Using model-driven deep learning to achieve high-fidelity 4K color holographic display
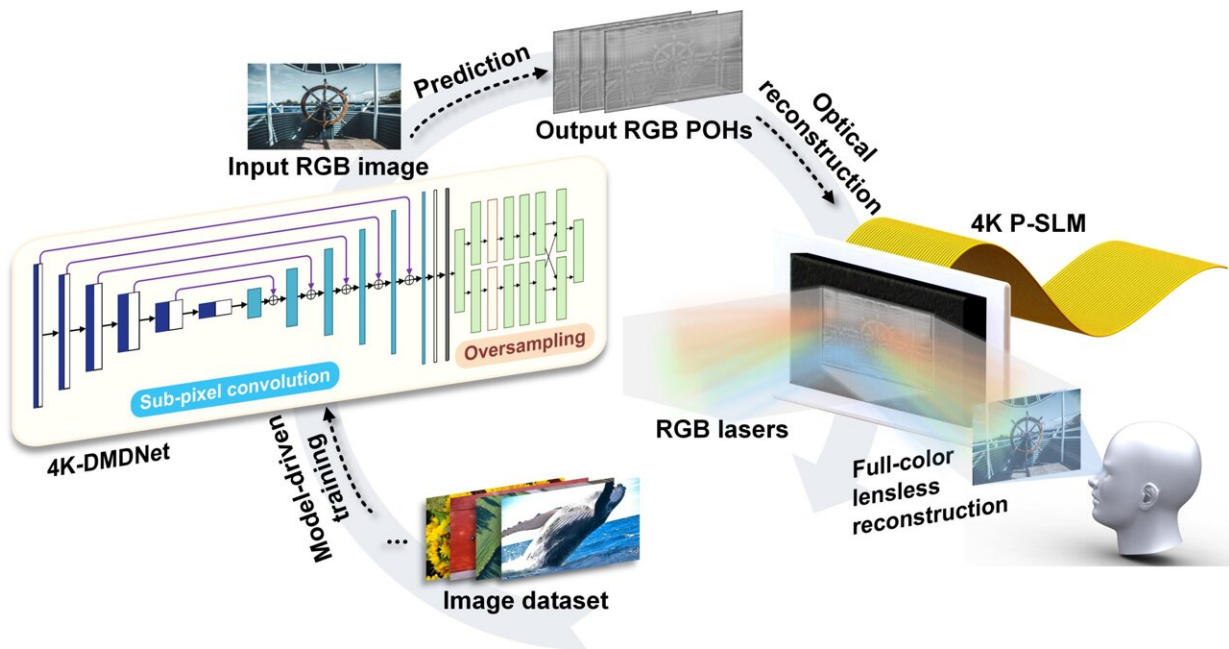
February 22 2023

Fig. 1 Generation and reconstruction processes of 4K holograms by the 4K-DMDNet. Credit: *Opto-Electronic Advances* (2023). DOI: 10.29026/oea.2023.220135

In 2009, the IMAX 3D film "Avatar" swept the global film market. A few years later, the Hatsune Miku 3D concert attracted all anime fans' attention. And recently AR/VR 3D headwear devices led to a booming development of the metaverse. Each progress in the 3D display field brings essential social concerns and economic benefits.

To obtain more realistic visual experiences, most of the mainstream commercial solutions for 3D display are based on the principles of binocular vision.

However, unlike the observation of real 3D objects, the depth of visual focus remains unchanged while the viewer is wearing the device to obtain 3D information. This type of vergence accommodation conflict makes the viewer susceptible to visual fatigue and vertigo, limiting the user experiences.

Computer-generated holography (CGH) can avoid the generation of vergence accommodation conflict from the origin. The experimental setups are simple and compact. CGH has received significant attention from academia and industry. It is regarded as the future form of 3D display.

In principle, CGH codes the 3D object into a digital two-dimensional (2D) hologram based on diffractive calculations. And then the 2D hologram is uploaded to a [spatial light modulator](#) (SLM) illuminated by plane waves. The optical reconstruction of the 3D object is obtained at a certain distance. CGH has potential applications in a wide range of 3D displays such as head-mounted displays, heads-up displays, and projection displays.

How to generate high-speed and high-quality 2D holograms is a key issue and essential research direction in this field at present.

Recently, Hololab at Tsinghua University proposed a model-driven deep learning neural network, called 4K-DMDNet. It realizes the high-quality high-speed hologram generation and achieves high-fidelity 4K color holographic displays. The paper is published in the journal *Opto-Electronic Advances*.
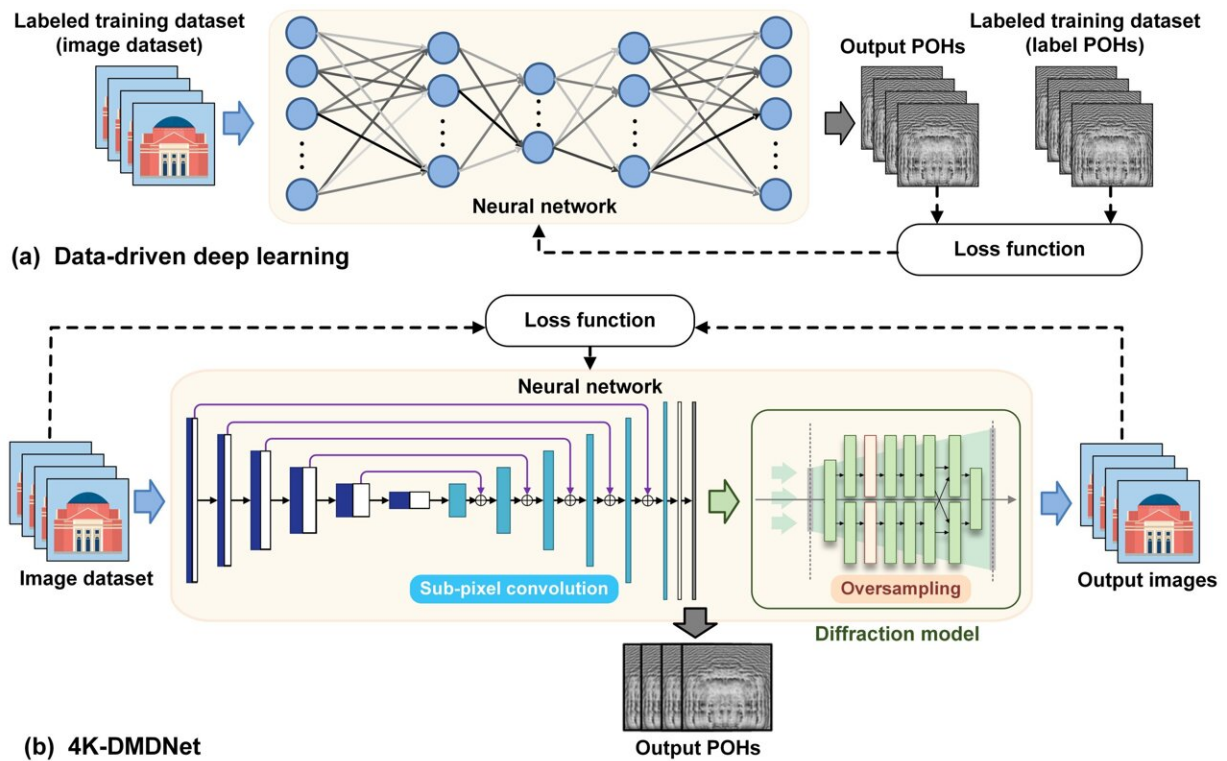
Fig. 2 Comparison of (a) data-driven deep learning with (b) 4K-DMDNet in terms of the training principle. Credit: *Opto-Electronic Advances* (2023). DOI: 10.29026/oea.2023.220135

Due to the limitations of SLM, the calculated complex-amplitude distributions on the holographic plane need to be converted into amplitude-only holograms or phase-only holograms (POHs). Among them, the POH generation process is typically an ill-posed inverse problem. It has the challenges that the solution may be not unique, stable, or existing.

The iterative algorithms can convert the POH generation process to the optimization problem. Numerical solutions with good convergence can be obtained. However, the algorithms face a trade-off between

computational speed and reconstruction quality.

The powerful parallel processing capabilities of deep learning have brought about revolutionary improvements in solving optimization problems. The profound impacts of the deep learning on CGH have also been made.

The training dataset of 3D objects and the corresponding hologram dataset are obtained in advance to serve as the inputs and outputs of the neural network. The neural network is trained to learn the mapping relationship between them. The trained network can achieve fast prediction of display target inputs outside the training dataset. It is expected to simultaneously realize high-speed and high-quality hologram generation.

The idea of using neural networks for hologram generation was proposed by Japanese researchers as early as 1998. But limited by the hardware and software performance of computers at that time, only preliminary results were obtained.

With the broad applications of GPUs and convolutional neural networks (CNNs), current hardware and software performance are more suitable with the mathematical characteristics of CGH. The learning-based CGH has seen rapid developments.

In 2021, researchers at MIT proposed a Tensor holography network that enables the real-time generation of 2K holograms on smartphones.
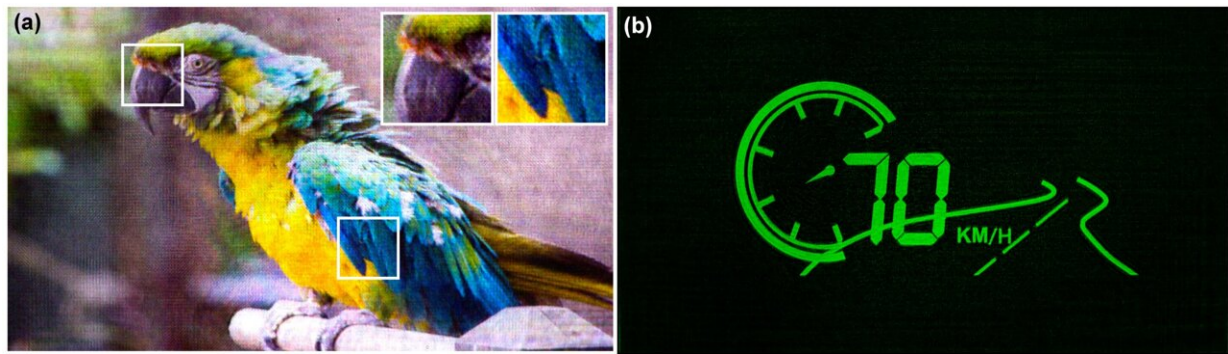
Fig. 3 Optical reconstructions of different types of images: (a) color image and (b) binary image. Credit: *Opto-Electronic Advances* (2023). DOI: 10.29026/oea.2023.220135

To obtain accurate network predictions, the training dataset and corresponding hologram dataset need a time-consuming generation process. Moreover, as the network simply learns the mapping between inputs and outputs, the quality of the hologram dataset limits the ceiling of the training results.

To break through the above limitations of data-driven deep learning, hologram generation schemes based on model-driven deep learning are proposed.

Instead of generating the hologram dataset in advance, the network is trained by using the forward physical model of the inverse problem as a constraint in the model-driven method. The network can thus learn how to encode holograms autonomously, breaking through the limitations of the hologram dataset size and quality.

However, conventional model-driven deep learning networks require transfer learning on the display targets to achieve better performances.

The additional time cost limits the practical applications of model-driven deep learning.

The 4K-DMDNet proposed in this work uses a residual U-Net [neural network](#) framework. The Fresnel diffraction model acts as the constraint for the training process. It is capable of high-fidelity 4K [hologram](#) generation without transfer learning.

In general, the prediction performances of the network are influenced by both the limited learning capacity of the network and the insufficient constraints in the training process.

To address the challenges of the limited learning capability, 4K-DMDNet introduces the sub-pixel convolution method. In the upsampling path, the number of channels is expanded four times by using convolutions, and the spatial expansion is obtained through the pixel shuffle. The sub-pixel convolution method solves the challenges of adding a large number of zero parameters for spatial expansion in the traditional transposed convolution. It increases the learnable parameters in the upsampling path to four times the original size without changing the overall data volume. It effectively enhances the learning capability of the network, resulting in a significant improvement in the sharpness and fidelity of the reconstructions.

To address the challenges of the insufficient constraints in the training process, 4K-DMDNet introduces the oversampling operation in the Fresnel diffraction model. The constraint region in the frequency domain is zero-padded to double the size in the calculation process. According to the mapping between the spatial sampling interval and frequency range, the reconstructions meet the Nyquist-Shannon sampling theorem. While tightening the frequency domain constraints, the accuracy of the diffraction model is enhanced.

More information: Kexuan Liu et al, 4K-DMDNet: diffraction model-driven network for 4K computer-generated holography, *Opto-Electronic Advances* (2023). DOI: 10.29026/oea.2023.220135

Provided by Compuscript Ltd