Supplementary: Assessing Learned Models for Phase-only Hologram Compression

1 Experimental Design and Training Protocol

To validate the effectiveness of neural implicit representation in hologram compression, this study systematically compares 3 approaches: 1. **vanilla MLP** as the baseline model; 2. **FilmSIREN**; 3. **SIREN**. Experiments utilize the central region of 512×512 pixel in the tested phased-only holograms, with training samples generated through a patch sampling strategy producing four patch sizes ($3 \times 64 \times 64$ to $3 \times 160 \times 160$) at 25% overlap through 10000 epochs. Each training with the weights learned from the previous patch. All models employ the Adam optimizer (initial learning rate: 1e-4) with a step decay scheduler (gamma=0.5 every 5,000 iterations), MSE loss function, and early stopping (threshold: 9e-6), trained on an NVIDIA RTX 4070 platform. Evaluation metrics include PSNR/SSIM, and model parameter count. Additionally, for reconstructed hologram we add LIPIS to evaluate. With each configuration tested all patches to compute mean values and statistical deviations. By balancing model complexity and patch dimensions, the experiments achieved a compression ratio of approximately 0.4, demonstrating efficient parameter utilization.

2 Model Architectures

2.1 vanilla MLP

Storage Mechanism:

σ represent ReLU activation function, $W_l \in \mathbb{R}^{d_l \times d_{l-1}}$ and $b_l \in \mathbb{R}^{d_l}$ are weight matrix and bias respectively.

$$\mathcal{F}_{\text{MLP}}(x, y) = W_n \sigma\left(\cdots W_2 \sigma(W_1 \phi(x, y) + b_1) \cdots\right) + b_n \tag{1}$$

Positional Encoding: $L = 10$: Number of frequency bands in positional encoding

$$\phi(x, y) = \bigoplus_{k=0}^{L-1} \left[\sin(2^k \pi x), \cos(2^k \pi x), \sin(2^k \pi y), \cos(2^k \pi y)\right] \tag{2}$$

2.2 SIREN [Sitzmann et al. 2020]

Storage Mechanism: ω_0: Global frequency scaling factor, n_{in}: Input dimension, $\mathcal{U}(a, b)$: Uniform distribution sampling.

$$\Phi_{\text{SIREN}}(x, y) = W_n \circ \sin \circ \cdots \circ W_2 \circ \sin \circ W_1(x, y) \tag{3}$$

Initialization Constraints: first layer

$$w_{ij} \sim \mathcal{U}\left(-\sqrt{\frac{6}{n_{\text{in}}}}, \sqrt{\frac{6}{n_{\text{in}}}}\right) \tag{4}$$

other sine layers

$$\omega_0 = 30 \tag{5}$$

$$w_{ij} \sim \mathcal{U}\left(-\frac{1}{n_{\text{in}}}, \frac{1}{n_{\text{in}}}\right) \tag{6}$$

This strategy supports the fidelity compression of high-frequency features of holograms while maintaining phase continuity by dynamically adjusting the parameter scales of each layer.

2.3 FilmSIREN [Chan et al. 2021]

Storage Architecture: L: Number of FiLM layers, d: Hidden dimension size, $\omega_l \in \mathbb{R}^d$: Frequency modulation vector for layer l, $\phi_l \in \mathbb{R}^d$: Phase shift vector for layer l.

$$\mathcal{F}_{\text{FILM}}(x, z) = \text{Sigmoid}\left(W_f \psi(x, \mathcal{M}(z))\right) \tag{7}$$

Frequency Modulation:

$$\psi(x, m) = \prod_{l=1}^{L} \sin(\omega_l \cdot W_l x + \phi_l) \tag{8}$$

Modulation Network:

$$\mathcal{M} : z \mapsto (\omega, \phi) \in R^{2Ld} \tag{9}$$

2.4 TAESD [Bohan 2023]

Compression Pipeline:

$$\mathcal{E}(I) = \text{ConvBlock}^3\left(\text{Downsample}^3(I)\right) \in \mathbb{R}^{16 \times 64 \times 64} \tag{10}$$

$$\mathcal{D}(z) = \text{ConvBlock}^3\left(\text{Upsample}^3(z)\right) \tag{11}$$

Latent Space Constraints:

$$z = \text{Clamp}\left(\mathcal{E}(I)/3\right) \times 3 \in [-3, 3]^{16 \times 64 \times 64} \tag{12}$$

3 Additional Experimental

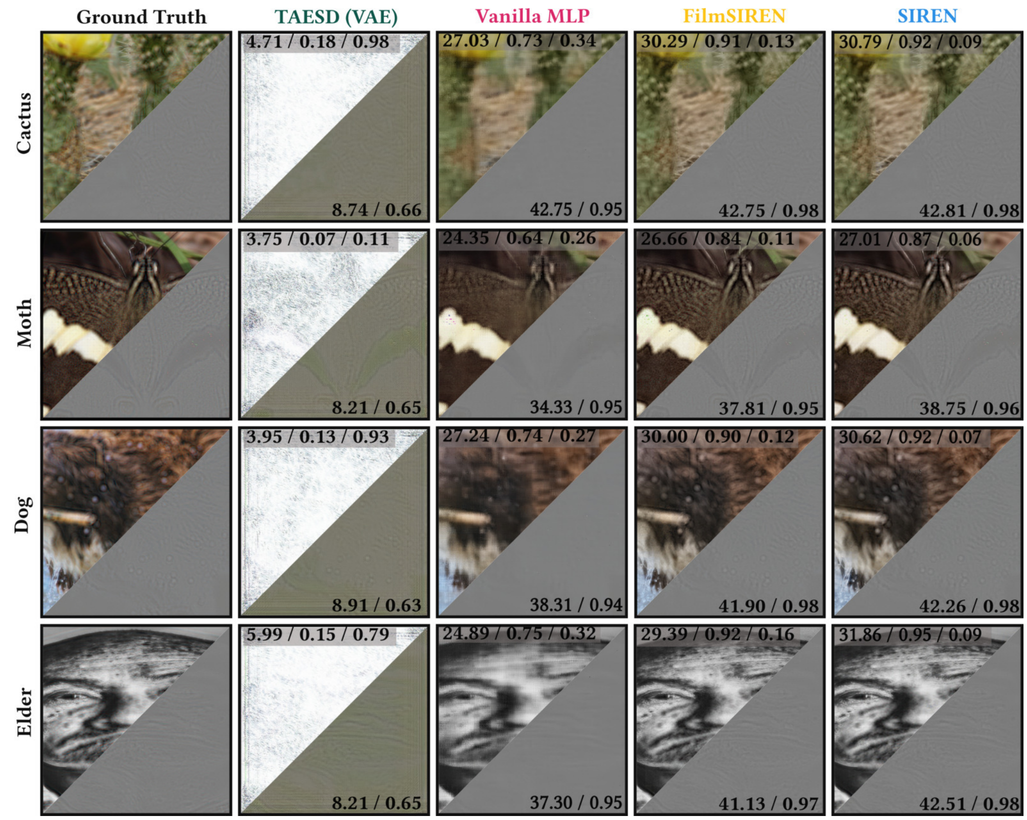

Figure 1: Four supplementary holograms are compressed to replicate the identical settings of the teaser image. (Source image: mark byzewski) (Source image: Charles J. Sharp (1951–)) (Source image: Matt Jones mattrobinjones) (Source image: Ton Henry binbin127)

In addition to the teaser image, we further validated the effectiveness of our approach by compressing and reconstructing four distinct holograms (cactus, moth, dog, and elderly face) under identical settings. **TAESD** exhibited significantly inferior performance in both reconstructed hologram quality (average PSNR: 4.6 dB, SSIM: 0.13, LIPIS: 0.70). **vanilla MLP** showed effective compression ability, achieving reconstructed metrics of PSNR=25.87 dB, SSIM=0.72, and LIPIS=0.30. Moreover, **FilmSIREN** and **SIREN**, further enhanced reconstruction precision and perceptual alignment: **FilmSIREN** attained PSNR=29.09 dB, SSIM=0.89, LIPIS=0.13, while SIREN achieved state-of-the-art performance with PSNR=30.07 dB, SSIM=0.915, and LIPIS=0.07. Notably, **SIREN** consistently outperformed others in preserving more accuracy information (e.g., moth wing patterns, facial wrinkles) and dynamic details, aligning with conclusions from our poster analysis. This confirms the high potential of implicit networks in maintaining holographic phase-amplitude coherence under compression.

References

Ollin Boer Bohan. 2023. Tiny Autoencoder for Stable Diffusion. https://github.com/madebyollin/taesd. Accessed: June 2025.

Eric R Chan, Marco Monteiro, Petr Kellnhofer, Jiajun Wu, and Gordon Wetzstein. 2021. pi-gan: Periodic implicit generative adversarial networks for 3d-aware image synthesis. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 5799–5809. doi:10.1109/CVPR46437.2021.00574

Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. 2020. Implicit neural representations with periodic activation functions. *Advances in neural information processing systems* 33 (2020), 7462–7473. doi:10.48550/arXiv.2006.09661