This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

HoloPointNet: A Deep Learning Framework for Efficient 3D Point Cloud Holography

Ankit Amrutkar[1✉], Ahmet Nazlioglu[1], Björn Kampa[2], Volkmar Schulz[1], Johannes Stegmaier[1‡], Markus Rothermel[3‡], and Dorit Merhof[4,5‡✉]

[1] Institute of Imaging and Computer Vision, RWTH Aachen University, Germany
[2] Systems Neurophysiology, Institute of Zoology, RWTH Aachen University, Germany
[3] Institute of Physiology, Otto von Guericke University Magdeburg, Germany
[4] Institute of Image Analysis and Computer Vision, University of Regensburg, Germany
[5] Fraunhofer Institute for Digital Medicine MEVIS, Bremen, Germany
ankit.amrutkar@lfb.rwth-aachen.de; dorit.merhof@ur.de[✉]

Abstract. HoloPointNet presents a novel deep-learning framework for 3D point cloud holography. Generally, computer generated holography (CGH) methods typically rely on stacked 2D slices and suffer from inefficiencies. These 2D slices often contain empty regions in natural 3D scenes or are intentionally sparse in applications like holographic optogenetics. This results in excessive memory consumption and increased processing latency. In contrast, HoloPointNet directly processes 3D point cloud data using a concatenation-based feature extractor, followed by hierarchical upsampling and wavefront reconstruction modules, eliminating redundant spatial regions and improving efficiency. This design allows for the direct mapping of point cloud data to phase modulations for spatial light modulators (SLMs). By employing a structured convolutional feature transformation pipeline, HoloPointNet enables hierarchical refinement of spatial embeddings, enhancing feature encoding accuracy. HoloPointNet offers the capability to generate multiplane holograms, effectively addressing the complexities of 3D volumetric data. This capability, combined with fast inference times, enables real-time holography for applications such as optogenetics. The code is available at https://github.com/AnkitAmrutkar/HoloPointNet.git

Keywords: Point-Cloud Holography · HoloPointNet · Phase Retreival.

1 Introduction

Computer-generated holography (CGH) is a computational technique that manipulates a coherent wavefront to create custom illumination patterns. At its core, CGH involves solving the phase retrieval problem, an inverse, non-convex, and ill-posed task of estimating phase information from intensity measurements. This enables precise optical control and finds applications in various fields, including augmented reality, 3D displays [16,3,25], systems neuroscience [20,8,23,12],

‡ These authors contributed equally to this work.

and optical trapping [17,13,6,18]. Several iterative and deep learning-based methods have been developed to solve the phase retrieval [7] problem. Traditional iterative approaches, such as the Gerchberg-Saxton algorithm [14] and its variants [31,5], as well as more advanced techniques like gradient descent [33] or Wirtinger derivatives [4], are often computationally expensive. Non-iterative methods [9,32] also exist for hologram generation. In contrast, deep learning approaches [2,34,28,11,30,24] have gained prominence in CGH due to their faster inference times and high image fidelity, leveraging neural network architectures that directly infer phase masks from target intensity distributions.

Most existing neural network architectures for CGH rely on grid-based representations of images. However, these grid-based 3D representations are computationally inefficient and significantly hinder their applicability to real-time CGH. Many current methods approximate volumetric intensity patterns as stacks of 2D planes, which is computationally costly and redundant in scenarios where the illumination pattern is sparse. For example, in holographic optogenetics [8,12], only a small portion of the 3D volume is relevant, as specific neurons must be targeted for activation or inhibition. Using dense 2D stacks increases memory consumption and inference times.

To address these limitations, point cloud-based representations offer an efficient alternative. Unlike grid-based volumetric data, point clouds encode only the essential target locations, significantly reducing computational overhead. This approach is particularly suited for the sparsity constraints inherent in optogenetics, where precise targeting is critical and minimizing off-target illumination is essential. DeepCGH2.0 [10] is a notable effort in implementing point cloud-based CGH. It maps point clouds to optimal phase modulation, significantly reducing memory requirements compared to traditional 2D slicing. However, it still faces challenges in approximating 3D holograms. In contrast, we propose a neural network architecture that maps point clouds to phase encodings for multi-plane holography. Our method is computationally efficient and adaptable for real-time applications. The main contributions of this paper are as follows:

1. HoloPointNet: We propose a deep learning framework and evaluate it on a biologically inspired dataset designed to reflect typical CGH applications targeting custom sparse cell-sized ensembles. It processes point cloud representations using a concatenation-based feature extractor, followed by hierarchical upsampling and a wavefront reconstruction module to efficiently approximate the phase at the target plane for holographic reconstruction.
2. Multiplane Holography with Real-Time Inference: HoloPointNet generates multiplane holograms with fast inference, enabling real-time applications like holographic optogenetics.

2 Method

2.1 Preliminaries

Forward Models: We structure the network architecture using a physics-inspired approach [29], leveraging a physics-based forward model to guide the

network towards an optimal solution. Specifically, we employ free-space propagation as the forward model, implemented using the band-limited angular spectrum method.[21,15]. *Free space propagation:* In this setup, no diffractive element is placed between the SLM and the hologram plane. Light propagation is simulated using the band-limited angular spectrum method [22]. Here, the forward model hyperparameters (FMH) are the wavelength of light (λ), propagation distance (d), SLM pixel-resolution (M), pixel-pitch (Δx) i.e. FMH $\equiv (\lambda, \Delta x, M, d)$. [§]

$$
\begin{aligned}
\Psi_{\mathrm{ASM}}\left(\Gamma(\mathbf{x},\mathbf{y})\right) &= \widetilde{\mathcal{F}}\left[\Gamma(\mathbf{u},\mathbf{v})\mathrm{H}\left(\lambda,\Delta x, M, d\right)\right] \\
\Psi^{-1}_{\mathrm{ASM}}\left(\Gamma(\mathbf{x},\mathbf{y})\right) &= \widetilde{\mathcal{F}}\left[\Gamma(\mathbf{u},\mathbf{v})\mathrm{H}\left(\lambda,\Delta x, M, -d\right)\right] \\
\mathrm{H}\left(\mathrm{FMH}\right) &= \mathrm{H}'\left(\mathbf{u},\mathbf{v};\mathrm{FMH}\right)\mathrm{rect}\left(\frac{\mathbf{u}}{2u_{\mathrm{BL}}}\right)\mathrm{rect}\left(\frac{\mathbf{v}}{2v_{\mathrm{BL}}}\right) \\
\mathrm{H}'\left(\mathbf{u},\mathbf{v};\mathrm{FMH}\right) &= \begin{cases} \exp^{(\mathrm{i}2\pi w(\mathbf{u},\mathbf{v})d)} & \text{, if } u^2+v^2 \leq \lambda^2 \\ 0 & \text{, otherwise} \end{cases} \\
w\left(\mathbf{u},\mathbf{v}\right) &= \left(\lambda^{-2}-\mathbf{u}^2-\mathbf{v}^2\right)^{1/2} \\
\left(u_{\mathrm{BL}}, v_{\mathrm{BL}}\right) &= \frac{1}{\left[\left(2d(\Delta u,\Delta v)\right)^2+1\right]^{1/2}\lambda}
\end{aligned}
\tag{1}
$$

2.2 HoloPointNet

We introduce HoloPointNet ($\mathcal{H}^{\mathrm{w}}_{\mathrm{TP}}$) (Fig. 1), a deep learning framework for phase retrieval in CGH. The network approximates the wavefront ($\mathrm{E}^{\mathcal{H}}_{\mathrm{TP}}(\mathbf{x},\mathbf{y})$), at the target plane (TP) based on a given 3D point cloud that represents the desired holographic structure. This estimated wavefront is then backward propagated through the inverse system Ψ^{-1} to the SLM plane, where laser amplitude constraints are applied. The constrained wavefront is subsequently forward propagated via Ψ to the target plane, where a loss function evaluates the discrepancy between the reconstructed wavefront and the rendered point cloud which is rendered using a custom-built pipeline (Ω). To achieve 3D reconstruction of the wavefront, we employ the forward model given in Eq. 1 to propagate the estimated wavefront before evaluating the loss function. Here the loss function is defined as a weighted combination of mean squared error and accuracy Eq. 3a. The overall unsupervised training pipeline is described by Eq. 2 and inspired by the approaches presented in [11,10,1]. The HoloPointNet architecture ($\mathcal{H}^{\mathrm{w}}_{\mathrm{TP}}$), which estimates the wavefront ($\mathrm{E}^{\mathcal{H}}_{\mathrm{TP}}$)($\cdot$), consists of three key modules: A Concatenation-Based Feature Extractor ($\Theta(\cdot)$) (CBFE), Hierarchical Upsampling ($\Lambda(\cdot)$) (HU), and a Wavefront Reconstruction Module ($\mathcal{W}(\cdot)$) (WRM). Each module plays a crucial role in processing the input point cloud and generating a high-resolution phase distribution. The core transformation in HoloPointNet is expressed as:

$$
\mathrm{E}^{\mathcal{H}}_{\mathrm{TP}}(\mathbf{x},\mathbf{y}) = \mathcal{H}^{\mathrm{w}}_{\mathrm{TP}}(\mathcal{P}) \equiv \mathcal{W}\circ\Lambda\circ\Theta(\mathcal{P}) \tag{2a}
$$

[§] Notation follows [1,21].

HoloPointNet (Fig. 1) takes as input a set of N points, $\mathcal{P} = (p_j | j = 1, ..., N) \in \mathbb{R}^{N\times 5}$, where each point p_j is defined by its spatial coordinates (x, y, z) and point properties (r, i), where r and i represent the radius and intensity of the circular points, respectively. We use circular points to mimic cell-sized ensembles.

The $\Theta(\cdot)$ module applies shared multilayer perceptions (MLPs) implemented as 1D convolutions with batch normalization and ReLU activation. These shared weights extract per-point features efficiently across the point cloud. Features from different layers are concatenated to capture multi-scale information, with an early-stage concatenation merging low-level features and a later-stage concatenation integrating deeper representations. The resulting feature map is then further processed through convolutional layers before a global max-pooling operation aggregates the point features into a structured latent representation using fully connected layer. The $\Theta(\cdot)$ module is primarily inspired by the architectures in [26,27,19].

The $\Lambda(\cdot)$ module refines the feature map by expanding the channel dimension while preserving spatial resolution through concatenated convolutional layers. This is followed by residual upsampling blocks, which progressively increase the spatial resolution using a combination of standard convolutions and transposed convolutions.

The $\mathcal{W}(\cdot)$ module reduces the channel dimensions to reconstruct both the magnitude and phase of the wavefront $(\mathrm{E}^{\mathcal{H}}_{\mathrm{TP}}(\mathbf{x}, \mathbf{y}))$ at the target plane. The reconstructed wavefront at the target plane is then propagated via Ψ^{-1} to the SLM plane $(\mathrm{E}_{\mathrm{SLM}}(\mathbf{x}', \mathbf{y}'))$, where the final SLM phase profile is estimated $(\phi_{\mathrm{SLM}}(\mathbf{x}', \mathbf{y}'))$:

$$\begin{aligned} \Psi^{-1}\left[\mathrm{E}^{\mathcal{H}}_{\mathrm{TP}}(\mathbf{x}, \mathbf{y})\right] &= \mathrm{E}_{\mathrm{SLM}}(\mathbf{x}', \mathbf{y}'), \\ \angle \mathrm{E}_{\mathrm{SLM}}(\mathbf{x}', \mathbf{y}') &= \phi_{\mathrm{SLM}}(\mathbf{x}', \mathbf{y}'). \end{aligned} \tag{2b}$$

The constrained SLM wavefront $(\mathcal{A}_{\mathrm{SLM}}(\mathbf{x}', \mathbf{y}') \exp(\mathrm{i}\phi_{\mathrm{SLM}}(\mathbf{x}', \mathbf{y}')))$, which incorporates laser amplitude constraints $(\mathcal{A}_{\mathrm{SLM}}(\mathbf{x}', \mathbf{y}'))$ is then forward propagated back to the target plane:

$$\Psi\left[\mathcal{A}_{\mathrm{SLM}}(\mathbf{x}', \mathbf{y}')\mathrm{e}^{\left(\mathrm{i}\phi_{\mathrm{SLM}}(\mathbf{x}', \mathbf{y}')\right)}\right] = \mathrm{E}^{\mathrm{r}}_{\mathrm{TP}}(\mathbf{x}, \mathbf{y}). \tag{2c}$$

Inter-plane propagation (z) at the target plane is also calculated via Ψ. A loss function $\mathcal{L}$ then evaluates the difference between the rendered point cloud $(\Omega(\mathcal{P}))$ and the reconstructed (r) amplitude $(|\mathrm{E}^{\mathrm{r}}_{\mathrm{TP}}(\mathbf{x}, \mathbf{y})|^2)$ at the target plane:

$$\mathcal{L}\left(\Omega(\mathcal{P}), (|\mathrm{E}^{\mathrm{r}}_{\mathrm{TP}}(\mathbf{x}, \mathbf{y})|^2)\right). \tag{2d}$$

Thus, this architecture performs implicit phase retrieval using a wavefront adjustment network as defined in [1]: first at the target plane via $\mathcal{W}(\cdot)$ (Eq. 2a), then at the SLM plane through Ψ^{-1} (Eq. 2b), all in an unsupervised fashion (Eq. 2c, Eq. 2d).

2.3 Experiments

Model: We generate a custom dataset of 60,000 images, each containing 36 points distributed across three planes. The dataset is split into training (90%)

and validation (10%) sets, with an additional independent test set of 1,000 images for evaluation. Our loss function is a weighted combination of accuracy (weight: 1) Eq. 3a and mean squared error (weight: 0.25). Images are standardized to a resolution of 256×256, with three planes each containing 12 simulated points. For forward and backward propagation, we employ free-space propagation with a wavelength of 1035 nm, a propagation distance of 0.4 m, and an inter-plane spacing of 0.05 m. The SLM parameters include a pixel pitch of 9.2 μm and a resolution of 256×256. The FMH used in this study were selected for proof-of-concept evaluation; for real-world optical CGH experiments, the model must be retrained using application-specific FMH. We employ a multistep learning rate scheduler with milestones at epochs 17, 18, and 19, and a decay factor of 0.1.

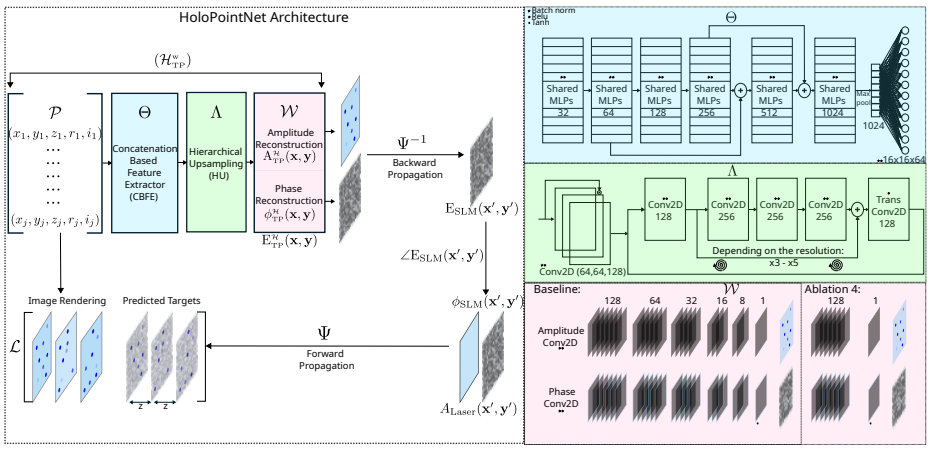

Fig. 1. HoloPointNet Framework: $(\mathcal{H}^{\mathrm{w}}_{\mathrm{TP}}(\mathcal{P}) \equiv \mathcal{W} \circ \Lambda \circ \Theta(\mathcal{P}))$. Here, Baseline and Ablation 4 are both variants of the proposed HoloPointNet. The Wavefront Reconstruction module $(\mathcal{W}(\cdot))$ (light red) has a hierarchical channel reduction for the Baseline model and abrupt channel reduction for the Ablation 4 model.

Comparison and Ablation Studies: We compare our model to an adapted 3D Gerchberg–Saxton (GS3D) [14,11] with free-space propagation, as existing methods (e.g., DeepCGH[11]/2.0 [10], etc.) differ in forward models or input representations, precluding direct comparison. Additionally, we conduct ablation studies to analyze the impact of different architectural components: (1) *Ablation 1*: Removal of hierarchical shared MLPs in the $\Theta(\cdot)$, replaced with a single shared MLP, to assess the significance of multiple shared MLPs. (2) *Ablation 2*: Substitution of the fully connected layer with a transpose convolution for feature extraction and image formation. (3) *Ablation 3*: Elimination of the residual block from the $\Lambda(\cdot)$ module to isolate the importance of residual connections. (4) *Ablation 4*: Modification of the $\mathcal{W}(\cdot)$ by replacing the hierarchical

channel reduction layer with abrupt convolutional layers (128 → 1), to evaluate the effect of this change on performance.

Evaluation Metrics: We evaluate our models presented in Sec. 2.3 using four key metrics, where I and $\widetilde{I}$ represent the 3D target and predicted intensity distributions, respectively. Contrast (C) quantifies the power density difference between target and non-target regions. Accuracy (A) measures the cosine similarity between the target and predicted images, providing an assessment of overall alignment. The efficiency metric (E) indicates the proportion of predicted intensity that overlaps with the target intensity, normalized by the total predicted intensity. Lastly, speckle contrast (C_{speckle}) characterizes noise or intensity variability in non-target regions by comparing the mean intensity (μ) to its standard deviation (σ).

$$C = \frac{\frac{\sum \widetilde{I}(x,y,z) \;\; I(x,y,z)}{\sum I(x,y,z)}}{\frac{\sum \widetilde{I}(x,y,z) \;\; (1-I(x,y,z))}{\sum (1-I(x,y,z))}}, A = \frac{\sum \widetilde{I}(x,y,z) \;\; I(x,y,z)}{\sqrt{\sum I(x,y,z)^2 \sum \widetilde{I}(x,y,z)^2}} \tag{3a}$$

$$C_{\text{speckle}} = \frac{\mu\left(\widetilde{I}(x,y,z)(1-I(x,y,z))\right)}{\sigma\left(\widetilde{I}(x,y,z)(1-I(x,y,z))\right)}, E = \frac{\sum \widetilde{I}(x,y,z) \;\; I(x,y,z)}{\sum \widetilde{I}(x,y,z)} \tag{3b}$$

3 Results and Discussion

We evaluated all models (Sec. 2.3) on a fixed test dataset consisting of 1,000 simulated images. Both qualitative and quantitative analyses indicate that the Baseline model and Ablation 4 outperform the iterative GS3D algorithm (Fig. 2, Fig. 3).

Ablation 1 performed poorly in both qualitative and quantitative evaluations, highlighting the importance of multi-layered shared MLPs in the feature extractor. Ablation 2 exhibited reduced contrast due to its high power density in non-target regions. Similarly, Ablation 3, which involved modifications to the hierarchical upsampling module, displayed significant non-target intensities, as reflected in qualitative assessments (Fig. 2,Fig. 3). Interestingly, Ablation 4 outperformed the Baseline model across most evaluation metrics (Fig. 3). While the validation loss for the Baseline model was lower than that of Ablation 4 during training, Ablation 4 ultimately achieved superior performance on this specific test dataset. We observed that the Baseline model tends to get trapped in local minima for certain test images, as indicated by the long-tailed distributions in the violin plots (Fig. 3). Initially, all models were evaluated with a batch size of 1, where the Baseline model, trained with a larger batch size (25), showed a performance drop.

To examine the impact of batch size, we re-evaluated both the Baseline and Ablation 4 models with larger batch sizes in Tab. 1. This reduced the long-tailed

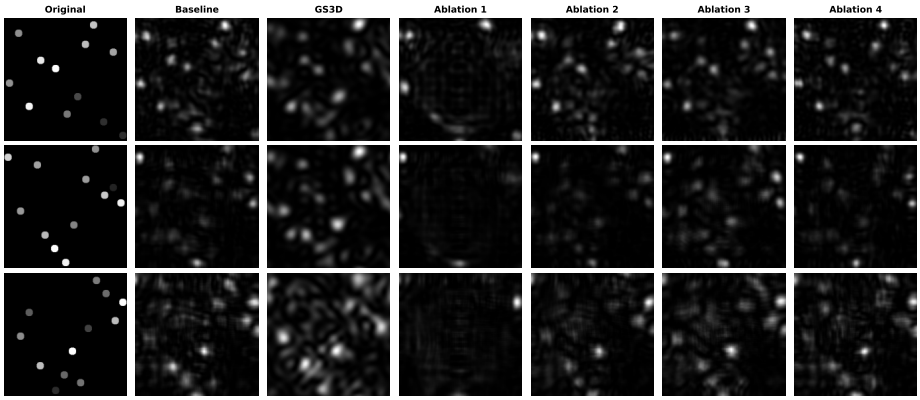

Fig. 2. Qualitative Analysis of Simulated Holographic Reconstruction: The rows represent the first, second, and third target intensity planes from top to bottom (inter-plane spacing of 0.05 m, first plane propagated at the distance of 0.4 m). The first three columns show comparative studies, while the last four columns present the results of the ablation studies.

distribution for the Baseline model and slightly improved its performance. However, Ablation 4 also benefited from the increased batch size and continued to outperform the Baseline model, demonstrating greater robustness across different batch size settings. Ablation 4 also outperformed Baseline model on a larger dataset with 2,000 images Tab. 1. Our results suggest that abrupt channel reduction introduces an information bottleneck, forcing the network to focus on the most critical features, particularly in sparse data scenarios. The qualitative analysis of Ablation 4 (Fig. 2) showed improved identification of high-amplitude regions, albeit with a trade-off of increased speckle contrast (Fig. 3), indicating greater intensity variation in non-target areas. We demonstrated the enhanced speed and performance of our network compared to the iterative GS3D algorithm (Tab. 1).

Metric	D:2000,BS:25	D:2000,BS:1	D:1000,BS:25	D:1000,BS:1	GS3D
Accuracy ↑	(0.565,0.583)	(0.565,0.583)	(0.559,0.583)	(0.56,0.583)	0.480
Contrast ↑	(6.743,7.116)	(6.576,6.968)	(6.636,7.192)	(6.526,6.97)	5.469
Efficiency ↑	(0.117,0.124)	(0.114,0.120)	(0.115,0.124)	(0.113,0.120)	0.1
Speckle Contrast ↓	(0.694,0.777)	(0.750,0.762)	(0.613,0.774)	(0.743,0.758)	0.739
Time (sec) ↓	(0.508,0.425)	(0.0321,0.0324)	(0.509,0.439)	(0.0332,0.0322)	0.533

Table 1. Average evaluation metrics across different datasets and batch sizes (D: dataset size, BS: batch size). For each value $(\cdot,\cdot)$, the first represents the Baseline model, and the second represents Ablation 4. Metrics in the final column represent averages over 1,000 images (D:1000) at iteration 30 of GS3D. Average iteration time in the final column is calculated cumulatively from iterations 1 through 30.

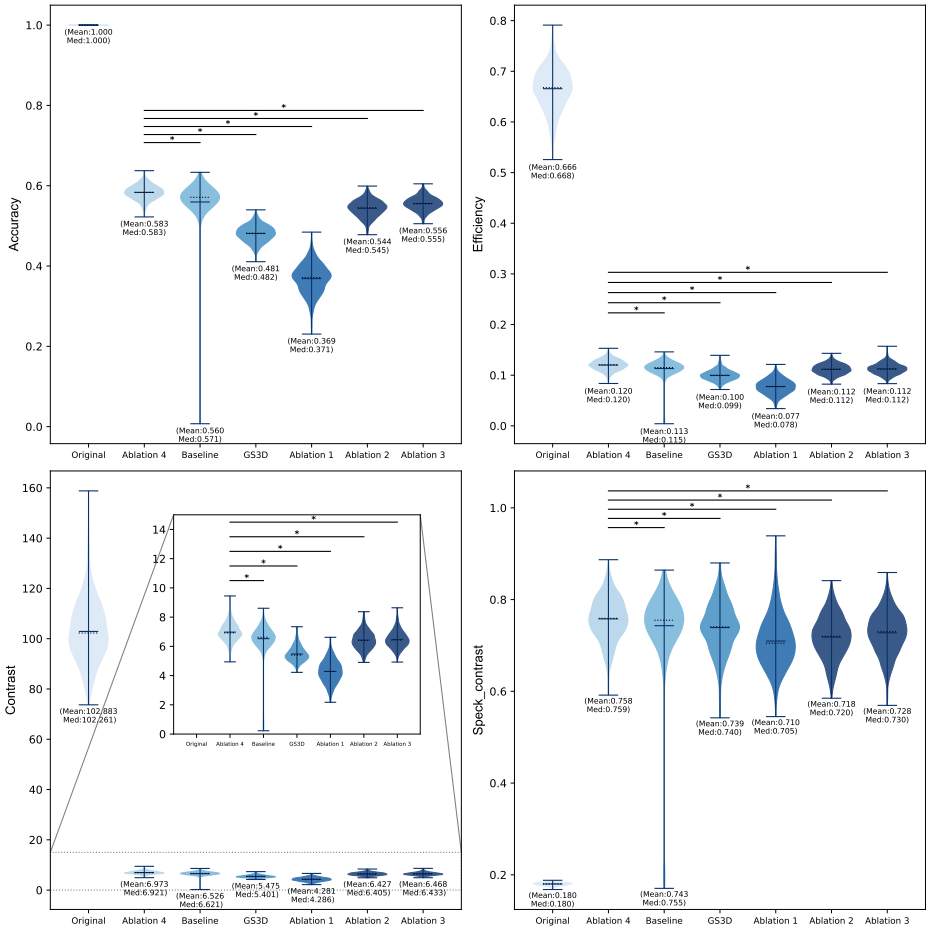

Fig. 3. Quantitative Analysis of Simulated Holographic Reconstruction: The four panels display evaluation metrics on 1,000 test images. The first violin plot shows the ideal scenario with similar images. The subsequent plots represent the Ablation 1-4 models, Baseline model and the GS3D algorithm at the 4,000th iteration. Mean and median (dotted line) values are indicated below each plot. A Wilcoxon signed-rank test (one-sided, alternative: "greater") confirmed that the Ablation 4 model significantly outperforms the others in accuracy, efficiency, and contrast ($p < 0.025$, marked *). However, the Ablation 4 model's speckle contrast is significantly higher than all other models, indicating poorer performance in this aspect.

4 Conclusion

In this paper, we present HoloPointNet, a novel deep learning framework that efficiently processes point cloud data for multi-plane holography. HoloPointNet significantly reduces computational overhead compared to traditional grid-based methods for holographic optogenetics and similar applications by leveraging a concatenation-based feature extractor, hierarchical upsampling, and a wavefront reconstruction module. Our results demonstrate that HoloPointNet outperforms the adapted iterative GS3D algorithm, in both speed and performance. Additionally, the model maintains high accuracy and contrast while improving efficiency. Notably, our analysis suggests that the HoloPointNet variant (Ablation 4) with abrupt channel reduction provides a more robust and effective solution, further enhancing performance in holographic reconstruction. Overall, HoloPointNet offers a computationally efficient and scalable approach, advancing the field of computer-generated holography. Additionally, a lighter version of the model has been successfully applied to 2D holography for free-space propagation and Fourier holography, further demonstrating its versatility.

Acknowledgments. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 368482240 / RTG2416. DM and BK are current RTG Principal Investigators (PIs), MR is a former RTG PI, and AA has been funded on an RTG PhD student position awarded to DM.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Amrutkar, A., Kampa, B., Schulz, V., Stegmaier, J., Rothermel, M., Merhof, D.: Towards robust and generalizable Gerchberg Saxton based physics inspired neural networks for computer generated holography: A sensitivity analysis framework. arXiv preprint arXiv:2505.00220 (2025)
2. Barbastathis, G., Ozcan, A., Situ, G.: On the use of deep learning for computational imaging. Optica **6**(8), 921–943 (2019)
3. Blinder, D., Birnbaum, T., Ito, T., Shimobaba, T.: The state-of-the-art in computer generated holography for 3D display. Light: Advanced Manufacturing **3**(3), 1 (2022)
4. Chakravarthula, P., Peng, Y., Kollin, J., Fuchs, H., Heide, F.: Wirtinger holography for near-eye displays. ACM Trans. Graph. **38**(6) (2019)
5. Chen, L., Zhang, H., He, Z., Wang, X., Cao, L., Jin, G.: Weighted constraint iterative algorithm for phase hologram generation. Applied Sciences **10**(10) (2020)
6. Curtis, J.E., Koss, B.A., Grier, D.G.: Dynamic holographic optical tweezers. Optics communications **207**(1-6), 169–175 (2002)
7. Dong, J., Valzania, L., Maillard, A., Pham, T.a., Gigan, S., Unser, M.: Phase retrieval: From computational imaging to machine learning: A tutorial. IEEE Signal Processing Magazine **40**(1), 45–57 (2023)
8. Emiliani, V., Cohen, A.E., Deisseroth, K., Häusser, M.: All-optical interrogation of neural circuits. Journal of Neuroscience **35**(41), 13917–13926 (2015)

9. Ersaro, N.T., Yalcin, C., Murray, L., Kabuli, L., Waller, L., Muller, R.: Fast non-iterative algorithm for 3d point-cloud holography. Opt. Express **31**(22), 36468–36485 (2023)
10. Eybposh, M.H., Cai, C., Deb, D., Schott, M.A.B., Ye, L., Both, G.J., Turaga, S.C., Rodriguez-Romaguera, J., Pégard, N.C.: Computer-generated holography using point cloud processing neural networks. In: Optica Imaging Congress (3D, COSI, DH, FLatOptics, IS, pcAOP). p. DW5A.4. Optica Publishing Group (2023)
11. Eybposh, M.H., Caira, N.W., Atisa, M., Chakravarthula, P., Pegard, N.C.: Deep-cgh: 3D computer-generated holography using deep learning. Opt. Express **28**(18), 26636–26650 (2020)
12. Eybposh, M.H., Curtis, V.R., Rodriguez-Romaguera, J., Pegard, N.C.: Advances in computer-generated holography for targeted neuronal modulation. Neurophotonics **9**, 041409 (2022)
13. Favre-Bulle, I.A., Scott, E.K.: Optical tweezers across scales in cell biology. Trends in cell biology **32**(11), 932–946 (2022)
14. Gerchberg, R.W.: A practical algorithm for the determination of phase from image and diffraction plane pictures. Optik, 35, 237-246. (1972)
15. Goodman, J.W.: Introduction to Fourier optics. Roberts and Company publishers (2005)
16. He, Z., Sui, X., Jin, G., Cao, L.: Progress in virtual reality and augmented reality based on holographic display. Applied optics **58**(5), A74–A81 (2019)
17. Keir, C., Steven, M.: Optical trapping. Review of scientific instruments **75**(9), 2787–2809 (2004)
18. Leach, J., Sinclair, G., Jordan, P., Courtial, J., Padgett, M.J., Cooper, J., Laczik, Z.J.: 3D manipulation of particles into crystal structures using holographic optical tweezers. Optics Express **12**(1), 220–226 (2004)
19. Ma, X., Qin, C., You, H., Ran, H., Fu, Y.: Rethinking network design and local geometry in point cloud: A simple residual MLP framework. arXiv preprint arXiv:2202.07123 (2022)
20. Marshel, J.H., Kim, Y.S., Machado, T.A., Quirin, S., Benson, B., Kadmon, J., Raja, C., Chibukhchyan, A., Ramakrishnan, C., Inoue, M., et al.: Cortical layer–specific critical dynamics triggering perception. Science **365**(6453), eaaw5202 (2019)
21. Matsushima, K.: Introduction to computer holography: creating computer-generated holograms as the ultimate 3D image. Springer International Publishing (2020)
22. Matsushima, K., Shimobaba, T.: Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields. Opt. Express **17**(22), 19662–19673 (Oct 2009)
23. Pégard, N.C., Mardinly, A.R., Oldenburg, I.A., Sridharan, S., Waller, L., Adesnik, H.: Three-dimensional scanless holographic optogenetics with temporal focusing (3D-shot). Nature communications **8**(1), 1228 (2017)
24. Peng, Y., Choi, S., Padmanaban, N., Wetzstein, G.: Neural holography with camera-in-the-loop training. ACM Transactions on Graphics **39**(6), 1–14 (2020)
25. Pi, D., Liu, J., Wang, Y.: Review of computer-generated hologram algorithms for color dynamic holographic three-dimensional display. Light: Science & Applications **11**(1), 231 (2022)
26. Qi, C.R., Su, H., Mo, K., Guibas, L.J.: Pointnet: Deep learning on point sets for 3D classification and segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 652–660 (2017)

27. Qi, C.R., Yi, L., Su, H., Guibas, L.J.: Pointnet++: Deep hierarchical feature learning on point sets in a metric space. Advances in neural information processing systems **30** (2017)
28. Situ, G.: Deep holography. Light: Advanced Manufacturing **3**(2), 1 (2022)
29. Wang, K., Song, L., Wang, C., Ren, Z., Zhao, G., Dou, J., Di, J., Barbastathis, G., Zhou, R., Zhao, J., et al.: On the use of deep learning for phase recovery. Light: Science & Applications **13**(1), 4 (2024)
30. Wu, J., Liu, K., Sui, X., Cao, L.: High-speed computer-generated holography using an autoencoder-based deep neural network. Opt. Lett. **46**(12), 2908–2911 (2021)
31. Wu, Y., Wang, J., Chen, C., Liu, C.J., Jin, F.M., Chen, N.: Adaptive weighted Gerchberg-Saxton algorithm for generation of phase-only hologram with artifacts suppression. Opt. Express **29**(2), 1412–1427 (Jan 2021)
32. Zhang, C., Wu, F., Zhou, J., Wei, S.: Non-iterative phase hologram generation for color holographic display. Opt. Express **30**(1), 195–209 (2022)
33. Zhang, J., Pégard, N., Zhong, J., Adesnik, H., Waller, L.: 3D computer-generated holography by non-convex optimization. Optica **4**(10), 1306–1313 (2017)
34. Zhang, Y., Zhang, M., Liu, K., He, Z., Cao, L.: Progress of the computer-generated holography based on deep learning. Applied Sciences **12**(17), 8568 (2022)