpose (evaluation the performance the different UNet networks is document in the reference [29]). The neural network aims to estimate the phase information $\tilde{\theta}$ by comparing the diffraction model estimate I derived from the estimated phase $\tilde{\theta}$ with the original diffraction model. Training the neural network involves minimizing the mean square error (MSE) between the estimated diffraction models I obtained from the input image and the diffraction models formed from the phase information I^*. Through this iterative process, the estimated phase undergoes refinement, enhancing the performance of the neural network in phase retrieval tasks.

A. Neural Network Res-U2Net

The Res-U2Net architecture is an enhanced version of the UNet model [18] designed for image segmentation tasks. It incorporates residual connections at multiple levels to improve information flow and gradient propagation during training. The architecture consists of downsampling and upsampling pathways similar to UNet. In the downsampling pathway, convolutional layers with batch normalization and ReLU activation are followed by max-pooling layers. The key improvement lies in the upsampling pathway, where transposed convolutional layers are used to upsample the feature maps. Residual connections are established between corresponding layers in the downsampling and upsampling pathways to preserve finer details. These connections aid in mitigating the vanishing gradient problem and enable better feature propagation. The final layer employs a convolutional operation with sigmoid activation to produce the segmentation output. Res-U2Net demonstrates improved performance in various image segmentation tasks compared to the original UNet (refer to Figure 2).

B. 3D Phase Reconstruction

We employed the Unified Shape-From-Shading Model (USFSM) for the purpose of conducting 3D reconstruction on the estimated image obtained via phase retrieval. The USFSM approach enables the creation of three-dimensional representations by analyzing the spatial intensity variations found in the recovered two-dimensional image [30]. In order to extract depth information from the phase retrieval image, which corresponds to the surface points of the scene, we utilized the fast sweeping method. This particular method makes use of the Lax–Friedrichs Hamiltonian technique [31] to solve for the surface. It employs an iterative sweeping strategy based on the fast sweeping scheme described in [19] (Refer to Fig. 3 for visual representation).

V. Numerical Results

Our research approach based on neural network and physics models to tackle 2D/3D phase retrieval challenges. We employ state-of-the-art techniques like UNet and Res-U2Net, incorporating pretrained and untrained models. Our primary focus is enhancing the quality of reconstructed 2D images. We evaluate model efficacy through NR-IAQ (No-Reference Image) measurements [32], using BRISQUE for distortion analysis. Lower BRISQUE scores indicate reduced distortion. We also employ NIQE (Natural Image Quality Evaluator) [33] to assess factors like texture and sharpness. In NIQE, higher values imply lower perceptual quality, while lower values reflect higher quality (see Table I). We extend our evaluation to 3D images, analyzing meshes via Skewness and MSE metrics. Through thorough model analysis, we strive to pinpoint the best strategy for superior image and mesh reconstruction. Visual representations are in Figure 4 and Figure 5.

In our 2D phase retrieval experiments, we utilize a 440x440 pixel test image (depicted in Figure 5a). These trials employ a wavelength of 632.8 nm and diffraction distances: 34.36 mm, 34.66 mm, and 37.15 mm. Various untrained deep learning models are employed, including GS-based UNet (Figure 5b) and FR-based Res-U2Net (Figure 5d). The approach involves up to 1000 iterations, a learning rate of 10^{-4}, and reconstitution times (500-800 ms) to assess neural network and physics model performance. The implementation employs the Keras framework in Python [34], utilizing an NVIDIA GTX 1080 GPU for computations.

Our methodology commences by reconstructing the 2D phase, followed by the application of the Shape-From-Shading model (USFSM) to generate a 3D image for each model. This process entails a harmonious fusion of neural network and physics models. The assessment of 3D image quality involves two key metrics: MSE [35] and Skewness (see Table II). Skewness quantifies the symmetry of 3D shapes, where a value approaching 0 indicates optimal mesh quality, while a value close to 1 implies a wholly degenerate mesh [36]. We compare the normalized 3D mesh derived from the phase reconstruction with the normalized 3D mesh of the test images (as illustrated in Fig. 5c and 5e).

TABLE I: Evaluating 2D Phase Retrieval Images Using NR-IAQ BRISQUE and NIQE Comparison of UNet (GS) and Res-U2Net (FR).

Method	$UNet(GS)$	$Res-U2Net(FR)$
BRISQUE	11	2.83
NIQE	5.45	1.85

TABLE II: 3D phase retrieval MSE and Skewness: Comparison of UNet (GS) and Res-U2Net (FR).

Method	$UNet(GS)$	$Res-U2Net(FR)$
MSE	0.134	0.055
Skewness	0.755	0.007

VI. Conclusion and Discussion

This comprehensive study compares 2D and 3D imaging phase-retrieval techniques, evaluating UNet and Res-U2Net

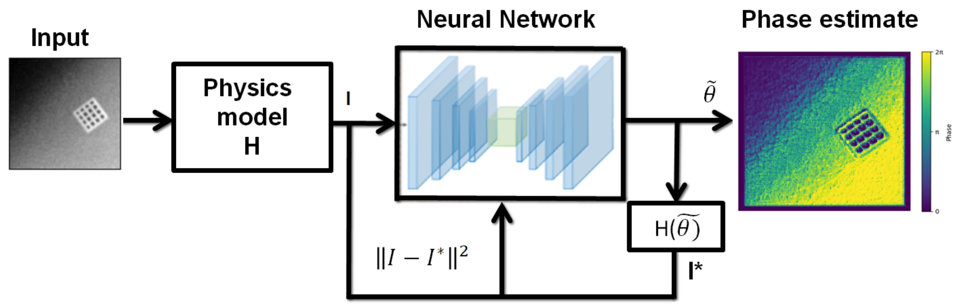


Fig. 1: The phase retrieval process can be summarized by the following block diagram: Initially, we employ the Physics-Model (GS or FR) H to process the input image, yielding a diffration model I. Subsequently, this model serves as the input for a complex neural network. The neural network (Unet or Res-U2Net) produces an output that represents the estimated phase, denoted as $\tilde{\theta}$. To obtain an estimate for the diffration model, denoted as I^*, we apply the Physics-Model to the estimated phase $\tilde{\theta}$. To adjust the parameters of the neural network, including $\tilde{\theta}$, we calculate the mean square error (MSE) between I and I^*, which serves as the loss value.

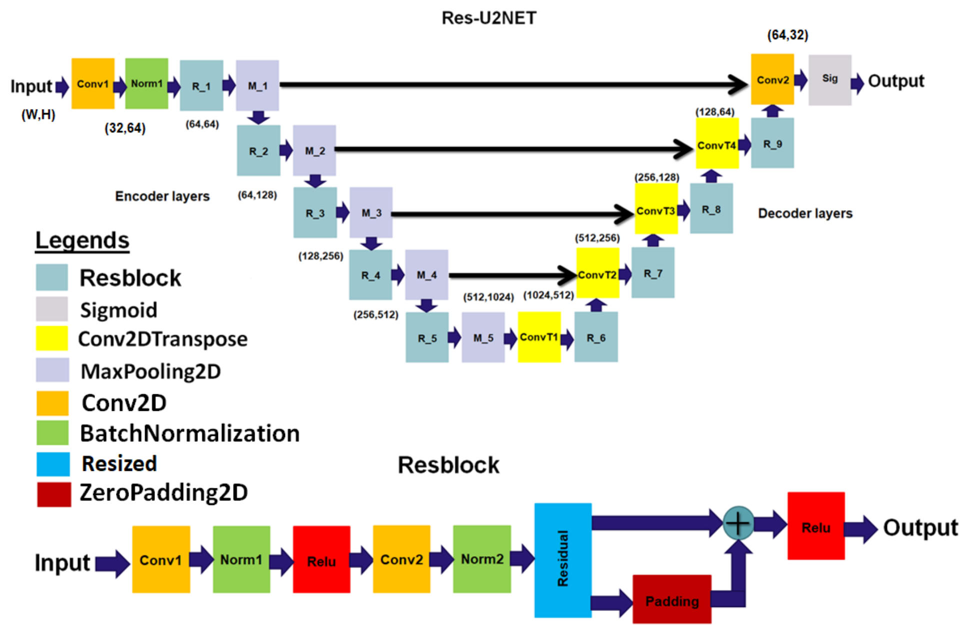
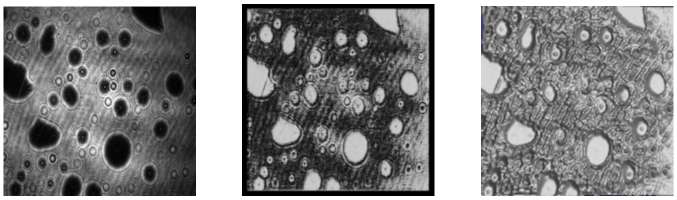


Fig. 2: The Encoder Layer of Res-U2Net demonstrates the sequential steps involved in image segmentation using a convolutional neural network. The process begins by reshaping the input tensor to align with the desired dimensions. Next, features are extracted through convolutional layers, incorporating batch normalization and ReLU activation, forming what is known as a Resblock. To reduce the spatial resolution, max pooling is applied. The Decoder Layer then utilizes transpose convolutions and skip connections to restore and preserve spatial information. Residual connections are employed to further enhance the network's performance. Finally, a 1x1 convolutional layer generates the segmentation mask, resulting in the final output.

TABLE III: Comparative Evaluation of Phase Retrieval Methods: Performance, Complexity, and Processing Analysis

Method	Image Size	Processing Time	Complexity	Performance
SiSPRNet [37]	762x762	3.56 ms/Image	High	Medium
ResNet [38]	762x762	-	High	Good
PhysnNet [5]	440x440	10 min	Medium	Good
UNet (GS)	440x440	500-700 ms	Medium	Low
Res-U2Net (FR)	440x440	500-800 ms	High	Good



(a) (b) (c)

Fig. 3: Example of 3D Phase Retrieval image diffraction with a distance of 27.75 mm and a wavelength of 632.8 nm: (a) 2D Raw image, (b) 2D phase retrieval estimate, and (c) 3D estimation of phase retrieval

neural networks against Gerchberg-Saxton (GS) and Fourier-Rytov (FR) methods. Results highlight the significant advancements achieved by UNet and Res-U2Net in both 2D and 3D reconstructions. The proposed phase model's excellence is particularly evident in 2D images, assessed through NR-IAQ for distortion (BRISQUE), texture, and sharpness. Findings demonstrate the model's superiority with minimal distortion levels, especially in background details (Table I). Notably, Res-U2Net exhibits superior performance, showcasing reduced distortion (Fig. 5d). Metrics unequivocally establish Res-U2Net's supremacy over UNet. For 3D mesh normalized test images, UNet excels in mean squared error (MSE) and symmetry, while Res-U2Net shows promise in both aspects. The study concludes that combining neural and physics-based models effectively addresses 3D phase retrieval, catering to application-specific needs (Table III). Remarkably, Res-U2Net achieves this with a processing time of around 800 ms. Future work should explore additional metrics and model enhancements.

References

[1] D. Zhang and Z. Tan, "A review of optical neural networks," *Applied Sciences*, vol. 12, no. 11, 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/11/5338

[2] S. Liu, X. Chen, T. Yang, C. Guo, J. Zhang, J. Ma, C. Chen, C. Wang, C. Zhang, and S. Liu, "Machine learning aided solution to the inverse problem in optical scatterometry," *Measurement*, vol. 191, p. 110811, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0263224122001087

[3] A. Burvall, U. Lundström, P. A. C. Takman, D. H. Larsson, and H. M. Hertz, "Phase retrieval in x-ray phase-contrast imaging suitable for tomography," *Opt. Express*, vol. 19, no. 11, pp. 10 359–10 376, May 2011. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-19-11-10359

[4] X.-F. Han, H. Laga, and M. Bennamoun, "Image-based 3d object reconstruction: State-of-the-art and trends in the deep learning era," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 5, pp. 1578–1604, 2021.

[5] F. Wang, Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, and G. Situ, "Phase imaging with an untrained neural network," *Light: Science & Applications*, vol. 9, no. 1, p. 77, May 2020. [Online]. Available: https://doi.org/10.1038/s41377-020-0302-3

[6] T. Shimobaba, Y. Endo, T. Nishitsuji, T. Takahashi, Y. Nagahama, S. Hasegawa, M. Sano, R. Hirayama, T. Kakue, A. Shiraki, and T. Ito, "Computational ghost imaging using deep learning," *Optics Communications*, vol. 413, pp. 147–151, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0030401817311628

[7] T. Zeng, Y. Zhu, and E. Y. Lam, "Deep learning for digital holography: a review," *Opt. Express*, vol. 29, no. 24, pp. 40 572–40 593, Nov 2021. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-29-24-40572

[8] R. Horisaki, R. Takagi, and J. Tanida, "Learning-based imaging through scattering media," *Opt. Express*, vol. 24, no. 13, pp. 13 738–13 743, Jun 2016. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-24-13-13738

[9] J. T. Smith, N. Un, R. Yao, N. Sinsuebphon, A. Rudkouskaya, J. Mazurkiewicz, M. Barroso, P. Yan, and X. Intes, "Fluorescent lifetime imaging improved via deep learning," in *Biophotonics Congress: Optics in the Life Sciences Congress 2019 (BODA,BRAIN,NTM,OMA,OMP)*. Optica Publishing Group, 2019, p. NM3C.4. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=NTM-2019-NM3C.4

[10] K. Wang, Y. Li, Q. Kemao, J. Di, and J. Zhao, "One-step robust deep learning phase unwrapping," *Opt. Express*, vol. 27, no. 10, pp. 15 100–15 115, May 2019. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-27-10-15100

[11] G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica*, vol. 6, no. 8, pp. 921–943, Aug 2019. [Online]. Available: https://opg.optica.org/optica/abstract.cfm?URI=optica-6-8-921

[12] R. Shang, K. Hoffer-Hawlik, F. Wang, G. Situ, and G. P. Luke, "Two-step training deep learning framework for computational imaging without physics priors," *Opt. Express*, vol. 29, no. 10, pp. 15 239–15 254, May 2021. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-29-10-15239

[13] K. Monakhova, V. Tran, G. Kuo, and L. Waller, "Untrained networks for compressive lensless photography," *Opt. Express*, vol. 29, no. 13, pp. 20 913–20 929, Jun 2021. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-29-13-20913

[14] V. Lempitsky, A. Vedaldi, and D. Ulyanov, "Deep image prior," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 9446–9454.

[15] R. Heckel and P. Hand, "Deep decoder: Concise image representations from untrained non-convolutional networks," *International Conference on Learning Representations*, 2019.

[16] G. zhen Yang, B. zhen Dong, B. yuan Gu, J. yao Zhuang, and O. K. Ersoy, "Gerchberg–saxton and yang–gu algorithms for phase retrieval in a nonunitary transform system: a comparison," *Appl. Opt.*, vol. 33, no. 2, pp. 209–218, Jan 1994. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-33-2-209

[17] T. E. Gureyev, T. J. Davis, A. Pogany, S. C. Mayo, and S. W. Wilkins, "Optical phase retrieval by use of first born- and rytov-type approximations," *Appl. Opt.*, vol. 43, no. 12, pp. 2418–2430, Apr 2004. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-43-12-2418

[18] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 234–241.

[19] C. Osorio Quero, D. Durini, J. Rangel-Magdaleno, J. Martinez-Carranza, and R. Ramos-Garcia, "Single-pixel near-infrared 3d image reconstruction in outdoor conditions," *Micromachines*, vol. 13, no. 5, 2022. [Online]. Available: https://www.mdpi.com/2072-666X/13/5/795

[20] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. Zaiane, and M. Jagersand, "U2-net: Going deeper with nested u-structure for salient object detection," vol. 106, 2020, p. 107404.

[21] D. R. Luke, "Relaxed averaged alternating reflections for diffraction imaging," *Inverse Problems*, vol. 21, no. 1, p. 37, nov 2004. [Online]. Available: https://dx.doi.org/10.1088/0266-5611/21/1/004

[22] Çağatay Işil, F. S. Oktem, and A. Koç, "Deep iterative reconstruction for phase retrieval," *Appl. Opt.*, vol. 58, no. 20, pp. 5422–5431, Jul 2019. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-58-20-5422

[23] E. Cha, C. Lee, M. Jang, and J. C. Ye, "Deepphasecut: Deep relaxation in phase for unsupervised fourier phase retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 9931–9943, 2022.

[24] F. Wang, A. Eljarrat, J. Müller, T. R. Henninen, R. Erni, and C. T. Koch, "Multi-resolution convolutional neural networks for inverse problems," *Scientific Reports*, vol. 10, no. 1, p. 5730, Mar 2020. [Online]. Available: https://doi.org/10.1038/s41598-020-62484-z

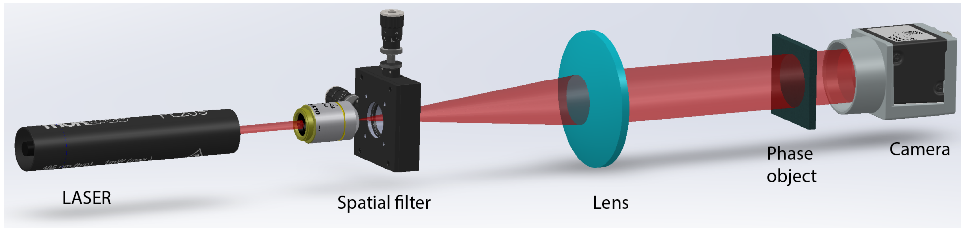


Fig. 4: The schematic diagram illustrates the phase retrieval experiment conducted using a laser with a wavelength of 632.8 nm. The laser beam passes through a spatial filter and a lens to expand its size. The expanded beam then illuminates the object, allowing us to capture the diffracted image using a sensor camera with a pixel size of 4.8 μm.

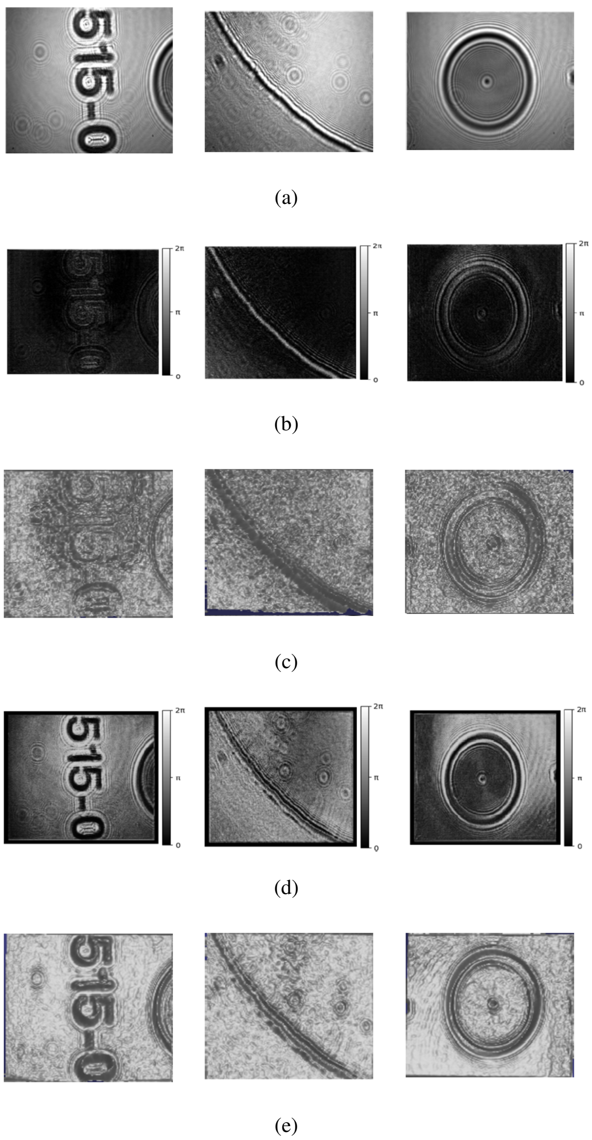


Fig. 5: The image diffraction distances for 2D/3D Phase Retrieval are 34.36 mm, 34.66 mm, and 37.15 mm, respectively, with a wavelength of 632.8 nm. The following components are involved: (a) a test figure, (b-d) 2D phase retrieval estimate using UNet and Res-U2Net, and (c-e) 3D estimation phase retrieval using UNet and Res-U2Net.

[25] J. Dong, L. Valzania, A. Maillard, T.-a. Pham, S. Gigan, and M. Unser, "Phase retrieval: From computational imaging to machine learning: A tutorial," *IEEE Signal Processing Magazine*, vol. 40, no. 1, pp. 45–57, 2023.

[26] G. Jagatap and C. Hegde, "Phase retrieval using untrained neural network priors," in *NeurIPS 2019 Workshop on Solving Inverse Problems with Deep Networks*, 2019. [Online]. Available: https://openreview.net/forum?id=r1l9n725IH

[27] T. E. Gureyev, T. J. Davis, A. Pogany, S. C. Mayo, and S. W. Wilkins, "Optical phase retrieval by use of first born- and rytov-type approximations," *Appl. Opt.*, vol. 43, no. 12, pp. 2418–2430, Apr 2004. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-43-12-2418

[28] B. Blaschke-Kaltenbacher and H. W. Engl, *Regularization Methods for Nonlinear Ill-Posed Problems with Applications to Phase Reconstruction*. Vienna: Springer Vienna, 1997, pp. 17–35. [Online]. Available: https://doi.org/10.1007/978-3-7091-6521-83

[29] J. Shao, K. Zhou, Y.-H. Cai, and D.-Y. Geng, "Application of an improved u2-net model in ultrasound median neural image segmentation," *Ultrasound in Medicine Biology*, vol. 48, no. 12, pp. 2512–2520, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301562922005117

[30] S. Tozza and M. Falcone, "Analysis and approximation of some shape-from-shading models for non-lambertian surfaces," *Journal of Mathematical Imaging and Vision*, vol. 55, no. 2, pp. 153–178, Jun 2016. [Online]. Available: https://doi.org/10.1007/s10851-016-0636-x

[31] C. Y. Kao, S. Osher, and J. Qian, "Lax–friedrichs sweeping scheme for static hamilton–jacobi equations," *Journal of Computational Physics*, vol. 196, no. 1, pp. 367–391, 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021999103006016

[32] A. Mittal, A. K. Moorthy, and A. C. Bovik, "Blind/referenceless image spatial quality evaluator," in *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, 2011, pp. 723–727.

[33] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209–212, 2013.

[34] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-net and its variants for medical image segmentation: A review of theory and applications," *IEEE Access*, vol. 9, pp. 82 031–82 057, 2021.

[35] R. Fulton, S. Eberl, S. Meikle, B. Hutton, and M. Braun, "A practical 3d tomographic method for correcting patient head motion in clinical spect," *IEEE Transactions on Nuclear Science*, vol. 46, no. 3, pp. 667–672, 1999.

[36] C. J. Budd, A. T. McRae, and C. J. Cotter, "The scaling and skewness of optimally transported meshes on the sphere," *Journal of Computational Physics*, vol. 375, pp. 540–564, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021999118305515

[37] Q. Ye, L.-W. Wang, and D. P. K. Lun, "Sisprnet: end-to-end learning for single-shot phase retrieval," *Opt. Express*, vol. 30, no. 18, pp. 31 937–31 958, Aug 2022. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-30-18-31937

[38] Y. Nishizaki, R. Horisaki, K. Kitaguchi, M. Saito, and J. Tanida, "Analysis of non-iterative phase retrieval based on machine learning," *Optical Review*, vol. 27, no. 1, pp. 136–141, Feb 2020. [Online]. Available: https://doi.org/10.1007/s10043-019-00574-8