Progress in Quantum Electronics 109 (2026) 100638

Contents lists available at ScienceDirect

Progress in Quantum Electronics

journal homepage: www.elsevier.com/locate/pqe

Invited review

Advanced deep learning for digital holography: A review

Yunping Zhang [a,b], Ni Chen [b], Yanmin Zhu [b], Jingyan Chen [b], Chutian Wang [b], Tianjiao Zeng [c], Edmund Y. Lam [b,*]

[a] Department of Applied Data Science, Hong Kong Shue Yan University, North Point, Hong Kong, 999077, China
[b] Department of Electrical and Computer Engineering, The University of Hong Kong, Pokfulam, Hong Kong, 999077, China
[c] School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu, 611731, China

ARTICLE INFO

Keywords:
Digital holography
Deep learning
Transfer learning
Physics-informed neural networks
Computational imaging

ABSTRACT

Deep learning (DL) has become an important tool in digital holography (DH), enabling data-driven learning of complex inverse mappings directly from holographic measurements. Early DL-assisted DH demonstrated the feasibility of this approach but often relied on large labeled datasets and lacked explicit physical interpretability. In recent years, advances in DL methodologies have begun to address these limitations. This review examines three major developments that are reshaping the field: innovative training strategies for data-efficient learning, the integration of physical priors for improved interpretability and physical consistency, and advanced network architectures capable of extracting richer representations from holographic data. By organizing recent progress from this methodological perspective, we provide an updated overview of how modern DL techniques are expanding the capabilities of DH. We conclude with a forward-looking perspective, identifying promising but underexplored DL directions expected to drive significant future advancements in DH applications.

Contents

* Corresponding author.
E-mail address: elam@eee.hku.hk (E.Y. Lam).

https://doi.org/10.1016/j.pquantelec.2026.100638

Available online 14 July 2026

1. Introduction

Digital holography (DH) is a powerful imaging modality capable of encoding comprehensive three-dimensional (3D) scene information into interference fringes under coherent illumination [1]. As illustrated in Fig. 1, depending on the arrangement of the object and reference waves, DH systems are commonly categorized into on-axis, off-axis, inline, and lensless configurations [1,2]. Although each setup offers different trade-offs among measurement simplicity, reconstruction complexity, and image quality, they share a fundamental physical principle: the recorded intensity pattern (hologram) $I(a,b)$ results from the interference between the object wavefront $O(a,b)$ and a reference background plane wave $R(a,b)$, governed by

$$I(a,b) = |O(a,b) + R(a,b)|^2 = |O(a,b)|^2 + |R(a,b)|^2 + R^*(a,b)O(a,b) + R(a,b)O^*(a,b), \tag{1}$$

where (a,b) represents the spatial coordinates at the hologram plane, and superscript $(\cdot)^*$ denotes the complex conjugation.

Following acquisition, computational algorithms decode the object information, typically retrieving the complex optical field (amplitude and phase). Since the phase component encodes optical path length and refractive index variations, DH enables label-free, quantitative imaging and metrology. This capability has found broad applications across scales, ranging from cell morphology and microfluidics [3,4] to particle characterization and industrial inspection [5,6]. However, traditional reconstruction methods, such as Fresnel diffraction or angular spectrum propagation, face inherent limitations. They rely heavily on precise knowledge of physical parameters and often suffer from artifacts like the twin-image problem and phase wrapping, which degrade image quality and hinder real-time deployment [7–9]. Although many of these challenges can be mitigated through established techniques, such as off-axis recording [2], phase-shifting holography [10,11], phase-unwrapping algorithms [12,13], and iterative reconstruction methods [14], these solutions often require additional measurements, a precise calibration process, increased hardware complexity, or computationally intensive optimization procedures. Consequently, maintaining reconstruction quality, robustness, and real-time performance under practical imaging conditions remains an active research challenge.

The introduction of deep learning (DL) techniques into DH represents a paradigm shift aimed at overcoming these fundamental challenges by learning complex mappings directly from measurement data. As detailed in previous overviews [7,15,16], early explorations in applying DL to DH have predominantly adopted supervised, data-driven approaches. Fig. 1 sketches a typical DL-assisted DH workflow: an object is imaged by an optical system to form a hologram I, which is processed by a neural network $f_\theta(\cdot)$ to produce an estimate $\hat{y}$ of the target quantity. During the training stage, the network parameters θ are iteratively optimized by minimizing loss functions that quantify discrepancies between network predictions $\hat{y}$ and ground truth labels y^{GT}. These labels may come from predetermined patterns with known characteristics [17], outputs of classical physics-based reconstruction methods [18–20], or other imaging modalities such as bright-field microscopy [21]. This end-to-end learning approach allows the network to discover complex transformations directly from holographic measurements, without explicit analytical modeling of the underlying wave propagation.

Upon completion of training, the resultant network can efficiently address various inverse problems in DH with a single forward pass, offering solutions for previously unseen input data with superior speed and often improved accuracy compared to conventional methods. This framework has been successfully applied across all common DH configurations for a broad spectrum of tasks, including amplitude and phase retrieval [4,17], object detection and identification [22,23], and fluidic flow characterization [24,25].

Despite these significant advancements, early DL-assisted DH methods face several important challenges. A primary bottleneck is the requirement for large amounts of registered training data. Acquiring sufficiently diverse holograms I paired with accurate ground truth labels y^{GT} demands considerable experimental resources. Moreover, the lack of explicit physical constraints makes many networks operate as "black boxes": they learn statistical mappings that may suppress noise and artifacts but do not necessarily enforce underlying physical imaging principles. This interpretability gap can impede adoption in domains requiring rigorous validation, particularly in biomedical applications where diagnostic decisions demand transparent reasoning [26]. Furthermore, holograms are inherently information-rich measurements that encode complex phase and amplitude distributions on intricate interference patterns. Effectively exploiting this information richness necessitates deeper and more sophisticated network architectures. These challenges have motivated continuous research efforts in DL techniques for broadening their effective application.

Differing from the previous reviews that primarily focused on the feasibility and early successes of DL-assisted DH [7,15,16], this review adopts a thematic perspective centered on recent advances in DL methodologies aimed at addressing the challenges outlined above. Rather than providing an exhaustive survey of all DL-related studies in DH, we focus on representative methodological developments and emerging trends. As summarized in Fig. 1, three tightly connected directions are reviewed in detail: (i) innovative training strategies designed to overcome data scarcity; (ii) the integration of physical priors to enhance interpretability, peaking at the paradigm of differentiable imaging; and (iii) advanced network architectures tailored for the rich feature content of holographic data. By synthesizing these technical advancements, we aim to provide a clear roadmap of the evolving DL toolkit available to the holography community and conclude with a forward-looking perspective on promising future research directions. It should be noted that the scope of this review is primarily focused on DL-assisted digital holographic imaging and related applications. Topics such as computer-generated holography (CGH) and holographic display systems are beyond the scope of the present review.

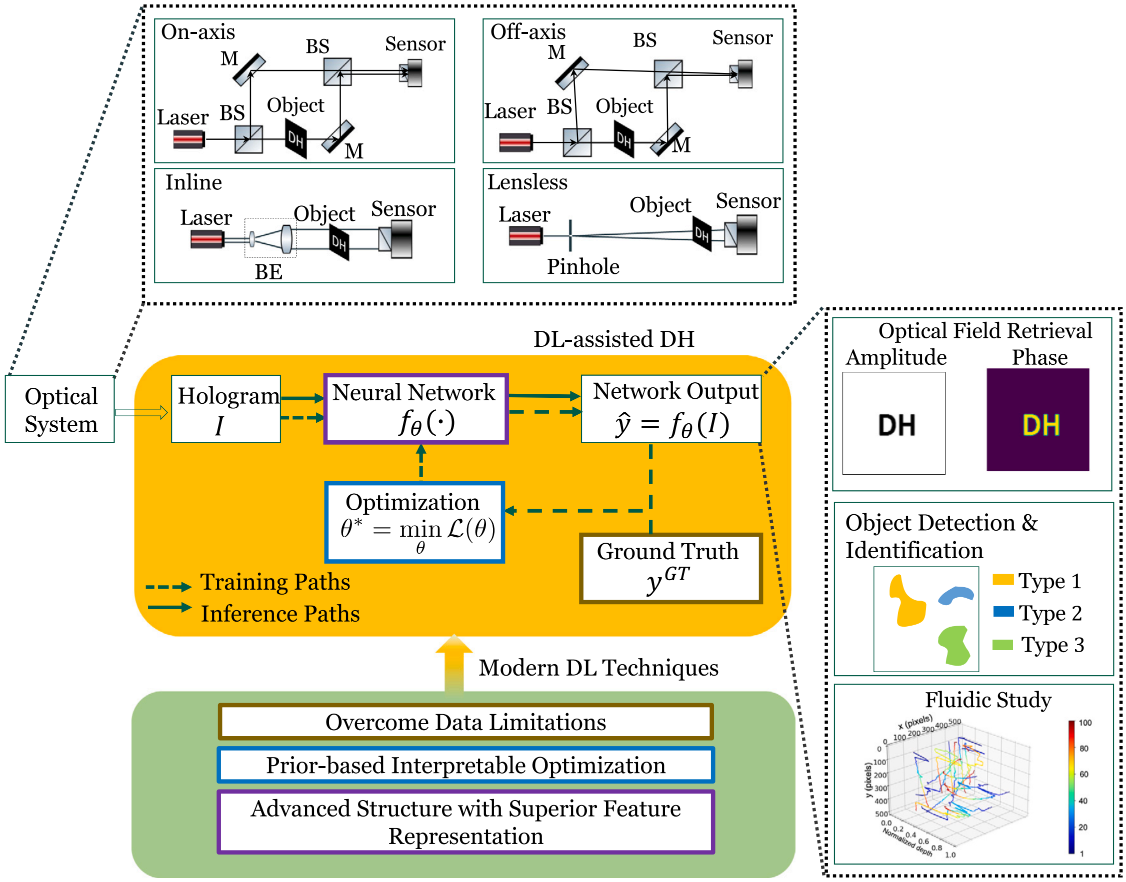

Fig. 1. Overview of DL-assisted digital holography (DH). The figure summarizes representative hologram acquisition configurations, the general DL-assisted DH workflow, major application scenarios, and the methodological developments covered in this review, including strategies for overcoming data limitations, incorporating physical priors for interpretability, and designing advanced network architectures. BS: Beam Splitter; BE: Beam Expander; M: Mirror.

2. Recent advancements in DL methodologies for DH

2.1. Addressing data scarcity: Innovative training strategies

As outlined in the Introduction, a central challenge for applying DL in DH is the limited availability of sufficiently annotated datasets, especially for early end-to-end supervised schemes. These supervised methods rely on paired holographic measurements and their ground truth target labels. The ground truths are typically obtained from predetermined patterns with known characteristics (e.g., patterns displayed by spatial light modulators) [27], traditional reconstruction methods such as multi-height phase retrieval [18,28–30], or off-axis setups with spectrum filtering [31–33].

To alleviate this dependence on large, precisely matched datasets, recent works have developed innovative training strategies and generative models for synthetic data. Among these, Deep Transfer Learning (DTL) has emerged as a powerful paradigm, enabling knowledge transfer from data-rich source domains to data-scarce holographic targets. Based on knowledge transfer mechanisms, DTL methodologies in DH can be broadly categorized into two types that have demonstrated practical utility: adversarial-based and network-based approaches [34,35].

2.1.1. Adversarial-based DTL

The adversarial-based DTL employs adversarial technology to discover transferable representations that are effective for the primary learning task while being indistinguishable between the source and target domains. The underlying principle is that an ideal representation should be discriminative for the task (e.g., reconstruction) but indiscriminate across domains (e.g., synthetic vs. real holograms). This approach is particularly effective for training with unpaired datasets. For instance, Yin et al. [27] and Sun et al. [30] effectively learned the mapping function between unpaired holograms and phase/amplitude images. Although differing in specific network structures, both methods utilize the CycleGAN architecture, incorporating adversarial loss and cycle consistency loss [37]. As shown in Fig. 2(a), in the standard CycleGAN framework, two generators and two discriminators are trained jointly.

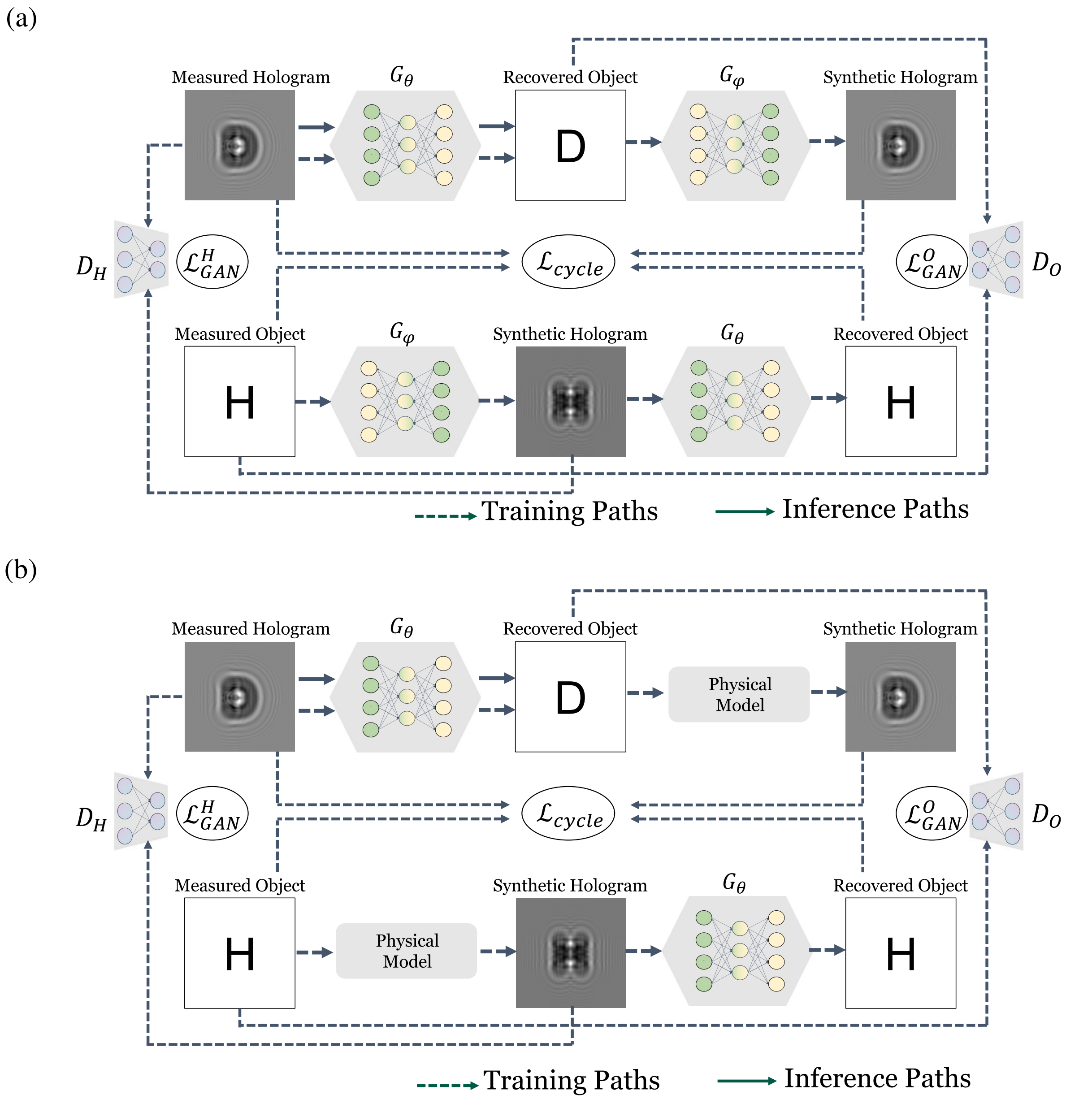

Fig. 2. Examples of DTL-based DH reconstruction methods. (a) Standard CycleGAN [27,30] consists of two sets of generators and discriminators. (b) Physics-informed CycleGAN [36], where one generator in standard CycleGAN is replaced with the fixed parameterized physical forward model. Dashed arrows denote the training-related paths, while solid arrows indicate the inference-related process. During training, the adversarial and cycle-consistency losses are jointly back-propagated to optimize the network components. For clarity, the complete gradient propagation paths associated with the loss functions are not explicitly shown in the diagram.

Generator G_θ maps phase/amplitude images to holograms for reconstruction, while the other generator G_ϕ generates synthetic holograms from phase/amplitude images. Two discriminators, D_H and D_O, are responsible for distinguishing between measured and synthetic signals in the holographic and object domains, respectively. The adversarial losses ($\mathcal{L}^H_{GAN}$ and $\mathcal{L}^O_{GAN}$) from the two GAN networks force the generators to produce outputs that are indistinguishable from real samples in both domains. At the same time, the cyclic consistency loss ($\mathcal{L}_{cycle}$) enforces consistency between the original samples and those reconstructed after a round-trip domain translation. In this way, CycleGAN can be trained using unpaired datasets drawn independently from the holographic and object domains. This effectively relaxes the requirement for a strict one-to-one correspondence between holograms and their labels, alleviating the burden of data collection. In Fig. 2(b), Lee et al. [36] replaced one of the generators in the standard CycleGAN

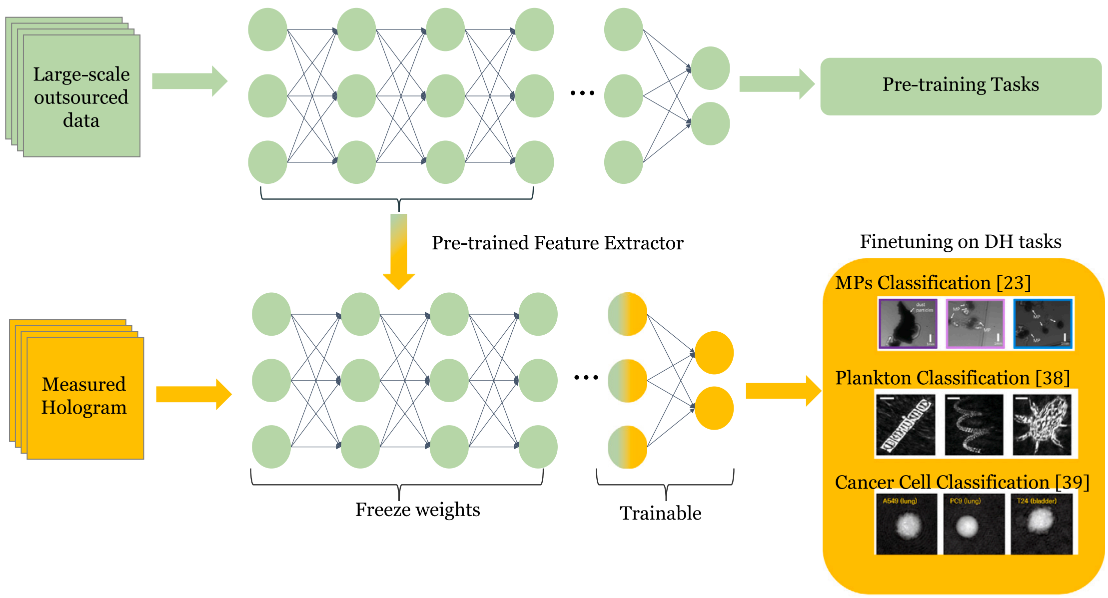

Fig. 3. Network-based DTL for DH. A network pre-trained on large-scale outsourced data provides a transferable feature extractor, which is adapted to hologram inputs through finetuning with partially frozen layers. This approach enables effective DH tasks with limited labeled data, including MPs [23], plankton (adapted with permission from Ref. [38] under terms of the CC-BY license 4.0), and cancer cell classification (adapted with permission from Ref. [39] under terms of the CC-BY license 4.0).

framework with the fixed parameterized physical forward model. It enables the network to adapt to deterministic perturbations in holographic configurations, such as changes in hologram distances, pixel size, and illumination wavelength [36].

2.1.2. Network-based DTL

The network-based DTL involves reusing knowledge encapsulated within a pre-trained network. Fig. 3 demonstrates an overview of the Network-based DTL framework for DH. The core idea is that the initial layers of a deep network, often trained on a large-scale dataset like ImageNet, act as versatile feature extractors that capture generic hierarchical features (e.g., edges, textures, shapes). This pre-trained knowledge can be transferred to a new task by freezing these layers and training only the final layers, or by fine-tuning the entire network on the new, smaller dataset. To compensate for the small size and imbalanced distribution of holographic datasets, several studies have leveraged this approach. Zhu et al. [23] and MacNeil et al. [38] leveraged pre-trained deep neural network (DNN) on large, publicly available datasets like ImageNet to extract generic features, providing a robust baseline for feature recognition from holograms. The networks were then fine-tuned to achieve challenging classifications of small particles in aquatic environments: Zhu et al. [23] applied them on microplastics detection, while MacNeil et al. [38] focused on plankton. By employing DTL, both achieved classification accuracy comparable to supervised learning methods but with fewer labeled data and significant reductions in training time. Moreover, the reliance on large labeled datasets of pre-trained network can be further alleviated through the integration of self-supervised learning strategies, which exploit abundant unlabeled holographic or related data to learn robust representations without manual annotation. For example, in classifying elliptical cancer cells, Rehman et al. [39,40] adopted a network-based DTL approach. However, instead of using a network pre-trained on labeled holographic datasets, they first pre-trained the feature extractor (ResNet-50) on unlabeled data using a combination of self-supervised learning frameworks, such as contrastive learning, SwAV framework [41], and Barlow Twins [42]. This combined approach significantly reduced the reliance on large labeled datasets while achieving classification accuracy comparable to that of supervised learning methods.

It is important to note that these DTL categories are not mutually exclusive. In practice, hybrid approaches that combine multiple techniques can be potentially employed to maximize performance and address the challenges of data scarcity in DH [43].

2.2. Enhance interpretability: Physical prior-based optimization

Another significant challenge is the “black-box” nature of DL networks, which makes it difficult to understand their decision-making processes. This lack of transparency poses significant barriers to building trust and promoting adoption, particularly in critical domains such as clinical and scientific applications. To address this, domain-specific knowledge, such as fundamental physics principles, can be directly incorporated into DL networks in the form of physical priors. We classify those physical priors into two categories: strong priors and weak priors. As summarized in Fig. 4, these priors can be incorporated into DL frameworks through three main components: data, network, and optimization. In the following, we first examine **strong priors**,

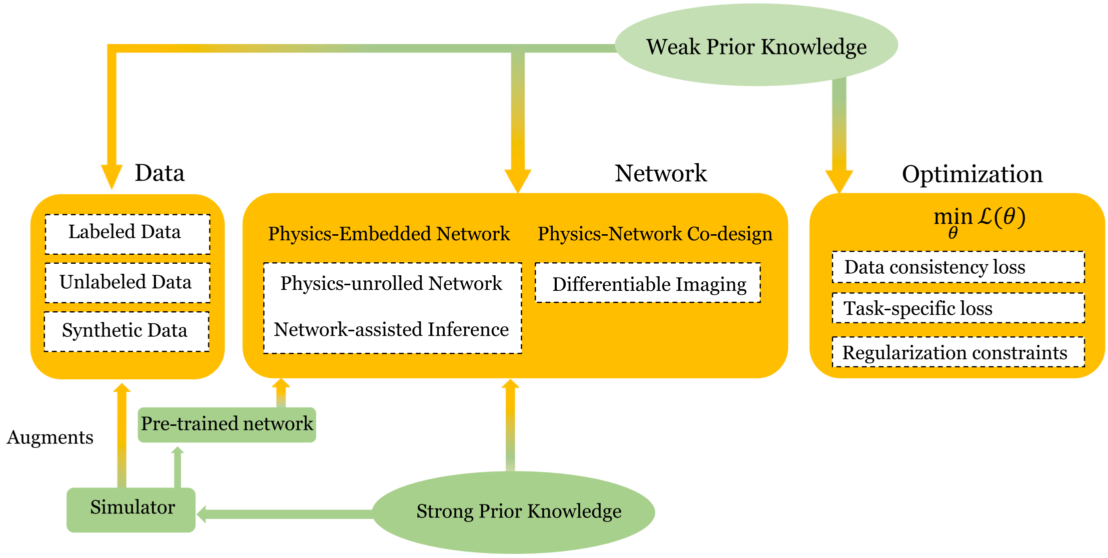

Fig. 4. Overview of physical-prior integration in DL-assisted digital holography (DH). Physical knowledge can be incorporated as either strong priors or weak priors through three major components of the deep learning pipeline: data, network, and optimization.

which explicitly incorporate physical models into the learning process through physics-driven data synthesis, physics-embedded network architectures, and physics-network co-design strategies. We then discuss **weak priors**, which introduce broader forms of physical guidance through data and label engineering, network architecture design, and optimization objectives.

2.2.1. Strong priors

Strong priors refer to domain knowledge derived from explicit physical models that accurately describe the imaging process. In holography, these include wave propagation models and sensor noise statistics [1]. As summarized in Fig. 4, strong priors are primarily incorporated through data generation and network design. At the data level, physical models can be used to generate physics-compliant synthetic datasets for data augmentation and pre-training. At the network level, physical knowledge can either be embedded directly into network architectures through physics-embedded networks or incorporated through physics-network co-design strategies. The subsequent discussion is organized according to this taxonomy.

2.2.1.1. Physics-driven data synthesis. When experimental data is scarce, synthetic datasets generated from physical forward models provide physically consistent training exemplars. When the dataset is sufficiently large and the neural network is expressive and properly trained, the network can approximate the behavior of the physical laws that govern the data generation process.

For example, simulated datasets can be synthesized using physical forward models. Since the fundamental cause of unwanted conjugate images in holographic reconstructions stems from the missing phase in intensity-only holograms, Nagahama et al. [44] trained the network using synthetic complex-valued holograms based on wave propagation to add phase information to holograms. They generate paired training samples by simulating inline holography system with random object masks combined with textures. To handle phase discontinuities, the synthetic complex amplitudes are encoded into three-channel inputs (amplitude, cosine-phase, sine-phase). The network, trained exclusively on this synthetic dataset, successfully generalizes to real experimental holograms, achieving increased PSNR in 8.37 dB and 9.06 dB PSNR gain in amplitude and phase reconstruction, respectively.

Similarly, Huang et al. [45] generated synthetic holograms from artificial random images with no resemblance to real-world samples using free-space wave propagation via angular spectrum methods. It then enforced physics-consistency loss between input holograms and those predicted from the network's output complex fields. After training, without any exposure to real samples or experimental data, it achieves zero-shot generalization to experimental biological holograms (e.g., human tissues, Pap smears), reconstructing high-fidelity phase and amplitude images while inherently learning wave-equation compatibility [45]. This approach also demonstrates resilience to physical perturbations (e.g., unknown axial shifts, pixel-size changes, illumination wavelengths) by encoding physics directly into the network through synthetic dataset [45].

2.2.1.2. Physics-embedded network.

Directly embedding physical equations into the computational pipeline creates hybrid schemes with inherent interpretability [46, 47]. Within this framework, two principal strategies have emerged: **physics-unrolled networks**, where physical models shape the network architecture itself, and **network-assisted inference**, where DNNs are integrated into physics-based iterative solvers as learned update rules.

Physics-unrolled networks use physical knowledge to design specialized network architectures that perform the entire inference process. Rather than employing off-the-shelf DNNs, these architectures are tailored to the problem by mimicking model-based

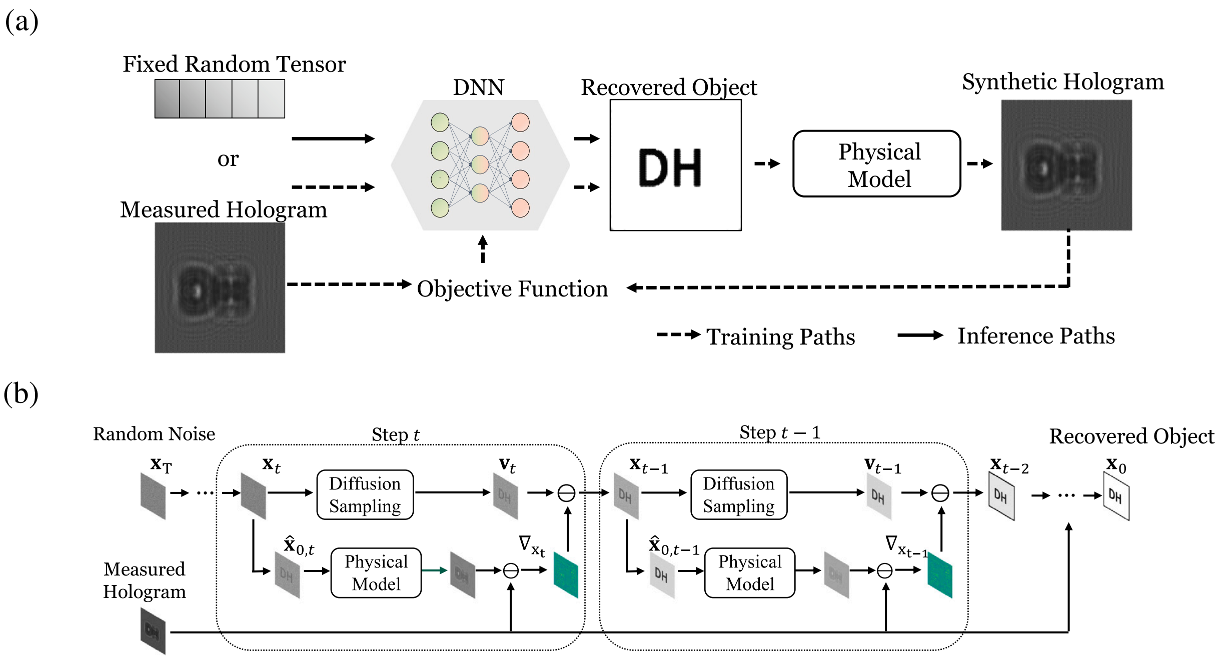

Fig. 5. Examples of DIP-based iterative reconstruction methods. (a) Standard framework of DIP-based holographic reconstruction. A fixed random tensor [53,54] or the measured hologram [55,56] is used as the network input. The reconstructed object generated by the DNN is passed through a physical forward model to synthesize a hologram, and the network parameters are iteratively optimized by minimizing an objective function that promotes agreement between the synthetic and measured holograms. (b) Physics-aware diffusion reconstruction [57]. A pretrained diffusion model is employed as a generative prior, while the holographic imaging physics is incorporated into the iterative sampling process through a data-consistency constraint to guide the reconstruction.

algorithms. The process involves identifying a relevant algorithm or flowchart structure for the problem, breaking it into distinct building blocks, and replacing each block with a dedicated sub-network designed to perform the specific task or computation of that block. These blocks can then be trained either end-to-end using a joint objective function or sequentially by optimizing each block. If the algorithm can be divided into sequential building blocks, where each block corresponds to a single layer in the DNN, the method is called deep unfolding [48] or deep unrolling [49–51].

For example, Chen et al. [25] proposed MB-HoloNet, which uses algorithm unrolling to embed strong physical knowledge directly into the network. By structuring each stage to correspond to an iteration of a physics-based solver, it incorporates the exact physical forward model as a non-trainable layer. This ensures each step of the reconstruction process is constrained by the laws of propagation physics, rather than relying on the network to learn them from scratch. Zhang et al. [52] introduced PSHoloNet, a physics-informed DL framework for holographic reconstruction in photon-starved conditions. The network's architecture is designed by unrolling the Alternating Direction Method of Multipliers (ADMM), embedding a strong physics prior directly into the network by transforming the algorithm's update steps into distinct layers. Specifically, the network incorporates the physics of wave propagation through a fixed layer that implements the holographic diffraction kernel, and it models the unique Poisson noise statistics of the quanta image sensor in another dedicated update layer. While these parts of the network are based on the explicit physical model, the regularization step of the algorithm is replaced by a trainable DNN, allowing the system to learn complex, data-driven priors. This hybrid design enables robust wavefront reconstruction from extremely noisy, single-photon level measurements.

Network-assisted inference, on the other hand, integrates DNNs into traditional model-based methods, where the inference process remains rooted in iterative algorithms, but DNNs are used to enhance, assist, or replace specific intermediate computations. A notable example is the Deep Image Prior (DIP) concept, or more broadly, untrained neural network priors, which enable unsupervised reconstruction without requiring pre-training on large datasets. The fundamental idea of DIPs is that the architecture of a randomly initialized DNN inherently encodes sufficient prior information to capture the statistical regularities of natural objects. This prior is integrated into traditional optimization frameworks for solving holographic inverse problems, such as reconstructing the complex amplitude (phase and amplitude) from intensity-only measurements.

Several studies have demonstrated the feasibility of DIP-based approaches for holographic imaging. Specifically, DeepDIH [53] and PhysenNet [55] showcase the feasibility of standard DIP-based approaches in holographic imaging. As shown in Fig. 5(a), a fixed random tensor [53,54] or the measured hologram [55,56] is used as the network input. The reconstructed object generated by the DNN is numerically forward-propagated through the known physical model of holographic system to synthesize a hologram. Then the network's weights and biases are iteratively adjusted by minimizing the objective function, which calculates the difference (e.g., mean squared error) between the simulated hologram and the actual measured input. As a result, DIP-based methods operate in an unsupervised manner, eliminating the need for pre-training on extensive labeled datasets.

To further improve robustness and mitigate overfitting, subsequent works have incorporated additional regularization strategies and learned priors. Niknam et al. [54] incorporated explicit handcrafted regularization terms, such as bound constraints, weight decay, and periodic random perturbations, to prevent overfitting and improve robustness against noise. Alternatively, Galande et al. [56] addressed overfitting to interference-related noise by integrating an explicit regularization-by-denoising term into the optimization process via the ADMM. More recently, Zhang et al. [57] proposed a physics-aware diffusion model (PadDH) that integrates pretrained generative priors (diffusion models) into the reconstruction process. As demonstrated in Fig. 5(b), during each diffusion sampling step, the current object estimation is propagated through the holographic forward model to generate a synthetic hologram. The discrepancy between the synthetic and measured hologram is then used to compute a data-consistency gradient, which serves as a physics-aware correction term for the diffusion sampling process. By combining the learned generative prior embedded in the pretrained diffusion model [58] with this physics-based correction step, PadDH [57] guides the sampling process towards solutions that are both physically consistent with the measurements and compatible with the learned image prior. As a result, PadDH [57] exhibits enhanced robustness under noisy measurement conditions.

2.2.1.3. Physics-network co-design.

Differentiable imaging (DI) is a computational imaging paradigm in which the entire physical optical system — from illumination and wave propagation to sensing and image formation — is mapped onto a differentiable computational graph [59,60]. By expressing each stage of the optical pipeline as a differentiable operator, gradients can be back-propagated through the complete system, enabling gradient-based joint optimization of both the reconstruction target and the underlying physical system parameters. Unlike physics-embedded networks that incorporate selected physical equations into a fixed network architecture, DI is not a neural-network architecture itself; rather, it serves as a physics-network co-design framework in which physical models and neural networks are jointly optimized, maintaining physical consistency throughout the entire process [59].

In the context of DH, this requires constructing differentiable wave propagation operators that accurately simulate the physics of light diffraction and interference [61]. Consider a holographic imaging system where the measured hologram I is related to the object field u_0 through a forward operator $\mathcal{F}$

$$I = |\mathcal{F}(u_0; \theta_{\mathrm{sys}})|^2 + n, \tag{2}$$

where θ_{sys} encapsulates system parameters including propagation distance, wavelength, pixel size, and illumination characteristics, while n represents measurement noise. Traditional reconstruction approaches treat θ_{sys} as fixed quantities determined from nominal specifications or through separate calibration procedures. In contrast, differentiable imaging treats both the object field u_0 and system parameters $\hat{\theta}_{\mathrm{sys}}$ as jointly optimizable variables, minimizing a composite loss function

$$\mathcal{L} = \mathcal{L}_{\mathrm{data}}\left(I, |\mathcal{F}(u_0, \hat{\theta}_{\mathrm{sys}})|^2\right) + \sum_{i=1}^{N} \mathcal{R}_i(u_0), \tag{3}$$

where $\mathcal{L}_{\mathrm{data}}$ quantifies data fidelity between measured and forward-modeled holograms, and $\mathcal{R}_i$ represents physics-informed or data-driven regularization terms that encode prior knowledge about the object field. Typical examples include total variation (TV) regularization for promoting smoothness [62], sparsity constraints for suppressing noise and artifacts [63], and support constraints that restrict the solution to physically feasible regions [64].

The technical foundation enabling this approach is automatic differentiation through the forward model $\mathcal{F}$. Using modern automatic differentiation frameworks such as PyTorch, TensorFlow, and JAX allows efficient computation of gradients of the loss function with respect to all optimization variables via the chain rule, including both reconstruction and system parameters [25]. When hologram formation operators are implemented as differentiable operations within these frameworks, backpropagation seamlessly flows through the entire computational pipeline: from measured data through the physical forward model, into the reconstruction space, and back through parameter updates [61]. Such differentiability transforms what was traditionally a multi-stage process involving separate calibration, forward modeling, and reconstruction into a unified optimization problem.

The differentiable imaging framework has demonstrated transformative capabilities across diverse holographic imaging applications [59,60,65].

Calibration-free reconstruction: When system parameters are unknown or inaccessible, differentiable holography treats key factors, including the object–sensor distance, illumination wavefront characteristics, and optical aberrations, as optimization variables instead of fixed, pre-calibrated parameters [61]. This approach maintains experimental simplicity by eliminating the need for extensive calibration procedures that are often impractical in dynamic or field-deployed systems, while simultaneously ensuring robust adaptability across varying experimental configurations. For instance, in lensless holographic imaging, an integrated pixel-super-resolution (PSR) approach extends this paradigm by jointly optimizing sub-pixel sensor positions and defocus distances as learnable parameters [66], thereby achieving resolution beyond the physical sensor limitations without requiring precise mechanical alignment.

Physics-aware modeling of complex phenomena: Differentiable imaging excels at handling sophisticated physical interactions that challenge traditional simplified models. In differentiable 3D particle imaging, for example, the framework explicitly addresses multiple scattering effects by incorporating accurate wave interaction models into the forward propagator [67]. Rather than relying on conventional linearized Born approximations that assume weak single scattering, this approach captures the full complexity of object-light interactions through differentiable electromagnetic wave propagation, enabling accurate reconstruction even at high particle concentrations where multiple scattering dominates [25,68]. This capability opens new regimes of measurement previously inaccessible to conventional holographic methods.

2.2.2. Weak priors

Weak priors are less rigid than strong priors but are still rooted in universal physical properties. Unlike strong priors, which explicitly integrate the physical model into the reconstruction process, weak priors encode the general structural characteristics of the imaging problem without requiring a strict mathematical description of the forward model. These priors are useful in scenarios where the physical laws are only partially understood or where domain knowledge exists in the form of broad structural assumptions rather than exact governing equations.

As summarized in Fig. 4, weak priors can be incorporated in different components of the DL framework, including data representation, network architecture design, and optimization objectives.

At the data level, weak priors can be introduced through data and label engineering, where prior knowledge is embedded into the representation of training samples rather than the network architecture itself. In this method, a simplified physical model or heuristic is encoded directly into the training data, particularly the target labels. This strategy guides a standard DL network to produce physically meaningful predictions without altering its core architecture. A prominent example of this technique is the adaptation of YOLO-based object detection methods for 3D particle tracking in holographic imaging. In this context, a simplified physical relationship is established where the axial depth (z) of a particle is proportional to the apparent size of its diffraction pattern in the hologram. The training labels are then engineered such that a bounding box is drawn around the particle, with its center indicating the lateral (x, y) coordinates and its size programmatically set as a function of the axial depth (z). This cleverly encodes three-dimensional spatial information into a format digestible by a 2D object detection network.

Building on this idea, several advanced frameworks have successfully implemented data-engineering strategies to infer volumetric information from 2D holograms. The CATCH framework, for example, integrates a YOLOv5-based localizer network with an estimator network [69,70]. The localizer identifies particle features and produces bounding boxes encoding the (x, y) coordinates and depth (z), while the estimator simultaneously predicts other continuous properties like particle radius and refractive index. Both networks are trained jointly on synthetic data generated from Lorenz-Mie theory, thus integrating a foundational physical model. Expanding on this, the OSNet [24] directly retrieves the volumetric distribution of particles from a single feed-forward process, which eschews the beforehand reconstruction and post-processing operations. Zhang et al. [71] employ the YOLOv7 network to generate precise regions of interest (ROIs). Notably, to address semantic information loss for small particles (less than $40\,\mu\mathrm{m}$), the hologram is partitioned into smaller patches as inputs to the network. The axial position is subsequently determined by identifying the maximum variation in pixel intensity gradients along the axial direction. This is followed by cluster-based segmentation, inspired by the fluctuation features between particles and the background, to recover the precise boundaries of the in-focus particles.

At the network level, weak priors are commonly incorporated through network architectures that reflect general structural properties of holographic data. A common strategy for implementing weak priors is to encode known physical properties directly into network architectures. For instance, networks can be designed to be invariant or equivariant to transformations like translation, rotation, or permutation, which are common in physical systems. A notable example is the convolutional neural network (CNN), which aligns with physical principles of locality and translational equivariance. CNNs process information through localized receptive fields and shared convolution kernels, enabling efficient extraction of spatially correlated features while preserving translational consistency. As a result, CNNs have become the dominant architectural backbone in DL-assisted DH [7,8].

Recurrent neural networks (RNNs) [72] offer another means of embedding weak priors by modeling sequential dependencies. In holographic imaging, RNNs are well-suited for processing holographic data acquired across multiple axial planes, as they can sequentially fuse complementary information from each plane. Their temporal-memory mechanism mirrors the evolution of physical systems over time and allows the network to fuse depth-dependent features into a coherent representation. Huang et al. [73] leveraged this property by feeding a sequence of holograms captured at different axial distances into an RNN, incrementally combining in-focus information to achieve extended depth-of-field reconstruction.

Weak priors can also be incorporated through optimization objectives. Instead of explicitly embedding a physical forward model, prior knowledge may be enforced through regularization terms that encourage desired structural properties, such as sparsity [63], smoothness [62], continuity, or geometric consistency [64]. These objectives act as soft constraints during training, steering the network toward physically plausible solutions.

Overall, DH relies on the principles of light-wave propagation, interference, and diffraction, all of which are governed by well-defined mathematical frameworks, such as the wave equation, Fresnel diffraction, and Fourier optics [1]. These physical principles provide a rich source of prior knowledge that can be incorporated into DL-assisted DH through both strong and weak priors. Such priors bridge the gap between data-driven learning and physical modeling, providing enhanced interpretability.

2.3. Superior feature representation: Advanced network architectures

Beyond training strategies and physical priors, the high information content intrinsic to holograms has also driven innovations in network architecture. Because holograms encode complex amplitude and phase information with high information density across large spatial dimensions, the limitations of traditional CNNs have motivated a shift toward more advanced architectural designs aimed at improving representational capacity and reconstruction performance.

One key direction involves explicitly separating distinct physical quantities within the network architecture. For instance, by extending the output channels, Wang et al. developed Y-Net [32] and Y4-Net [31] for simultaneous complex amplitude reconstruction in single-wavelength and dual-wavelength DH, respectively. In Y-Net, a shared encoder first extracts common features from the input hologram and then branches into two task-specific decoders, as illustrated in Fig. 6(a). One decoder is dedicated to reconstructing the intensity map, and the other is dedicated to the phase map. This design acknowledges that while intensity and

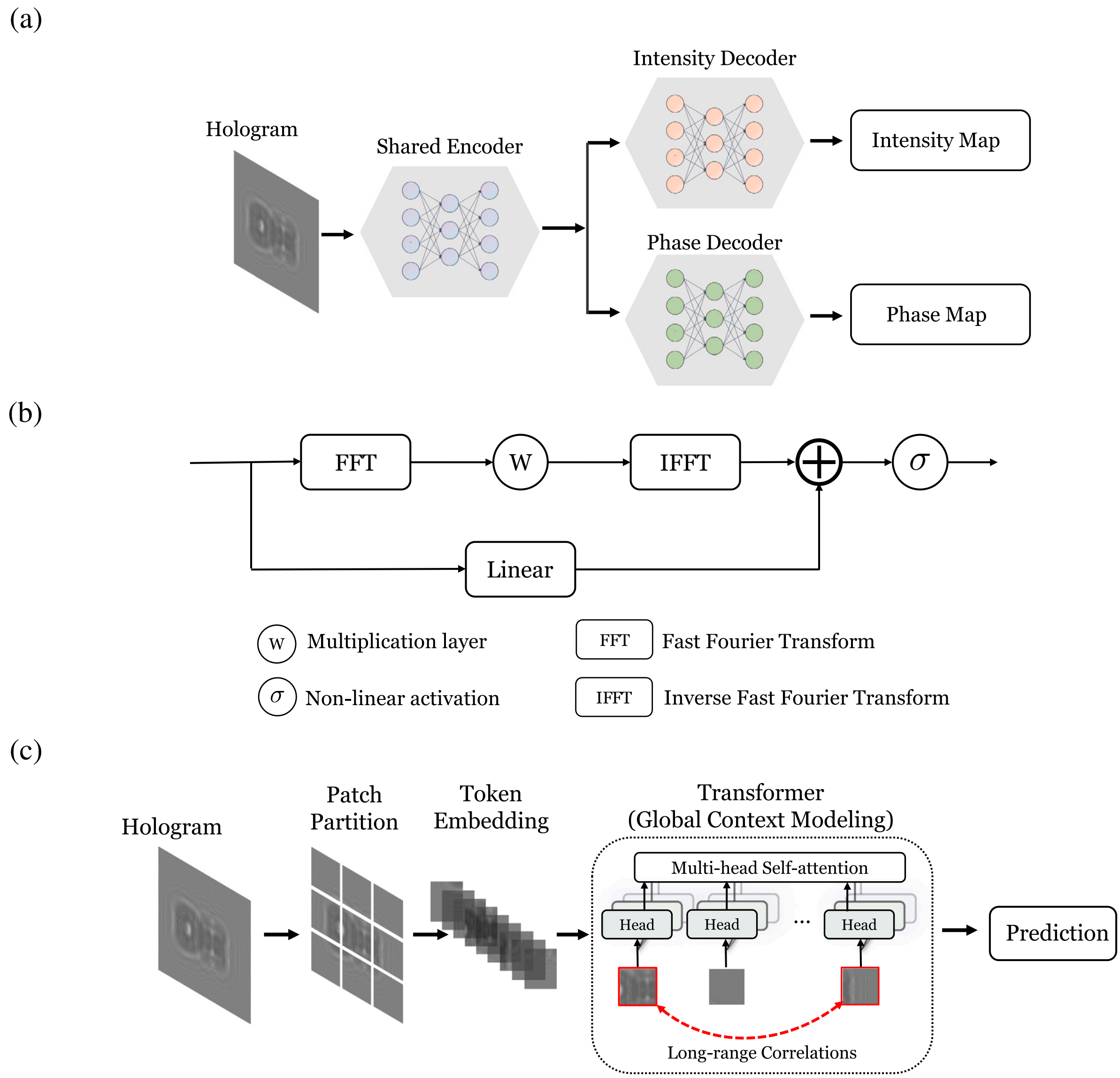

Fig. 6. Conceptual illustration of key components in representative advanced network architectures used in DL-assisted DH. (a) Y-Net [32] and its variants employ a shared encoder to extract common features from the input hologram, followed by task-specific decoders for reconstructing intensity and phase maps separately. (b) Fourier Neural Operator (FNO) [74] combines a spectral transformation (FFT–multiplication–IFFT) with a parallel linear transformation in the spatial domain, followed by feature fusion and non-linear activation. (c) TransFocusNet [75] partitions a hologram into patch tokens and employs multi-head self-attention to capture long-range spatial dependencies and global contextual information across the entire field of view.

phase are derived from the same hologram, they are distinct physical quantities. By giving them separate reconstruction paths, the network can learn specialized features for each task, preventing the learning process for one from interfering with the other and ultimately improving the accuracy for both. Meanwhile, the shared encoder allows information useful to both tasks to be learned jointly, improving feature utilization and training efficiency. Y4-Net [31] provides a direct and intelligent evolution of this concept for the more complex problem of dual-wavelength holography. It retains the Y-Net's core structure but expands the two output paths to each have two channels. This creates a one-to-four network that simultaneously reconstructs complex-valued object information under two wavelengths.

Another significant architectural shift moves from the local, spatially-bound processing of CNNs to global analysis in the frequency domain. To this end, Chen et al. [28,29] introduced spatial Fourier transform modules, based on the Fourier Neural Operator (FNO), to enhance the network's generalization capabilities. The core working principle of FNOs is parameterizing the integral kernel of a neural operator directly in the Fourier domain. Based on the convolution theorem, which states that a convolution in the spatial domain corresponds to a multiplication in the Fourier domain, FNO replaces the convolution operation used in CNNs with a Fourier-based linear transformation to extract implicit features. Specifically, as shown in Fig. 6(b), a FNO consists of a Fourier transform, spectral multiplication with learnable weights, and an inverse Fourier transform, which are combined with a parallel linear transformation in the spatial domain. This design grants the network to process the spatial frequencies of the entire input at once and capture global context more efficiently than standard convolutions. Such a global receptive field is particularly useful for DH because wave propagation and interference are inherently global physical processes.

Originally developed for natural language processing to capture global information and long-range dependencies, Transformers [76–78] have recently been adapted for holographic analysis. Li et al. [75] introduced TransFocusNet, which applies the Vision Transformers (ViTs) [78] architecture to the real-time analysis of crystalline structures using an in-line infrared DH system. Unlike CNNs, which primarily extract information through local convolutional kernels, ViTs partition the input hologram into patches and represent them as token sequences, as illustrated in Fig. 6(c). These tokens are processed by multi-head self-attention modules that enable direct interactions between spatially distant regions of the hologram. By effectively capturing long-range correlations and complex structural information, TransFocusNet [75] facilitates the monitoring of microstructural evolution during crystal growth and the detailed examination of microstructures and interlayer relationships in 2D materials.

Overall, by incorporating task-specific decoupling, global frequency-domain processing, and attention-based mechanisms, advanced network designs achieve superior feature representation of holographic data.

3. Outlook

The rapid evolution of modern DL methodologies is not only improving the performance of existing DH systems but also reshaping how holographic imaging systems are designed and deployed. Looking ahead, emerging trends suggest that DL will play an increasingly central role in expanding DH sensing capabilities under unconventional and extreme acquisition conditions, serving as the computational core of multimodal holographic imaging systems, and enabling the joint optimization of optical hardware and computational algorithms.

3.1. DL-assisted sensing for unconventional and extreme acquisition conditions

DH faces fundamental challenges when deployed under demanding conditions, such as challenging illumination or extreme environments ranging from cryogenic or high-radiation environments to low-photon or non-visible spectral regimes (infrared, terahertz, or X-ray) [52,79–82]. Conventional cameras, with their frame-based exposure, are inherently limited in temporal resolution and dynamic range. As a result, maintaining the calibration, coherence, and phase accuracy required for high-quality holographic reconstruction becomes difficult, as evidenced by early efforts in cryogenic quantum-fluid holography [83] and infrared holography through smoke and flames [84].

To push holography into these unconventional regimes, integration with next-generation imaging sensors has become a key enabling strategy. Event cameras, operating through computational neuromorphic imaging, provide asynchronous, microsecond-scale response with exceptional dynamic range, fundamentally breaking from the limitations of frame-based acquisition [85]. These properties allow blur-free, sub-millisecond holographic field recovery, exemplified by recent NeuroDH demonstrations [86,87]. In parallel, quanta image sensors [88] and single-photon avalanche diode [89] arrays enable holographic capture into photon-starved conditions, while also supporting high-speed temporal resolution unattainable by conventional detectors [52,90]. Together, these technologies represent a growing ecosystem of sensors tailored for extreme or non-standard optical conditions.

However, the diversity of their outputs, including asynchronous spike streams, sparse photon counts, and modality-dependent noise characteristics, creates a second challenge: how to synthesize these disparate measurements into a unified, coherent holographic representation. DL provides a natural computational core for such fusion, offering the ability to integrate asynchronous events, denoise ultra-low-photon measurements, compensate for modality-specific artifacts, and exploit priors learned from data or physical models. DL-assisted reconstructions can effectively convert the hardware's unconventional sampling patterns into high-fidelity holographic fields, enabling performance that would otherwise be unattainable. Therefore, DL is expected to play a critical role in enabling the integration of emerging sensing hardware into future DH systems. The resulting synergy between advanced sensing technologies and DL-based computational frameworks will broaden the operational range of DH, enabling reliable operation under challenging acquisition conditions and extending holographic measurements to scenarios that are difficult to access using conventional imaging hardware.

3.2. DL as the computational core of multi-modality DH

Multi-modality or cross-modality DH has evolved to leverage the strengths of diverse imaging modalities synergistically. Examples include Raman scattering holography [91], speckle holography [92,93], fluorescence holography [94], polarization holography [95,96], holographic optical coherence tomography [97], hyperspectral holography [98,99], and photoacoustic holographic imaging [100]. By incorporating multiple imaging schemes or various optical field modulation methods, multi-modality DH provides richer image features and enhances information complementarity. This fusion enables simultaneous probing of structural, chemical, and physical properties, offering a more comprehensive characterization than any single modality alone.

This complementary nature has already demonstrated clear advantages across a range of applications. For instance, the combination of holographic and polarization features enables simultaneous acquisition of multi-dimensional characteristics of samples, such as material composition, morphology, oxidation level, and refractive index, thus offering a powerful technical tool for particle detection, including microplastic pollutant monitoring [95]. Similarly, Raman scattering holography can capture the chemical fingerprints and morphological information of samples, thereby accurately supporting applications such as 3D localization and tracking of individual nanoparticles in living cells, tissue and cell detection, and anti-counterfeiting [91]. Beyond improving imaging quality, these multi-modal approaches establish a more robust data foundation for downstream analysis and identification tasks.

However, the benefits of multi-modality come with significant computational challenges. The resulting data streams are often massive, heterogeneous, and high-dimensional, making it difficult to fuse cross-modal features, align heterogeneous data, and extract meaningful information in a unified manner. Traditional, hand-engineered pipelines struggle to scale with this complexity, limiting the practical deployment of multi-modal DH systems.

DL emerges as the indispensable computational core to overcome these barriers. Its ability to automatically learn hierarchical features from raw, high-dimensional data makes it uniquely suited for multi-modality fusion, reconstruction, and analysis. Though not yet deeply or widely adopted, pioneering attempts have recently been made in DL-assisted multi-modality DH. For instance, Zhu et al. [101] developed SPLASH (smart polarization and spectroscopic holography), which simultaneously acquires polarization states (angle of polarization, degree of linear polarization), holographic features, and texture information without requiring physical spectroscopic instrumentation. Huang et al. [102,103] developed computational polarized holography to automate microplastic monitoring in scattering environments. Chen et al. [104] integrated dispersion-encoded holography and DL to reconstruct spectral channels from compressed data, enabling particle identification from fused spectral and 3D morphological features. Takahashi et al. [105] merged holography with Raman spectroscopy, using DL-driven fusion to achieve label-free marine particle clustering.

These studies point toward the transformative potential of compact multi-modal DH hardware with DL-based computational frameworks. The inherent capability of DNNs to automatically extract relevant features from complex, high-dimensional data makes them ideally suited for processing the rich information content generated by multi-modality holographic systems. This combination enables end-to-end learning of optimal feature representations that may transcend human-engineered features, potentially discovering subtle physical–chemical correlations that remain obscured in conventional analyses. Furthermore, DL approaches can effectively address challenges in multi-modality data registration, fusion, and interpretation, while simultaneously enhancing system performance through improved reconstruction quality, robust classification in scattering environments, and accelerated processing speeds. Such technological convergence will not only advance environmental monitoring and biomedical imaging but also open new possibilities for intelligent material characterization.

3.3. Differentiable imaging toward scalable and intelligent DH

As summarized earlier in this review, differentiable imaging is a computational imaging paradigm in which the entire physical optical system is mapped onto a differentiable computational graph [59,60]. By expressing each stage of the optical pipeline as a differentiable operator, this approach enables gradient-based joint optimization of optical systems and reconstruction algorithms. However, several fundamental challenges continue to limit its broader applicability. Addressing these challenges while integrating differentiable imaging with emerging AI paradigm will be essential for realizing scalable, efficient, and intelligent holographic systems.

Computational complexity remains a primary bottleneck. Differentiable physics simulations, particularly for full 3D electromagnetic wave propagation, can be orders of magnitude slower than forward passes through purely data-driven DNNs. Achieving real-time performance requires careful trade-offs between physical fidelity and computational efficiency through strategic approximations, GPU acceleration, and hybrid architectures that combine learned components with differentiable physics. Additionally, certain physical phenomena, such as discrete processes, non-differentiable operations (e.g., saturation and clipping), and intrinsically stochastic dynamics, pose fundamental challenges for gradient-based optimization and therefore require specialized approaches, including stochastic gradient estimation and relaxation strategies.

Uncertainty management represents another important dimension where differentiable imaging can make a distinctive contribution. Because the entire physical optical system is expressed as a differentiable computational graph, measurement uncertainty arising from detector noise, system perturbations, or imprecisely known physical parameters can be systematically propagated through the pipeline. Jacobian-based sensitivity analysis, made computationally tractable by automatic differentiation, allows the contribution of each uncertainty source to be quantified at the output; alternatively, Bayesian formulations within the differentiable framework enable posterior distributions over reconstruction targets and system parameters to be estimated jointly [60]. This stands in marked contrast to conventional multi-stage reconstruction pipelines, in which uncertainty is difficult to trace through non-differentiable intermediate steps. Principled uncertainty quantification is especially valuable in biomedical imaging and precision metrology, where calibrated confidence estimates for reconstructed quantities are as important as the reconstructions themselves.

Beyond these technical hurdles, differentiable imaging offers a principled foundation for connecting DH with the broader AI ecosystem. By serving as a differentiable physics engine, it enables the creation of digital twins for holographic systems, providing virtual replicas that support predictive simulation, design exploration, and self-calibration [60]. This physics-aware capability naturally extends to emerging paradigms like foundation models and vision-language models. The structured, physically meaningful 3D representations produced by differentiable pipelines provide a grounded interface for these large-scale models, enabling them to reason about microscopic scenes, transfer learned priors across modalities, and interpret holographic data through natural language.

Looking further ahead, this integration paves the way for world models in embodied AI, where agents utilize holographic perception to interact with the physical world. By embedding differentiable optical models within their internal simulators, future intelligent systems could reason about volumetric environments with the same rigor as physical laws, bridging the gap between computational imaging and autonomous action.

4. Conclusion

DH has undergone a paradigm shift driven by DL, transitioning the field from classical reconstruction algorithms to data-driven inference. Moving beyond early exploration, this review has highlighted an updated generation of DL-assisted DH systems that are more data-efficient, physically interpretable, and architecturally sophisticated.

We have examined how innovative training strategies are effectively overcoming the issue of data scarcity, enabling high-performance reconstruction even when experimental ground truth is limited. Furthermore, by incorporating both strong priors derived from explicit physical models and weak priors reflecting general physical properties across the data, network, and optimization components of the learning pipeline, the "black-box" nature of neural networks is being addressed with enhanced interpretability. Concurrently, the adoption of advanced architectures, such as Fourier Neural Operators and Vision Transformers, has allowed for superior feature representation, capturing global dependencies and complex wavefront information that traditional CNNs often miss.

Looking forward, the future of DL-assisted DH lies in the deep integration of advanced sensing hardware, multi-modality imaging, and scalable intelligent frameworks. This enables DH to operate under increasingly challenging acquisition conditions while extracting richer information from complex scenes. At the same time, rather than treating imaging, reconstruction, and analysis as separate stages, next-generation holographic systems are expected to unify these processes within a single learning-enabled framework. Emerging paradigms such as differentiable imaging are expected to further advance this vision, paving the way for end-to-end optimized holographic imaging pipelines. As these trends continue to mature, ongoing advances in deep learning are expected to further expand the capabilities of DH, enabling more robust reconstruction, richer information extraction, and increasingly intelligent analysis across a broader range of imaging scenarios.

CRediT authorship contribution statement

Yunping Zhang: Writing – original draft, Formal analysis, Data curation. **Ni Chen:** Writing – original draft, Investigation, Formal analysis. **Yanmin Zhu:** Writing – original draft, Investigation, Formal analysis. **Jingyan Chen:** Investigation, Formal analysis. **Chutian Wang:** Investigation, Formal analysis. **Tianjiao Zeng:** Investigation, Formal analysis. **Edmund Y. Lam:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Funding

This work described in this paper was fully supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (GRF: 17201822, RIF: R7003-21).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J.W. Goodman, Introduction to Fourier Optics, fourth ed., W.H.Freeman & Company Ltd, 2017.
[2] U. Schnars, W.P. Jüptner, Digital recording and numerical reconstruction of holograms, Meas. Sci. Technol. 13 (9) (2002) R85–R101.
[3] D. Pirone, R. Mottareale, G. Giugliano, C.D. Vita, G.L. Verde, J. Běhal, C. Arrichiello, P. Muto, I. Kurelac, L. Bagnale, M. Medugno, P. Memmolo, V. Bianco, L. Miccio, P.L. Maffettone, M. Durante, P. Ferraro, M. Pugliese, Holographic imaging flow cytometry for radiation response assessment in lymphocytes, in: Proceedings of the Optical Methods for Inspection, Characterization, and Imaging of Biomaterials VII, SPIE, 2025, 1357123.
[4] G. Giugliano, M. Schiavo, D. Pirone, V. Bianco, P. Memmolo, L. Miccio, P. Ferraro, Single cell analysis by holographic microscopy and flow cytometry: An high-throughput and label-free approach, in: Proceedings of the Biophotonics in Point-of-Care III, SPIE, 2024, 130080F.
[5] D. Pirone, B. Cavina, G. Giugliano, M. Schiavo, L. Miccio, V. Bianco, F. Nanetti, F. Reggiani, G. Gasparre, I. Kurelac, et al., Lightweight CNN efficiently discriminates ovarian cancer cells from a tumor microenvironment via holographic imaging flow cytometry, Biomed. Opt. Express 16 (11) (2025) 4273–4284.
[6] R. Ren, Z. Jia, J. Yang, N. Kasabov, X. Huang, Quasi-noise-free and detail-preserved digital holographic reconstruction, IEEE Access 7 (2019) 52155–52167.
[7] T. Zeng, Y. Zhu, E.Y. Lam, Deep learning for digital holography: A review, Opt. Express 29 (24) (2021) 40572–40593.
[8] Y. Rivenson, Y. Wu, A. Ozcan, Deep learning in holography and coherent imaging, Light: Sci. Appl. 8 (1) (2019) 85.
[9] Z. Ren, Z. Xu, E.Y. Lam, Learning-based nonparametric autofocusing for digital holography, Optica 5 (4) (2018) 337–344.
[10] A. Greenbaum, A. Ozcan, Maskless imaging of dense samples using pixel super-resolution based multi-height lensfree on-chip microscopy, Opt. Express 20 (3) (2012) 3129–3143.
[11] P.S. Huang, S. Zhang, Fast three-step phase-shifting algorithm, Appl. Opt. 45 (21) (2006) 5086–5091.
[12] P. Ferraro, S. De Nicola, A. Finizio, G. Coppola, S. Grilli, C. Magro, G. Pierattini, Compensation of the inherent wave front curvature in digital holographic coherent microscopy for quantitative phase-contrast imaging, Appl. Opt. 42 (11) (2003) 1938–1946.
[13] X. Su, W. Chen, Reliability-guided phase unwrapping algorithm: A review, Opt. Lasers Eng. 42 (3) (2004) 245–261.
[14] L. Rong, Y. Li, S. Liu, W. Xiao, F. Pan, D. Wang, Iterative solution to twin image problem in in-line digital holography, Opt. Lasers Eng. 51 (5) (2013) 553–559.
[15] Y. Rivenson, Y. Wu, A. Ozcan, Deep learning in holography and coherent imaging, Light: Sci. Appl. 8 (1) (2019) 85.
[16] G. Situ, Deep holography, Light: Adv. Manuf. 3 (2) (2022) 278–300.
[17] H. Wang, M. Lyu, G. Situ, eHoloNet: A learning-based end-to-end approach for in-line digital holographic reconstruction, Opt. Express 26 (18) (2018) 22603–22614.

[18] Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, A. Ozcan, Phase recovery and holographic image reconstruction using deep learning in neural networks, Light: Sci. Appl. 7 (2) (2018) 17141–17141.
[19] T. Zeng, H.K.-H. So, E.Y. Lam, RedCap: Residual encoder-decoder capsule network for holographic image reconstruction, Opt. Express 28 (4) (2020) 4876–4887.
[20] Z. Ren, Z. Xu, E.Y. Lam, End-to-end deep learning framework for digital holographic reconstruction, Adv. Photonics 1 (1) (2019) 016004(1–12).
[21] A. Sinha, J. Lee, S. Li, G. Barbastathis, Lensless computational imaging through deep learning, Optica 4 (9) (2017) 1117–1125.
[22] Y. Zhu, C.H. Yeung, E.Y. Lam, Microplastic pollution monitoring with holographic classification and deep learning, J. Phys.: Photonics 3 (2) (2021) 024013(1–12).
[23] Y. Zhu, C.H. Yeung, E.Y. Lam, Digital holographic imaging and classification of microplastics using deep transfer learning, Appl. Opt. 60 (4) (2021) A38–A47.
[24] Y. Zhang, Y. Zhu, E.Y. Lam, Holographic 3D particle reconstruction using a one-stage network, Appl. Opt. 61 (5) (2022) B111–B120.
[25] N. Chen, C. Wang, W. Heidrich, Holographic 3D particle imaging with model-based deep network, IEEE Trans. Comput. Imaging 7 (2021) 288–296.
[26] Z. Salahuddin, H.C. Woodruff, A. Chatterjee, P. Lambin, Transparency of deep neural networks for medical image analysis: A review of interpretability methods, Comput. Biol. Med. 140 (2022) 105111.
[27] D. Yin, Z. Gu, Y. Zhang, F. Gu, S. Nie, J. Ma, C. Yuan, Digital holographic reconstruction based on deep learning framework with unpaired data, IEEE Photonics J. 12 (2) (2019) 1–12.
[28] H. Chen, L. Huang, T. Liu, A. Ozcan, eFIN: Enhanced Fourier imager network for generalizable autofocusing and pixel super-resolution in holographic imaging, IEEE J. Sel. Top. Quantum Electron. 29 (2023) 1–10.
[29] H. Chen, L. Huang, T. Liu, A. Ozcan, Fourier Imager Network (FIN): A deep neural network for hologram reconstruction with superior external generalization, Light: Sci. Appl. 11 (1) (2022) 254.
[30] X. Sun, C. Xiong, Z. Ren, J. Di, J. Zhao, Unpaired learning for digital holographic reconstruction and generation, Opt. Express 33 (2) (2025) 3414–3428.
[31] K. Wang, Q. Kemao, J. Di, J. Zhao, Y4-Net: A deep learning solution to one-shot dual-wavelength digital holographic reconstruction, Opt. Lett. 45 (15) (2020) 4220–4223.
[32] K. Wang, J. Dou, Q. Kemao, J. Di, J. Zhao, Y-Net: A one-to-two deep learning framework for digital holographic reconstruction, Opt. Lett. 44 (19) (2019) 4765–4768.
[33] D. Pirone, D. Sirico, L. Miccio, V. Bianco, M. Mugnano, P. Ferraro, P. Memmolo, Speeding up reconstruction of 3D tomograms in holographic flow cytometry via deep learning, Lab A Chip 22 (4) (2022) 793–804.
[34] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, C. Liu, A survey on deep transfer learning, in: Proceedings of the Artificial Neural Networks and Machine Learning, Springer International Publishing, 2018, pp. 270–279.
[35] M. Iman, H.R. Arabnia, K. Rasheed, A review of deep transfer learning and recent advancements, Technologies 11 (2) (2023) 40.
[36] C. Lee, G. Song, H. Kim, J.C. Ye, M. Jang, Deep learning based on parameterized physical forward model for adaptive holographic imaging with unpaired data, Nat. Mach. Intell. 5 (1) (2023) 35–45.
[37] J.-Y. Zhu, T. Park, P. Isola, A.A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2223–2232.
[38] L. MacNeil, S. Missan, J. Luo, T. Trappenberg, J. LaRoche, Plankton classification with high-throughput submersible holographic microscopy and transfer learning, BMC Ecol. Evol. 21 (1) (2021) 123.
[39] K. Jaferzadeh, S. Son, A. Rehman, S. Park, I. Moon, Automated stain-free holographic image-based phenotypic classification of elliptical cancer cells, Adv. Photonics Res. 4 (1) (2023) 2200043.
[40] A. Rehman, H. An, S. Park, I. Moon, Automated classification of elliptical cancer cells with stain-free holographic imaging and self-supervised learning, Opt. Laser Technol. 174 (2024) 110646.
[41] M. Caron, I. Misra, J. Mairal, P. Goyal, P. Bojanowski, A. Joulin, Unsupervised learning of visual features by contrasting cluster assignments, Adv. Neural Inf. Process. Syst. 33 (2020) 9912–9924.
[42] J. Zbontar, L. Jing, I. Misra, Y. LeCun, S. Deny, Barlow twins: Self-supervised learning via redundancy reduction, in: Proceedings of the International Conference on Machine Learning, PMLR, 2021, pp. 12310–12320.
[43] M. Loey, G. Manogaran, N.E.M. Khalifa, A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images, Neural Comput. Appl. (2020) 1–13.
[44] Y. Nagahama, Phase retrieval using hologram transformation with U-Net in digital holography, Opt. Continuum 1 (7) (2022) 1506–1515.
[45] L. Huang, H. Chen, T. Liu, A. Ozcan, Self-supervised learning of hologram reconstruction using physics consistency, Nat. Mach. Intell. 5 (8) (2023) 895–907.
[46] N. Shlezinger, J. Whang, Y.C. Eldar, A.G. Dimakis, Model-based deep learning, Proc. IEEE 111 (5) (2023) 465–499.
[47] N. Shlezinger, Y.C. Eldar, S.P. Boyd, Model-based deep learning: On the intersection of deep learning and optimization, IEEE Access 10 (2022) 115384–115398.
[48] J.R. Hershey, J.L. Roux, F. Weninger, Deep unfolding: Model-based inspiration of novel deep architectures, 2014, arXiv:1409.2574.
[49] V. Monga, Y. Li, Y.C. Eldar, Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing, IEEE Signal Process. Mag. 38 (2) (2021) 18–44.
[50] S. Wu, A. Dimakis, S. Sanghavi, F. Yu, D. Holtmann-Rice, D. Storcheus, A. Rostamizadeh, S. Kumar, Learning a compressed sensing measurement matrix via gradient unrolling, in: Proceedings of the International Conference on Machine Learning, PMLR, 2019, pp. 6828–6839.
[51] Z. Zhang, Y. Liu, J. Liu, F. Wen, C. Zhu, AMP-Net: Denoising-based deep unfolding for compressive image sensing, IEEE Trans. Image Process. 30 (2020) 1487–1500.
[52] Y. Zhang, S.H. Chan, E.Y. Lam, Photon-starved snapshot holography, APL Photonics 8 (5) (2023) 056106(1–13).
[53] H. Li, X. Chen, Z. Chi, C. Mann, A. Razi, Deep DIH: Single-shot digital in-line holography reconstruction by deep learning, IEEE Access 8 (2020) 202648–202659.
[54] F. Niknam, H. Qazvini, H. Latifi, Holographic optical field recovery using a regularized untrained deep decoder network, Sci. Rep. 11 (1) (2021) 10903.
[55] F. Wang, Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, G. Situ, Phase imaging with an untrained neural network, Light: Sci. Appl. 9 (1) (2020) 77.
[56] A.S. Galande, V. Thapa, H.P.R. Gurram, R. John, Untrained deep network powered with explicit denoiser for phase recovery in inline holography, Appl. Phys. Lett. 122 (13) (2023) 133701.
[57] Y. Zhang, X. Liu, E.Y. Lam, Single-shot inline holography using a physics-aware diffusion model, Opt. Express 32 (6) (2024) 10444–10460.
[58] J. Ho, A. Jain, P. Abbeel, Denoising diffusion probabilistic models, Adv. Neural Inf. Process. Syst. 33 (2020) 6840–6851.
[59] N. Chen, L. Cao, T.-C. Poon, B. Lee, E.Y. Lam, Differentiable imaging: A new tool for computational optical imaging, Adv. Phys. Res. 2 (6) (2023) 2200118.
[60] N. Chen, D.J. Brady, E.Y. Lam, Differentiable imaging: Progress, challenges, and outlook, Adv. Devices Instrum. 6 (0117) (2025) 34.
[61] N. Chen, C. Wang, W. Heidrich, $\partial\mathcal{H}$: Differentiable Holography, Laser Photonics Rev. 17 (9) (2023) 2200828.
[62] Y. Gao, L. Cao, Iterative projection meets sparsity regularization: Towards practical single-shot quantitative phase imaging with in-line holography, Light: Adv. Manuf. 4 (1) (2023) 37–53.

[63] L. Denis, D. Lorenz, E. Thiébaut, C. Fournier, D. Trede, Inline hologram reconstruction with sparsity constraints, Opt. Lett. 34 (22) (2009) 3475–3477.
[64] J.R. Fienup, Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint, J. Opt. Soc. Amer. A 4 (1) (1987) 118–123.
[65] N. Chen, Y. Wu, C. Tan, L. Cao, J. Wang, E.Y. Lam, Uncertainty-aware Fourier ptychography, Light: Sci. Appl. 14 (1) (2025) 236.
[66] N. Chen, E.Y. Lam, Differentiable pixel-super-resolution lensless imaging, Opt. Lett. 50 (4) (2025) 1180–1183.
[67] Y. Wu, J. Wang, S.T. Thoroddsen, N. Chen, Single-shot high-density volumetric particle imaging enabled by differentiable holography, IEEE Trans. Ind. Inform. 20 (12) (2024) 13696–13706.
[68] N. Chen, C. Wang, W. Heidrich, Snapshot space-time holographic three-dimensional particle tracking velocimetry, Laser Photonics Rev. 15 (8) (2021) 2100008.
[69] L.E. Altman, D.G. Grier, CATCH: Characterizing and tracking colloids holographically using deep neural networks, J. Phys. Chem. B 124 (9) (2020) 1602–1610.
[70] L.E. Altman, D.G. Grier, Machine learning enables precise holographic characterization of colloidal materials in real time, Soft Matter 19 (16) (2023) 3002–3014.
[71] H. Zhang, B. Wang, L. Zhang, Y. Zhao, Y. Wang, J. Feng, W. Xiao, G. Wang, Y. Wu, X. Wu, Adaptive in-focus particle detection and segmentation in holographic 3D image with mechanism-guided machine learning, Opt. Lasers Eng. 181 (2024) 108400.
[72] Z.C. Lipton, J. Berkowitz, C. Elkan, A critical review of recurrent neural networks for sequence learning, 2015, arXiv:1506.00019.
[73] L. Huang, T. Liu, X. Yang, Y. Luo, Y. Rivenson, A. Ozcan, Holographic image reconstruction with phase recovery and autofocusing using recurrent neural networks, ACS Photonics 8 (6) (2021) 1763–1774.
[74] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, A. Anandkumar, Fourier neural operator for parametric partial differential equations, 2020, arXiv:2010.08895.
[75] Z. Li, H. Huang, D. Sun, Z. Zheng, F. Wang, S. Panezai, J. Xing, Y. Yang, K. Qiu, Deep learning infrared holography with transformer for crystal material characterization, Cryst. Growth Des. 24 (16) (2024) 6851–6864.
[76] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, Adv. Neural Inf. Process. Syst. 30 (2017).
[77] G. Bai, H. Guo, C. Xiao, Research on the application of transformer in computer vision, J. Phys.: Conf. Ser. 2649 (1) (2023) 012033.
[78] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby, An image is worth 16x16 words: Transformers for image recognition at scale, 2020, arXiv:2010.11929.
[79] C. Alvarez, B. Baskar, C. Berrospe-Rodriguez, S. Kavian, R. Sellers, M.J. Powell-Palm, G. Aguilar, Digital holography interferometry for refractive index characterization and temperature mapping of aqueous solutions at supercooled temperatures, Laser Photonics Rev. 19 (9) (2025) 2400921.
[80] M. Georges, Y. Zhao, J.-F. Vandenrijt, Holography in the invisible. From the thermal infrared to the terahertz waves: Outstanding applications and fundamental limits, Light: Adv. Manuf. 2 (2) (2022) 335–348.
[81] Z. Liu, T. Zeng, X. Zhan, X. Zhang, E.Y. Lam, Generative approach for lensless imaging in low-light conditions, Opt. Express 33 (2) (2025) 3021–3039.
[82] T.E. Chambers, I.M. Reid, M. Hamilton, A lightweight holographic imager for cloud microphysical studies from an untethered balloon, Atmospheric Meas. Tech. 17 (10) (2024) 3237–3253.
[83] V.S. Barroso, P. Švančara, C. Goodwin, S.C. Ajithkumar, I. Dimina, S. Schiattarella, P. Smaniotto, L. Solidoro, M. Cromb, R. Prizia, A.J. Kent, S. Weinfurtner, Digital holographic imaging of superfluid helium free surfaces, 2025, arXiv:2509.10235.
[84] M. Locatelli, E. Pugliese, P. Poggi, S. Euzzor, R. Meucci, SWIR digital holography and imaging through smoke and flames: Unveiling the invisible, Opt. Express 31 (25) (2023) 42090–42098.
[85] S. Zhu, C. Wang, H. Liu, P. Zhang, E.Y. Lam, Computational neuromorphic imaging: principles and applications, in: Computational Optical Imaging and Artificial Intelligence in Biomedical Sciences, 2024, pp. 4–10.
[86] C. Wang, S. Li, S. Zhu, E.Y. Lam, Motion-resolved event-based holography, in: Proceedings of the Advanced Optical Imaging Technologies VIII, SPIE, 2025, 1371709.
[87] Z. Ge, C. Wang, J. Huang, E.Y. Lam, Event-driven neuromorphic holography for dynamic particle imaging, Opt. Lett. 50 (5) (2025) 1496–1499.
[88] E.R. Fossum, J. Ma, S. Masoodian, L. Anzagira, R. Zizza, The quanta image sensor: Every photon counts, Sensors 16 (8) (2016) 1260.
[89] C. Bruschini, H. Homulle, I.M. Antolovic, S. Burri, E. Charbon, Single-photon avalanche diode imagers in biophotonics: Review and outlook, Light: Sci. Appl. 8 (1) (2019) 87.
[90] L.-J. Kong, J. Zhang, Z. Zhang, X. Zhang, Quantum holographic microscopy, Laser Photonics Rev. 19 (9) (2025) 2401909.
[91] M. Liebel, N. Pazos-Perez, N.F. van Hulst, R.A. Alvarez-Puebla, Surface-enhanced Raman scattering holography, Nature Nanotechnology 15 (12) (2020) 1005–1011.
[92] P. Gough, R. Bates, Speckle holography, Opt. Acta: Int. J. Opt. 21 (3) (1974) 243–254.
[93] S. Aime, M. Sabato, L. Xiao, D. Weitz, Dynamic speckle holography, Phys. Rev. Lett. 127 (8) (2021) 088003.
[94] B.W. Schilling, T.-C. Poon, G. Indebetouw, B. Storrie, K. Shinoda, Y. Suzuki, M.H. Wu, Three-dimensional holographic fluorescence microscopy, Opt. Lett. 22 (19) (1997) 1506–1508.
[95] Y. Li, Y. Zhu, J. Huang, Y.-W. Ho, J.K.-H. Fang, E.Y. Lam, High-throughput microplastic assessment using polarization holographic imaging, Sci. Rep. 14 (2024) 2355(1–11).
[96] L. Nikolova, P.S. Ramanujam, Polarization Holography, Cambridge University Press, 2009.
[97] D. Hillmann, C. Lührs, T. Bonin, P. Koch, G. Hüttmann, Holoscopy—holographic optical coherence tomography, Opt. Lett. 36 (13) (2011) 2390–2392.
[98] S. Kalenkov, G. Kalenkov, A. Shtanko, Hyperspectral holography: An alternative application of the Fourier transform spectrometer, J. Opt. Soc. Am. B 34 (5) (2017) B49–B55.
[99] E. Vicentini, Z. Wang, K. Van Gasse, T.W. Hänsch, N. Picqué, Dual-comb hyperspectral digital holography, Nat. Photonics 15 (12) (2021) 890–894.
[100] E. Olsson, P. Gren, M. Sjödahl, Photoacoustic holographic imaging of absorbers embedded in silicone, Appl. Opt. 50 (17) (2011) 2551–2558.
[101] Y. Zhu, Y. Li, J. Huang, E.Y. Lam, Smart polarization and spectroscopic holography for real-time microplastics identification, Commun. Eng. 3 (2024) 32(1–8).
[102] J. Huang, Y. Zhu, Y. Li, E.Y. Lam, Snapshot polarization-sensitive holography for detecting microplastics in turbid water, ACS Photonics 10 (12) (2023) 4483–4493.
[103] J. Huang, S. Zhu, Y. Li, C. Wang, E.Y. Lam, Computational polarized holography for automatic monitoring of microplastics in scattering aquatic environments, APL Photonics 10 (3) (2025) 036114.
[104] J. Chen, Y. Zhao, Y. Li, Y. Zhu, Y. Cao, E.Y. Lam, Snapshot HoloSpec: dispersion-coded 4D feature learning for waterborne particle monitoring, J. Phys.: Photonics 8 (1) (2026) 015051.
[105] T. Takahashi, Z. Liu, T. Thevar, N. Burns, D. Lindsay, J. Watson, S. Mahajan, S. Yukioka, S. Tanaka, Y. Nagai, et al., Multimodal image and spectral feature learning for efficient analysis of water-suspended particles, Opt. Express 31 (5) (2023) 7492–7504.