Amplitude/Phase Retrieval for Terahertz Holography With Supervised and Unsupervised Physics-Informed Deep Learning

Mingjun Xiang, Hui Yuan, Lingxiao Wang, Kai Zhou, and Hartmut G. Roskos

***Abstract*—Most neural networks proposed for computational imaging (CI) in the terahertz (THz) bands require a large amount of experimental data to optimize their weights and biases. However, obtaining a sufficient number of ground-truth images for training is challenging in the THz domain due to the requirements of environmental and system stability, as well as the lengthy data acquisition process. To overcome this limitation, this article proposes novel supervised and unsupervised physics-informed deep learning (DL) methods for amplitude and phase recovery by incorporating angular spectrum diffraction theory as prior knowledge. First, we demonstrate that our *unsupervised* dual network can predict both amplitude and phase simultaneously, overcoming the limitations of previous studies that could only predict phase objects. This is demonstrated using synthetic 2-D image data as well as measured diffraction images. The advantage of unsupervised DL is its ability to be used directly without labeling by human experts. In addition, we address *supervised* DL, which is a concept of general applicability. We introduce training with a database set of 2-D images taken in the visible spectra range and numerically modified by us to emulate THz images. This approach allows us to avoid the prohibitively time-consuming collection of a large number of THz-frequency images. Furthermore, we employ a combination method that enhances the sharpness of image edges, improves contrast, and effectively aligns the approach with the ground truth. The results obtained using both approaches represent the initial steps toward fast holographic THz imaging with reference-beam-free, low-cost power detection.**

Manuscript received 26 July 2023; revised 9 October 2023 and 20 November 2023; accepted 18 December 2023. Date of publication 3 January 2024; date of current version 5 March 2024. This work was supported in part by XF-IJRC, in part by the German Research Foundation DFG under Grant RO 770/48-1, in part by the BMBF under ErUM-Data, and in part by the AI Grant of SAMSON AG, Frankfurt. *(Corresponding authors: Hui Yuan; Kai Zhou.)*

Mingjun Xiang is with the Frankfurt Institute for Advanced Studies (FIAS), 60438 Frankfurt am Main, Germany, also with Physikalisches Institut, Goethe-Universität Frankfurt am Main, 60438 Frankfurt am Main, Germany, and also with Xidian-FIAS International Joint Research Center, 60438 Frankfurt am Main, Germany (e-mail: xiang@fias.uni-frankfurt.de).

Hui Yuan and Hartmut G. Roskos are with Physikalisches Institut, Goethe-Universität Frankfurt am Main, 60438 Frankfurt am Main, Germany (e-mail: yuan@physik.uni-frankfurt.de; roskos@physik.uni-frankfurt.de).

Lingxiao Wang is with the Frankfurt Institute for Advanced Studies (FIAS), 60438 Frankfurt am Main, Germany, and also with Xidian-FIAS International Joint Research Center, 60438 Frankfurt am Main, Germany (e-mail: lwang@fias.uni-frankfurt.de).

Kai Zhou is with the Frankfurt Institute for Advanced Studies (FIAS), 60438 Frankfurt am Main, Germany, also with Xidian-FIAS International Joint Research Center, 60438 Frankfurt am Main, Germany, and also with the School of Science, Engineering, The Chinese University of Hong Kong, Shenzhen 518172, China (e-mail: zhou@fias.uni-frankfurt.de).

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TTHZ.2024.3349482.

Digital Object Identifier 10.1109/TTHZ.2024.3349482

***Index Terms*—Amplitude estimation, deep learning, phase esitmation, signal reconstruction, terahertz wave imaging.**

I. Introduction

IMAGING at terahertz (THz) frequencies (0.1–10 THz, wavelengths of 3 mm–30 μm) has received considerable attention in recent years. Application areas for THz imaging explored at present include nondestructive testing [1], quality monitoring [2], security screening [3], [4], biomedical imaging [5], and sensing for robotics and vehicle control [6]. The applied imaging modalities increasingly include coherent holographic approaches [7], variously because they allow the capture of comparatively large scenes with good spatial resolution, offer the possibility of 3-D scene reconstruction, may have a relaxed demand on the optical imaging optics, or lend themselves for advanced numerical image processing, e.g., exploiting sparsity effects. However, at lower THz frequencies, and especially in the sub-1-THz frequency range, which is most relevant for many applications, one is forced to perform coherent detection as a serial rather than a parallel process [8], [9], [10]. The main reason is that power availability at these THz frequencies is a critical issue [11]. While detector arrays (with pixel numbers limited to hundreds or at most a few thousand [7]) in principle are available, the limited power makes it difficult to provide a reference beam for multipixel interferometric or heterodyne phase measurements. The need for serial data recording makes coherent imaging time-consuming and is not conducive to image reconstruction, since phase distortion and noise introduced during signal recording strongly affect data quality, especially in the case of weak signals [12], [13].

To overcome these limitations, it would be highly advantageous to perform computational phase retrieval from the amplitude or intensity diffraction pattern obtained with the object beam alone. This approach would significantly reduce power requirements, simplify the measurement system, and accelerate the data acquisition process. It would enable a transition from coherent detection with single ultrafast receivers to low-cost power detection for THz holography.

Conventional methods suffer from either slow convergence or slow image capture speed in the THz bands. For instance, the Gerchberg–Saxton (GS) algorithm relies on iterative processing of diffraction patterns recorded at different distances to improve convergence and reliability [14], [15], [16]. Using

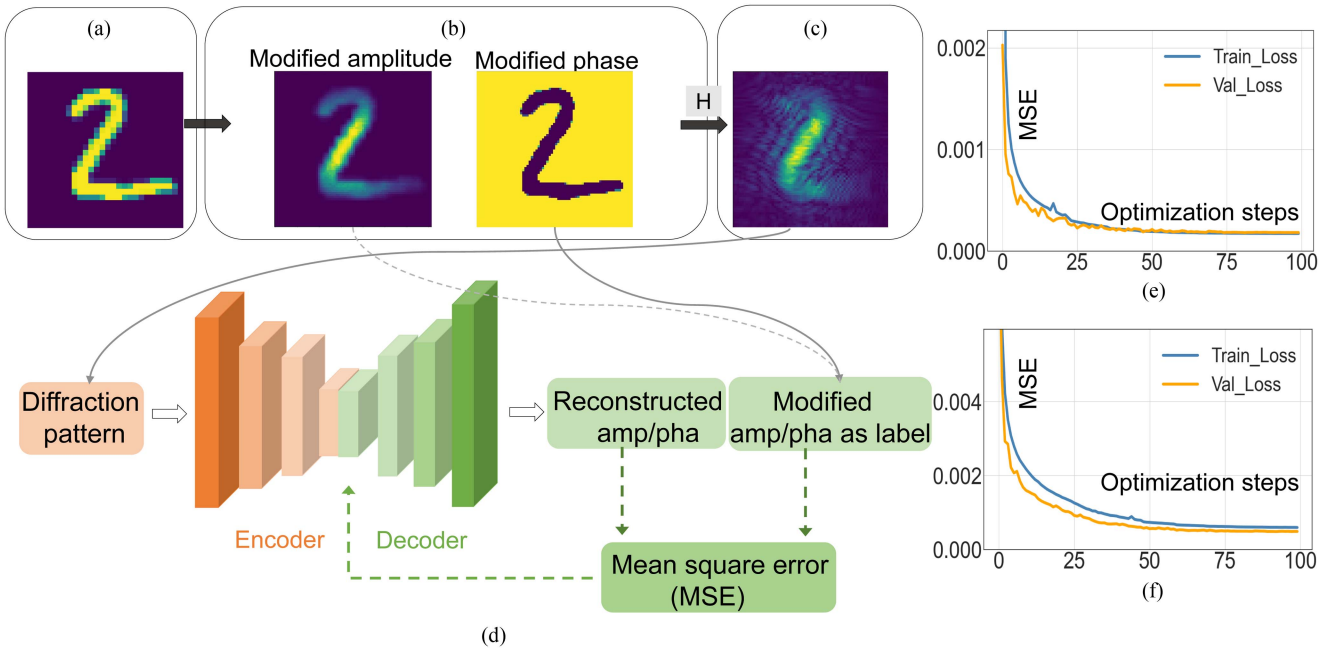

Fig. 1. Illustration of the supervised phase reconstruction model. (a) Random original image from the MNIST dataset. (b) Modified amplitude and phase image. (c) Amplitude diffraction pattern in the THz domain calculated by the physical model. (d) Scheme of the convolutional neural network (CNN) of the supervised DL method. A measured diffraction pattern is the input of the CNN, and the output is an amplitude/phase image. The mean squared error (MSE) between this amplitude/phase reconstruction and the ground truth amplitude/phase is taken as the loss function to optimize the CNN (see Supplementary for further details). (e), (f) Evolution of the MSE for amplitude and phase during pretraining (blue lines) with 60 000 images and testing with 10 000 ones (orange lines).

only one diffraction pattern for amplitude/phase retrieval always yields poor results due to the lack of constraint conditions and prior knowledge of the imaged targets [17], [18], [19]. On the other hand, approaches like the multiple-plane phase retrieval algorithm [20], [21], [22] employ wave propagation in multiple planes with recorded data, eliminating the need for support constraints and prior knowledge of the project. However, one drawback of this method is its slower imaging speed, taking, for example, around 36 h [22] to capture an image. Recently, deep learning (DL) techniques have emerged as a promising and efficient method for precise image reconstruction [23], [24], [25]. These techniques include supervised DL methods that rely on labeled experimental data [26], [27], [28], as well as unsupervised DL methods suitable for unlabeled data, such as PhysenNet [29]. Nevertheless, the expensive acquisition of THz images poses a challenge for supervised training with experimental data.

In this article, we propose two novel physics-informed DL methods for phase retrieval in THz holographic imaging. These methods incorporate prior knowledge of angular spectrum diffraction theory [30], [31], [32]. Physics-informed DL combines mathematical physics models and data synergistically, ensuring that the computations respect the physical laws and symmetries of the modeled system, typically described by ordinary/partial differential equations or integral equations [33], [34]. This approach has demonstrated its powerful capability in solving inverse problems in physics [35], [36], [37]. DL methods can also be employed for classification tasks in THz imaging, such as distinguishing the shapes of 2-D objects based on their diffraction patterns [38]. However, it is important to note that classification tasks rely on a limited number of known types of objects, shapes, or other characteristics, and can neither be applied for classification outside of this pool nor for image reconstruction. In this article, we explore techniques for image reconstruction without type restriction, utilizing both supervised and unsupervised DL methods.

The proposed supervised DL method utilizes the visible-light MNIST dataset for training, which addresses the issue of slow recording of THz frequency images by requiring only the recording of a reference beam. Compared to unsupervised learning methods, it requires longer training time but provides high prediction speed and robustness against noise in experimentally generated images after training [see Fig. 3(i) for details]. On the other hand, the unsupervised method can be used without labeling or extensive data preparation, and it enables simultaneous reconstruction of amplitude and phase, expanding the working range of previous unsupervised DL methods.

II. Methods

Fig. 1 displays the flowchart that illustrates the proposed supervised amplitude/phase reconstruction, demonstrated using handwritten digits as an example. Since a sufficiently large THz dataset is not available, we utilize the public MNIST dataset, which comprises thousands of photographs taken in the visible spectral range. These MNIST images are preprocessed to transform them into effective THz images, as depicted in Fig. 1(a)–(c). This preprocessing involves two steps: 1) superimposing the Gaussian field-amplitude profile of a measured THz beam (recorded in the experimental setup discussed below) onto

the images to obtain the modified amplitude, and 2) determining the phase contour based on the object's material and thickness to obtain the modified phase.

The specific process is to first measure the amplitude and phase maps of the THz beam. A measurement example can be found in the supporting information, demonstrating that the beam has an isotropic amplitude distribution and near-constant phase over its cross section. A numerical evaluation of the beam profile reveals a near-Gaussian beam shape. The MNIST dataset consists of images with dimensions of 28×28 pixels. We resize them to a pixel number of 72×72 equal to that of the data that we collect in our experiment. Afterward, we multiply the MNIST data by the reference amplitude map. Regarding the phase assignment, which can be challenging for 3-D objects, we benefit from the 2-D nature of the objects considered in this article. Since our validation experiment uses metal screens with cut-outs as objects, the phase at the metallic portions is random. It can be set to any value because it will be multiplied by an amplitude of zero. The phase at the open sample regions is set to be the phase in the object plane back-propagated from the recorded reference beam.

Next, we employ the physical model H (described below) to calculate the field-amplitude diffraction pattern of the effective THz image using the Huygens-Fresnel principle [30]. This diffraction pattern serves as input to the convolutional neural network (CNN) depicted in Fig. 1(d), which is inspired by U-Net [39]. The CNN architecture consists of a downsampling path and a symmetric upsampling path (for a detailed description, refer to the Supplementary material). The output of the network is the reconstructed phase images. Subsequently, the same network is used for amplitude retrieval from the diffraction patterns.

The learning curves shown in Fig. 1(e) and (f) demonstrate the convergence of the network to low validation loss values after approximately 50 epochs, indicating that the network has not overfitted the training dataset. For amplitude and phase, we achieve validation losses (MSE) of 0.00018 and 0.00048, respectively, after 100 epochs. In image processing, MSE is commonly used to measure the quality of the reconstructed image compared to the ground truth. It represents the Euclidean distance between images and is simple and efficient to evaluate, making it suitable for large-scale image processing tasks. In addition, MSE is a convex function and possesses good mathematical properties [40].

The physical model, H, simulates the experimental THz imaging process. If a planar object is illuminated by a beam, the complex-valued field amplitude immediately behind the object can be written as

$$E_0(x, y, z = 0) = A_0(x, y, 0)e^{i\phi_0(x,y;0)}, \tag{1}$$

where A_0 and ϕ_0 are the amplitude and the phase at the object plane. Over a distance d, diffraction reshapes the field as follows [30], [31]:

$$E_d(x, y, z = d) = \iint \hat{E}_0(f_x, f_y)\, G\, e^{i2\pi(f_x x + f_y y)} df_x\, df_y \tag{2}$$

where $G = e^{ikd\sqrt{1-\lambda^2 f_x^2 - \lambda^2 f_y^2}}$ is the wave propagation function, λ the wavelength, $\hat{E}_0$ the spatial Fourier transform of E_0 with $f_x = x/\lambda d$ and $f_y = y/\lambda d$ as the spatial frequencies in the x and y directions. The diffraction pattern is the field's absolute value

$$A(x, y, z = d) = |E_d(x, y, z = d)| = H(\phi_0, A_0) \tag{3}$$

where $H(\cdot)$ represents the mapping function relating the object to the diffraction pattern A. The challenge is now to achieve an inverse mapping, $H^{-1}(\cdot)$, such that

$$\phi_0(x, y, 0) = H^{-1}(A(x, y, z = d)). \tag{4}$$

The supervised method of Fig. 1 utilizes a parameterized network function R_θ (θ denoting the network weights and bias parameters) to approach the desired inverse mapping $H^{-1}(\cdot)$ via learning based upon the labeled training set $S_T = (A_k, \phi_k), k = 1, 2, \ldots, K$, thus solving the optimization problem of

$$R_{\theta^*} = \arg\min_\theta \|R_\theta(A_k) - \phi_k\|^2 \qquad \forall(\phi_k, A_k) \in S_T. \tag{5}$$

A corresponding inverse mapping reconstructs the amplitude information of the object, A_0. A different strategy is applied with the proposed unsupervised DL method shown in Fig. 2. Without the need for pretraining on large labeled datasets, it works directly on the amplitude $A(x, y, z = d)$, i.e., the propagated diffraction pattern of the object. Unlike PhysenNet, which was developed for visible-range images and only phase objects [29], the model proposed here is suitable for all holographic systems and has no object restrictions. The field-amplitude diffraction pattern is the only input to the CNN, which is designed to generate estimated phase and amplitude maps simultaneously (see Supplementary for a detailed description). The two output paths share the CNN's front layers in order to ensure that the same object information is analyzed, and the path splits in two only at the backend. The model H is then applied to convert the network outputs to an estimated diffraction pattern. The MSE between the input and the estimated diffraction pattern is fed back to the network to optimize the weights and bias values via gradient descent. Thus, the retrieval of the phase is formulated as

$$R_{\theta^*} = \arg\min_\theta \|H(R_\theta(A)) - A)\|^2. \tag{6}$$

This will force the calculated diffraction pattern to converge to the measured pattern, as the iterative process proceeds.

III. Results and Discussion

We demonstrate the performance of the proposed supervised and unsupervised methods using data derived from simulations (THz-emulated MNIST) and experiments, respectively. In the simulations, we compared the proposed methods with two typical phase retrieval methods, namely the Gerchberg–Saxton (GS) algorithm [15] and in-line holography direct reconstruction [41].

The results are illustrated in Fig. 3. We used randomly chosen digits "5" and "6" for demonstration because they encompass

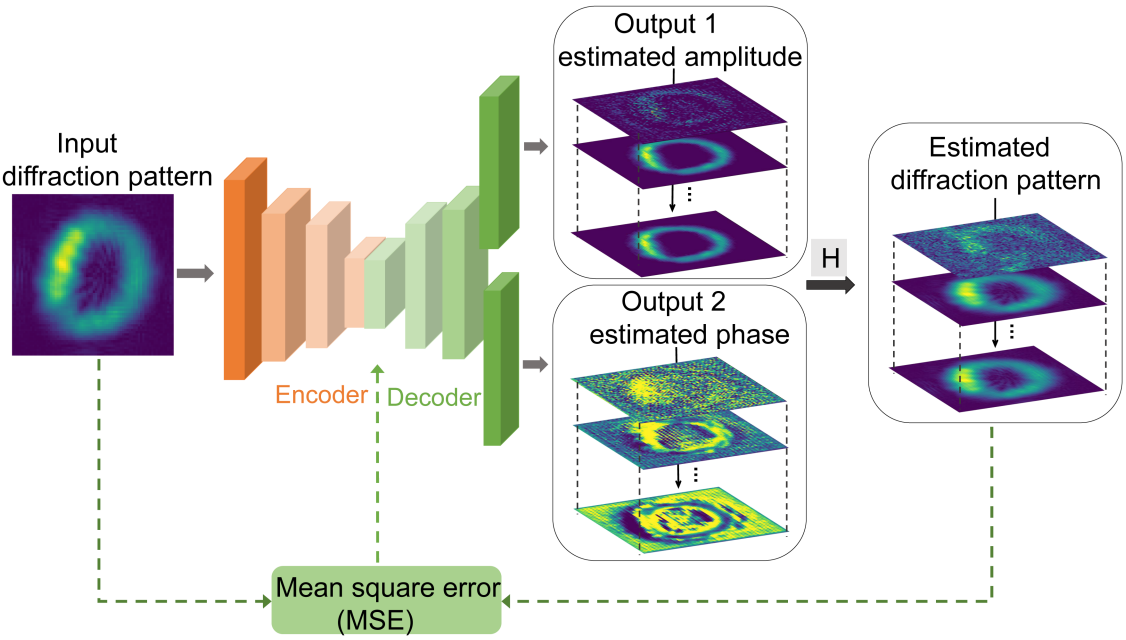

Fig. 2. Schematic illustration of the pipeline of unsupervised DL. The input to the NN is a measured field-amplitude diffraction pattern, the outputs are estimated phase and amplitude maps, which are then numerically propagated via the model H to simulate the diffraction and measurement processes, yielding an estimated diffraction pattern. The MSE between the true and estimated patterns provides the loss function used to optimize the NN parameters (see Supplementary).

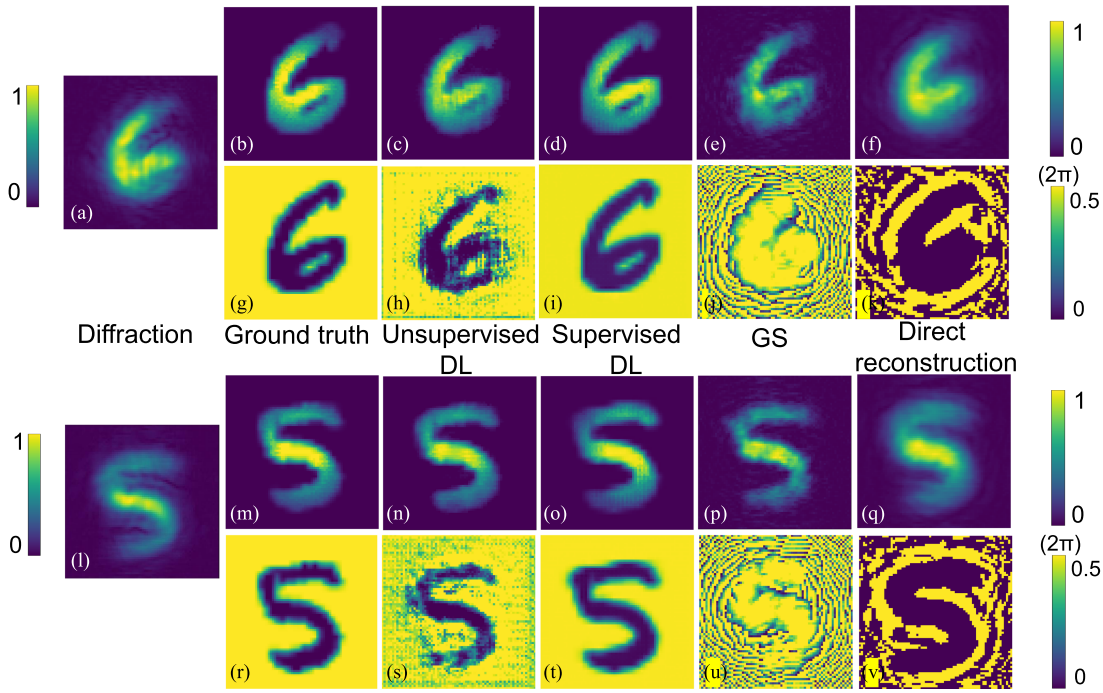

Fig. 3. Comparison of different amplitude/phase retrieval methods. (a), (l) Diffraction patterns of a digit "6" and "5" from the testing dataset. (b), (g) and (m), (r) Ground truth of the objects' amplitude and phase. (c)–(f) and (n)–(q) The amplitudes reconstructed by means of (c), (n) the unsupervised DL method; (d), (o) the supervised DL method; (e), (p) the GS algorithm; (f), (q) the direct reconstruction. (h)–(k), (s)–(v) The phases reconstructed by means of (h), (s) the unsupervised DL method; (i), (t) the supervised DL method; (j), (u) the GS algorithm; (k), (v) the direct reconstruction.

various shapes found in the dataset, such as horizontal, vertical, and curved lines. The units of the phases' colorbars are $[0, 0.5] * 2\pi$. Note that the diffraction patterns shown in Fig. 3(a) and (l) are the only input to both methods. We used the MSE to measure the quality of the reconstructed amplitude/phase images compared to the ground truth shown in Fig. 3(b), (g) and (m), (r). As mentioned earlier, the proposed unsupervised DL method works directly on diffraction data without prior training. As the optimization progresses, the MSE drops to 10^{-6} after 500 epochs, resulting in concurrently updated amplitude and

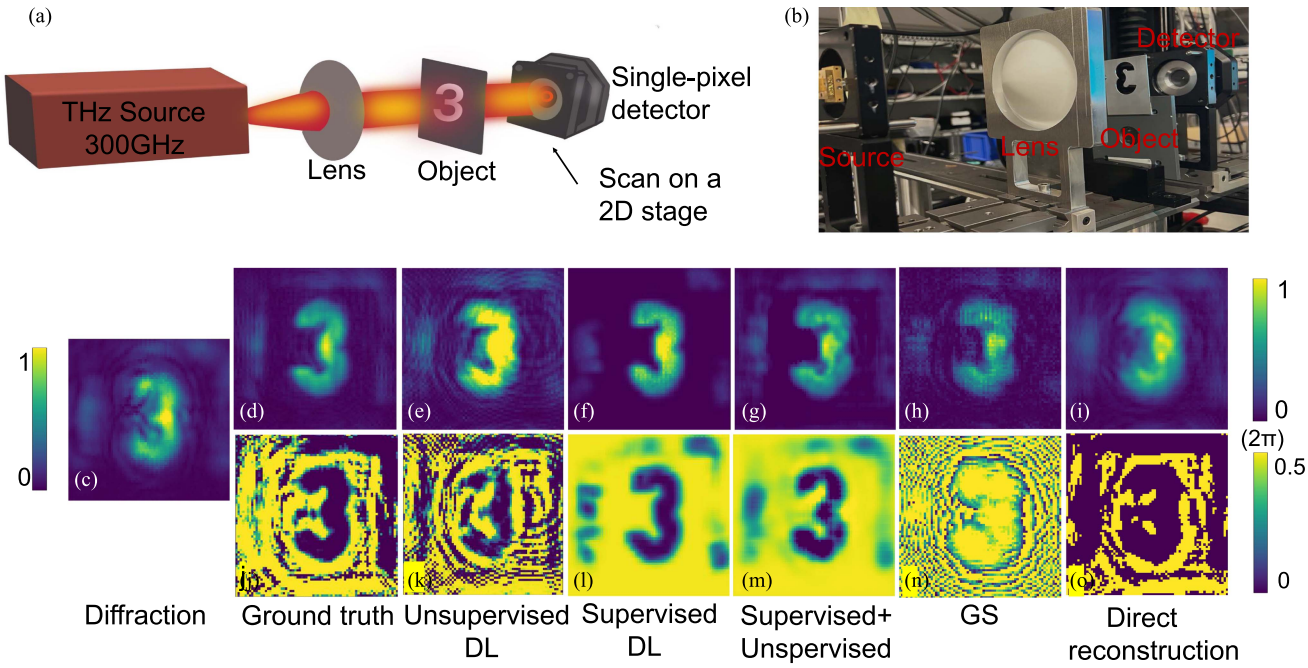

Fig. 4. Experimental setup and detected images. (a) And (b) display a schematic of the THz holographic imaging system and the photograph of our setup. (c) Shows the diffraction pattern we measured. (d) And (j) are the ground truth amplitude and phase. (e)–(i) and (k)–(o) Visualize the amplitude and phase images recorded by the different methods; (e), (k) the unsupervised DL method; (f), (l) the supervised DL method; (g), (m) the sequential combination of the two methods; (h), (n) GS algorithm; (i), (o) direct reconstruction.

TABLE I
COMPARISON OF DIFFERENT AMP/PHA RETRIEVAL METHODS

Method	Quality (Average MSE) amp / pha	Pre-training	Inference time
Unsupervised DL	0.0011 / 0.0182	No	∼1 min
Supervised DL	0.0004 / 0.0021	Yes	$<$ 0.1 s
GS algorithm	0.0035 / 0.3067	No	∼1 min
Direct reconstruction	0.0096 / 0.2066	No	$<$ 0.1 s

phase patterns shown in Fig. 3(c), (h) and (n), (s). The MSE values between the amplitude/phase reconstructed using unsupervised DL and the ground truths are 0.0016 and 0.0166 for the digit "6," and 0.0006 and 0.0199 for the digit "5." The supervised DL predictions are shown in Fig. 3(d), (i) and (o), (t). The MSE values between the reconstructed amplitude/phase and the ground truths are 0.00058 and 0.0023 for the digit "6," and 0.00040 and 0.0019 for the digit "5," while the corresponding values associated with the GS algorithm [Fig. 3(e), (j) and (p), (u)] are 0.0039 and 0.3029 for the digit "6," and 0.0031 and 0.3104 for the digit "5." As shown in Fig. 3(f), (k) and (q), (v), the MSE values between the direct reconstructions and the ground truths are 0.0096, 0.2079 for the digit "6," and 0.0095, 0.2053 for the digit "5."

For the proposed supervised and unsupervised DL methods, amplitude/phase prediction has at least an order of magnitude advantage over traditional methods. All the methods were performed on a PC with a sixteen-core 3.50-GHz CPU and 64 GB of RAM, using an Nvidia GeForce RTX 3080 GPU. On average, the pretraining of the supervised CNN took ∼4 h for 60 000 pairs of training data and 10 000 pairs of validation data during ∼100 training epochs. The inference time of the trained network for a diffraction pattern of 72×72 pixels was $<$ 0.1 s. The optimization process of the unsupervised NN (with amplitude and phase channel) took approximately ∼1 min over ∼500 epochs. The optimization process of the GS algorithm took approximately ∼1 min over ∼10 000 iterations. Table I summarizes the comparison.

To validate the proposed methods by experiments, we performed THz measurements at 300 GHz. The schematic diagram and experimental photographs are shown in Fig. 4(a) and (b). One can see that this is actually a single-beam lens-less imaging geometry. The illuminating radiation from 300 GHz, THz source is collimated by a focusing lens, with a 10-cm focal length and 4-in aperture. To ensure a large numerical aperture (NA), the lens has a hyperboloidal-planar shape to avoid spherical aberrations that would occur with a spherical lens. The NA of the lens must be large enough to illuminate the objects over their full extension with a collimated Gaussian beam, free from beam modulations caused by diffraction at the lens edge. Such modulations introduce high-spatial-frequency signatures which can make the learning process difficult. In addition, the alignment of the lens is crucial since the phase is sensitive to optical-axis displacements.

To acquire the diffraction pattern, we placed the THz single-pixel TeraFET detector [42] at a distance $d = 70$ mm from the object, scanning on a 2-D translation stage. The radiation enters the detector chip from the back-side, passing a Si substrate lens with a diameter of 4 mm [43], [44]. The recorded diffraction pattern is shown in Fig. 4(c). The objects consisted of thin metal screens with cut-outs in the form of digits. The size of the object

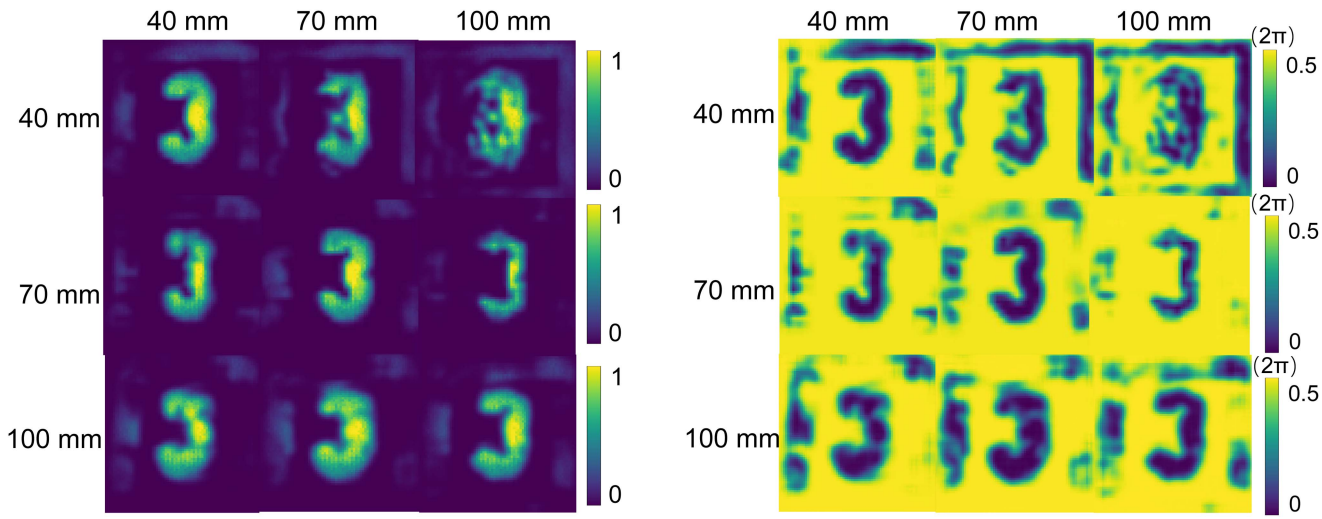

Fig. 5. Effect of d on the reconstructed amplitude/phase. Reconstructed amplitude/phase images of the object for different object distances (varied along the columns) and recovery distances (varied along the rows).

is 50 mm× 70 mm, and the scanned image area is 72 × 72 pixels with a 1-mm^2 pixel size pitch.

The ground-truth amplitude and phase maps, shown in Fig. 4(d) and (j), are obtained from the heterodyne system using the inverse 2-D FFT tool of MATLAB and wave back-propagation, inverting (2) given in the main text. More information can be found in Supplementary material.

We reconstructed the amplitude and phase in five different ways: using unsupervised and supervised DL [Fig. 4(e) and (f), (k) and (l)], and a combination of both in a sequential process [Fig. 4(g), (m)] where the prediction obtained by supervised DL served as input to the following unsupervised process, compared with (h), (n) from the traditional GS algorithm and (i), (o) from direct reconstruction. In the combination method, our initial values are no longer random, and we start closer to the global optimal solution, which allows us to converge to the optimal solution more quickly. For example, in this case, it only takes 200 epochs to obtain a prediction result of Fig. 4(g) and (m), which is much closer to the ground truth. In the GS algorithm, we set the constraint condition for the input plane (object plane) to 1 [45], [46], and the constraint condition for the output plane (diffraction pattern plane) to the single diffraction pattern for the sake of fairness in comparison with other algorithms.

The MSE between the amplitude (phase) reconstructed from the unsupervised method and the ground truths is 0.006 (0.05). For the supervised DL, the corresponding values are 0.0084 (0.0249) for the reconstructed amplitude (phase). The combined method yields the best MSE values: 0.0035 (0.0230) for amplitude (phase), while the MSE for the GS algorithm is 0.0049 (0.0931), and 0.0075 (0.063) for direct reconstruction.

The image from the combined method exhibits sharper edges and better contrast, and is closer to the ground truth. We attribute this considerable improvement to the suppression of the influence of noise, which is present in the measured images, by the supervised DL step. This denoising effect invites further studies in the future.

Next, we numerically analyzed the effect of the diffraction distance d on the quality of the reconstructed image in the supervised DL method. We took three diffraction distances, $d = 40$ mm, $d = 70$ mm, and $d = 100$ mm, as examples to examine the performance. Fig. 5 shows a series of reconstructed images (left: amplitudes; right: phases) obtained with the supervised DL method. For each diffraction map, we assumed these three values of d for the reconstruction. This yields the matrix plot, where the variation of the object positions occurs along the columns and the variation of the recovery distance along the rows. Object and recovery distances coincide along the diagonal of the matrix plot. Similar to physical defocusing, an unblurred image is obtained only in this case, and the image quality degrades if the recovery distance deviates from the object distance. Supervised learning is, hence, sensitive to the input of the correct value of d in a similar way as are traditional algorithms, including direct reconstruction and the GS algorithm, and also unsupervised learning.

Furthermore, we validated the generalization ability of our proposed supervised DL method. We trained the network using the THz-emulated MNIST dataset and then tested it on three objects that had never appeared in the dataset before: a wrench, a grid, and a bolt. These objects had completely different contours, sizes, and positions compared to the MNIST digit dataset, with the bolt being an irregular 3-D structure. Fig. 6 illustrates the comparison between the predictions of our neural network and those of traditional algorithms. The first column of each algorithm [Fig. 6(c)–(f)] represents the phase recovery results, while the second column shows the amplitude recovery results. It is evident that the GS algorithm struggles to accurately reconstruct the amplitude and phase of complex objects without additional constraints. In terms of phase recovery, our supervised approach demonstrates a clear advantage over direct reconstruction methods, as it more accurately delineates object contours, especially in the case of the wrench. The amplitude recovery results of the direct reconstruction method are decent; however, compared to

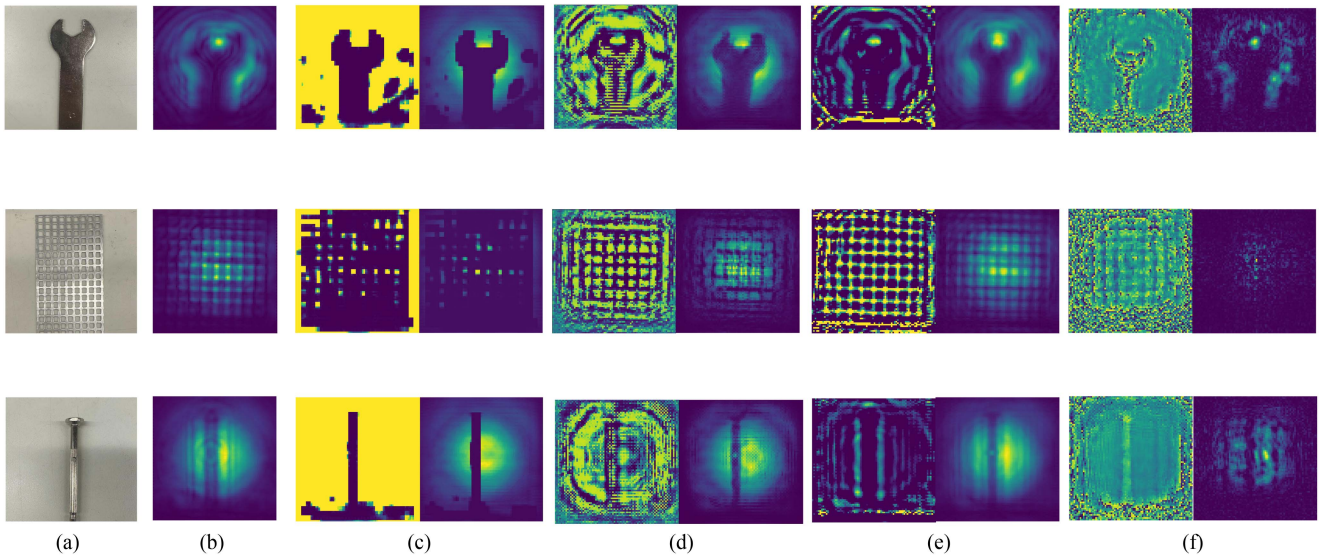

Fig. 6. Comparison of different amplitude/phase retrieval methods for different shapes. Column (a) represents the ground truth, while column (b) displays the diffraction patterns, which serve as the input information for our predictions. Columns (c), (d), (e), and (f) depict the reconstruction results of our proposed supervised learning method, unsupervised learning method, direct reconstruction method, and GS method, respectively. The first column of each method represents the phase recovery results, while the second column represents the amplitude recovery results. (a) Object. (b) Diffraction. (c) Supervised DL. (d) Unsupervised DL. (e) Unsupervised DL. (f) GS.

the supervised learning approach, the contours are less defined, indicating the presence of more noise and a limited ability to remove correlated noise. It is believed that our supervised DL method will be more prominent than the traditional methods when more efficient datasets are added. Regarding the unsupervised DL method, the phase prediction evidently outperforms the traditional algorithms as well. Amplitude prediction shows a slight improvement compared to direct reconstruction results.

IV. Conclusion

In conclusion, we have proposed supervised and unsupervised deep-learning methods, as well as a combined approach, for the reconstruction of 2-D THz images. These methods start from amplitude diffraction patterns and derive the amplitude and phase in the object plane. Both simulated and experimental data demonstrate the superior performance of deep learning methods compared to traditional algorithms in terms of amplitude and phase recovery. By integrating a physical model with conventional deep neural networks, we are able to reconstruct object phases using a single intensity measurement.

It is important to note that our supervised learning approach differs from traditional deep learning methods used in computational imaging. Unlike those methods, our approach does not rely on a mapping function derived from the statistical information of a large set of experimental data (represented by the weights of the network) for image reconstruction. Instead, we prepare the dataset by converting a large number of publicly accessible visible-range images into simulated complex-valued THz images, which only take 1 h for 70 000 pairs.

The resulting reconstructed images, from different distances diffraction patterns, demonstrate that optimal image quality is achieved when the recovery distance matches the object distance, resembling the effect of physical defocusing. Deviation from this alignment leads to degraded image quality. Our supervised learning network demonstrated strong generalization capabilities when tested on previously unseen objects (wrench, grid, and bolt) that differed greatly from the MNIST dataset. Our approach outperformed traditional algorithms, particularly in accurately reconstructing phase and delineating object contours, especially for the wrench.

Furthermore, we would like to emphasize the unsupervised method introduced in this study. This method incorporates a dual-output layer and integrates the physical model of wave propagation to predict amplitude and phase information simultaneously. It does not learn a mapping function from the statistical information of the training data but rather relies on the interplay between a handcrafted network structure and a physical image formation model. This network is specifically tailored to perform well in reconstruction from the given input, although it may sacrifice some generalization ability.

The best reconstruction performance using measured image data is achieved through the sequential application of both methods. The initial step with the supervised method effectively reduces the noise in the raw images and produces a good initial image estimate for the subsequent unsupervised method. These findings validate the potential of amplitude/phase retrieval from amplitude images of THz diffraction patterns, particularly for relatively simple scenes.

Acknowledgment

Disclosures: The authors declare no conflicts of interest.

Data Availability Statement: Raw data will be made available upon reasonable request.

REFERENCES

[1] I. Amenabar, F. Lopez, and A. Mendikute, "In introductory review to THz non-destructive testing of composite mater," *J. Infrared, Millimeter, Terahertz Waves*, vol. 34, no. 2, pp. 152–169, 2013.

[2] F. Ellrich et al., "Terahertz quality inspection for automotive and aviation industries," *J. Infrared, Millimeter, Terahertz Waves*, vol. 41, no. 4, pp. 470–489, 2020.

[3] J. F. Federici et al., "THz imaging and sensing for security applications-explosives, weapons and drugs," *Semicond. Sci. Technol.*, vol. 20, no. 7, pp. S266–S280, Jun. 2005. [Online]. Available: https://doi.org/10.1088/0268-1242/20/7/018

[4] F. Friederich et al., "THz active imaging systems with real-time capabilities," *IEEE Trans. THz Sci. Technol.*, vol. 1, pp. 183–200, Sep. 2011.

[5] X. Yang et al., "Biomedical applications of terahertz spectroscopy and imaging," *Trends Biotechnol.*, vol. 34, no. 10, pp. 810–824, 2016.

[6] D. Jasteh et al., "Low-THz imaging radar for outdoor applications," in *Proc. 16th Int. Radar Symp.*, 2015, pp. 203–208.

[7] G. Valušis, A. Lisauskas, H. Yuan, W. Knap, and H. G. Roskos, "Roadmap of terahertz imaging 2021," *Sensors*, vol. 21, 2021, Art. no. 4092.

[8] A. Nahata, D. H. Auston, T. F. Heinz, and C. Wu, "Coherent detection of freely propagating terahertz radiation by electro-optic sampling," *Appl. Phys. Lett.*, vol. 68, no. 2, pp. 150–152, 1996.

[9] H. Yuan, D. Voß, A. Lisauskas, D. Mundy, and H. G. Roskos, "3D Fourier imaging based on 2D heterodyne detection at THz frequencies," *APL Photon.*, vol. 4, 2019, Art. no. 106108.

[10] H. Yuan, A. Lisauskas, M. D. Thomson, and H. G. Roskos, "600-GHz Fourier imaging based on heterodyne detection at the 2nd sub-harmonic," *IEEE Trans. THz Sci. Technol.*, vol. 31, no. 24, pp. 40856–40870, 2023.

[11] S. Verghese, K. McIntosh, and E. Brown, "Optical and terahertz power limits in the low-temperature-grown GaAS photomixers," *Appl. Phys. Lett.*, vol. 71, no. 19, pp. 2743–2745, 1997.

[12] M. Wan, J. J. Healy, and J. T. Sheridan, "Terahertz phase imaging and biomedical applications," *Opt. Laser Technol.*, vol. 122, 2020, Art. no. 105859.

[13] H. Yuan, M. Wan, A. Lisauskas, J. T. Sheridan, and H. G. Roskos, "300-GHz holography with heterodyne detection," in *Digit. Holography Three-Dimensional Imag.*, 2019, Art. no. Th3A.21. [Online]. Available: http://opg.optica.org/abstract.cfm?URI=DH-2019-Th3A.21

[14] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–250, 1972.

[15] J. R. Fienup, "Phase retrieval algorithms: A comparison," *Appl. Opt.*, vol. 21, no. 15, pp. 2758–2769, 1982.

[16] K. Jaganathan, Y. C. Eldar, and B. Hassibi, "Phase retrieval: An overview of recent developments," *Opt. Compressive Imag.*, pp. 279–312, 2016.

[17] L. Rong et al., "Iterative solution to twin image problem in in-line digital holography," *Opt. Lasers Eng.*, vol. 51, no. 5, pp. 553–559, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0143816612003454

[18] L. Rong, F. Pan, W. Xiao, Y. Li, and F. Wang, "Twin image elimination from two in-line holograms via phase retrieval," *Chin. Opt. Lett.*, vol. 10, no. 6, Jun. 2012, Art. no. 060902. [Online]. Available: https://opg.optica.org/col/abstract.cfm?URI=col-10-6-060902

[19] J. Hu, Q. Li, and S. Cui, "Research on object-plane constraints and hologram expansion in phase retrieval algorithms for continuous-wave terahertz inline digital holography reconstruction," *Appl. Opt.*, vol. 53, no. 30, pp. 7112–7119, Oct. 2014. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-53-30-7112

[20] C. Xing, F. Qi, and S. Guo, "Enhanced terahertz phase retrieval imaging by unequal spaced measurement," *Sensors*, vol. 22, no. 10, 2022, Art. no. 3816. [Online]. Available: https://www.mdpi.com/1424-8220/22/10/3816

[21] G. Hislop, G. C. James, and A. Hellicar, "Phase retrieval of scattered fields," *IEEE Trans. Antennas Propag.*, vol. 55, no. 8, pp. 2332–2341, Aug. 2007.

[22] N. V. Petrov et al., "Terahertz phase retrieval imaging in reflection," *Opt. Lett.*, vol. 45, no. 15, pp. 4168–4171, Aug. 2020. [Online]. Available: https://opg.optica.org/ol/abstract.cfm?URI=ol-45-15-4168

[23] A. Goy, K. Arthur, S. Li, and G. Barbastathis, "Low photon count phase retrieval using deep learning," *Phys. Rev. Lett.*, vol. 121, no. 24, 2018, Art. no. 243902.

[24] G. Zhang et al., "Fast phase retrieval in off-axis digital holographic microscopy through deep learning," *Opt. Exp.*, vol. 26, no. 15, pp. 19388–19405, 2018.

[25] P. Hand, O. Leong, and V. Voroninski, "Phase retrieval under a generative prior," in *Proc. 32nd Int. Conf. Neural Inf. Process. Syst.*, 2018, pp. 9154–9164.

[26] L. Boominathan, M. Maniparambil, H. Gupta, R. Baburajan, and K. Mitra, "Phase retrieval for Fourier ptychography under varying amount of measurements," 2018, *arXiv:1805.03593*.

[27] M. Deng, S. Li, A. Goy, I. Kang, and G. Barbastathis, "Learning to synthesize: Robust phase retrieval at low photon counts," *Light: Sci. Appl.*, vol. 9, no. 1, pp. 1–16, 2020.

[28] G. Ju, X. Qi, H. Ma, and C. Yan, "Feature-based phase retrieval wavefront sensing approach using machine learning," *Opt. Exp.*, vol. 26, no. 24, pp. 31767–31783, 2018.

[29] W. Fei, Y. Bian, H. Wang, L. Meng, and G. Situ, "Phase imaging with an untrained neural network," *Light: Sci. Appl.*, vol. 9, 2020, Art. no. 77.

[30] J. W. Goodman, *Introduction to Fourier Optics.*, Greenwood Village, CO, USA: Roberts and Company Publishers, 2005.

[31] J. Li, Z. Peng, and Y. Fu, "Diffraction transfer function and its calculation of classic diffraction formula," *Opt. Commun.*, vol. 280, no. 2, pp. 243–248, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0030401807008504

[32] J. Weng, J. Zhong, and C. Hu, "Digital reconstruction based on angular spectrum diffraction with the ridge of wavelet transform in holographic phase-contrast microscopy," *Opt. Exp.*, vol. 16, no. 26, pp. 21971–21981, Dec. 2008. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-16-26-21971

[33] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, vol. 378, pp. 686–707, 2019.

[34] G. E. Karniadakis et al., "Physics-informed machine learning," *Nature Rev. Phys.*, vol. 3, no. 6, pp. 422–440, 2021.

[35] L.-G. Pang et al., "An equation-of-state-meter of quantum chromodynamics transition from deep learning," *Nature Commun.*, vol. 9, no. 1, 2018, Art. no. 210.

[36] S. Shi, K. Zhou, J. Zhao, S. Mukherjee, and P. Zhuang, "Heavy quark potential in the quark-gluon plasma: Deep neural network meets lattice quantum chromodynamics," *Phys. Rev. D*, vol. 105, no. 1, 2022, Art. no. 014017.

[37] L. Wang, S. Shi, and K. Zhou, "Reconstructing spectral functions via automatic differentiation," *Phys. Rev. D*, vol. 106, no. 5, 2022, Art. no. L051502.

[38] M. Xiang, L. Wang, H. Yuan, K. Zhou, and H. G. Roskos, "Phase retrieval for Fourier THz imaging with physics-informed deep learning," in *Proc. 47th Int. Conf. Infrared, Millimeter Terahertz Waves*, 2022, pp. 1–2.

[39] L. Alzubaidi et al., "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, pp. 1–74, 2021.

[40] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.

[41] D. Gabor, "Microscopy by reconstructed wave-fronts," *Proc. Roy. Soc. London. Ser. A. Math. Phys. Sci.*, vol. 197, no. 1051, pp. 454–487, 1949.

[42] K. Ikamas et al., "Broadband terahertz power detectors based on 90-nm silicon CMOS transistors with flat responsivity up to 2.2 THz," *IEEE Electron Device Lett.*, vol. 39, pp. 1413–1416, Sep. 2018.

[43] H. Yuan et al., "Dynamic-range enhancement of heterodyne THz imaging by the use of a soft paraffin-wax substrate lens on the detector," in *Proc. Photon. Electromagnetics Res. Symp. - Fall*, 2019, pp. 2607–2611.

[44] H. Yuan et al., "Fourier imaging based on sub-harmonic detection at 600 GHz," in *Proc. 15th Int. Workshop Mobile Terahertz Syst.*, 2022, pp. 1–1.

[45] H.-E. Hwang, H. T. Chang, and W.-N. Lie, "Fast double-phase retrieval in Fresnel domain using modified Gerchberg-Saxton algorithm for lensless optical security systems," *Opt. Exp.*, vol. 17, no. 16, pp. 13700–13710, Aug. 2009. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-17-16-13700

[46] Y. Wan, C. Liu, T. Ma, Y. Qin, and S. Lv, "Incoherent coded aperture correlation holographic imaging with fast adaptive and noise-suppressed reconstruction," *Opt. Exp.*, vol. 29, no. 6, pp. 8064–8075, Mar. 2021. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-29-6-8064