MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

Quantum Implicit Neural Compression

Fujihashi, Takuya; Koike-Akino, Toshiaki

TR2026-022 February 19, 2026

Abstract

Signal compression based on implicit neural representation (INR) is an emerging technique to represent multimedia signals with a small number of bits. While INR-based signal compression achieves high-quality reconstruction for relatively low-resolution signals, the accuracy of high- frequency details is significantly degraded with a small model. To improve the compression efficiency of INR, we introduce quantum INR (quINR), which leverages the exponentially rich expressivity of quantum neural networks for data compression. Evaluations using some benchmark datasets show that the proposed quINR-based compression could improve rate-distortion performance in image compression compared with traditional codecs and classic INR-based coding methods, up to 1.2dB gain.

Springer Nature 2026

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Quantum Implicit Neural Compression

Takuya Fujihashi[1][0000−0002−6960−0122] and
Toshiaki Koike-Akino[2][0000−0002−2578−5372]

[1] The University of Osaka
Yamadaoka 1-5, Suita, Osaka 565-0871, Japan
tfuji@ist.osaka-u.ac.jp
[2] Electric Research Laboratories (MERL)
201 Broadway, Cambridge, MA 02139, USA
koike@merl.com

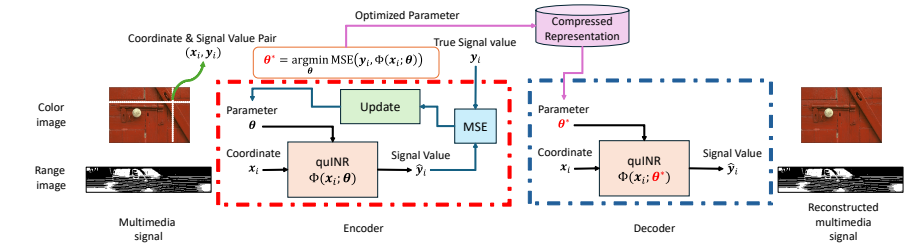
Abstract. Signal compression based on implicit neural representation (INR) is an emerging technique to represent multimedia signals with a small number of bits. While INR-based signal compression achieves high-quality reconstruction for relatively low-resolution signals, the accuracy of high-frequency details is significantly degraded with a small model. To improve the compression efficiency of INR, we introduce quantum INR (quINR), which leverages the exponentially rich expressivity of quantum neural networks for data compression. Evaluations using some benchmark datasets show that the proposed quINR-based compression could improve rate-distortion performance in image compression compared with traditional codecs and classic INR-based coding methods, up to 1.2dB gain.

Keywords: Quantum Machine Learning · Neural Implicit Representation · Data Compression.

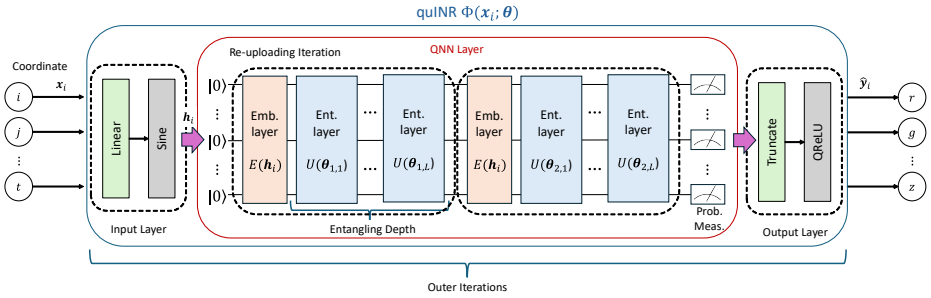
1 Background

Representing multimedia signals (such as images and video frames) in a compact format is an important task for communicating and storing such signals [36]. Implicit neural representation (INR) is an emerging, memory-efficient format [34] for compressing data. Most INR architectures exploit a small and simple multilayer perceptron (MLP)-based neural network (NN) architecture and train the coordinate-to-value mappings using the target signals. Since the trained NN architecture can obtain the signal values of all coordinates by using a simple feedforward process, recent studies have used the trained INR architectures for multimedia signal compression. For example, COmpression with Implicit Neural representations (COIN) variants [6, 7, 16] have been designed for image coding, and Neural Representations for Videos (NeRV) variants [3, 17, 19, 24, 18] and COOL-CHIC variants [20, 21, 39, 22, 14] have considered video coding.

A key issue in such INR-based signal compression methods is the inaccurate representation of high-frequency details in a small MLP-based NN architecture. Some studies have developed sinusoidal coding [25] and activation functions [34] to approximate high-frequency details even in a small NN architecture.



(a) Encoding and decoding procedures using quINR



(b) Architecture of quINR

Fig. 1: Overview of the proposed scheme for data compression using a hybrid quantum-classical implicit neural representation.

In this paper, we introduce a new hybrid quantum-classical INR architecture, namely, quantum INR (quINR), for signal compression. The proposed quINR integrates feature embedding and quantum neural network (QNN) [9] for training the coordinate-to-value mapping. Since QNN is a promising technique for accelerating computation and saving parameters, our quINR may have the potential to reconstruct accurate high-frequency representations with fewer parameters.

Experiments using the range image (RI) dataset in the KITTI Light Detection and Ranging (LiDAR) point cloud [13] and Kodak color image dataset [8] show that the proposed quINR-based compression can provide better coding efficiency compared to the existing compression methods.

2 Related Work

2.1 Implicit Neural Compression

INR compresses signals by overfitting a compact coordinate-to-value network and transmitting its weights as the bitstream. Early INR variants for images encode pixel coordinates with sinusoidal or periodic activations to better capture high-frequency content, e.g., SIREN and Neural Radiance Fields (NeRF)-style

sinusoidal encodings [6, 7]. On images, COIN and COIN++ are the pioneering works and pursue this paradigm by directly mapping 2D coordinates to pixel intensities and extending to multiple modalities [6, 7]. To improve upon the COIN paradigm for both image and video data, COOL-CHIC and C3 [20, 14] propose hierarchical/coordinate-based codecs that achieve competitive rate-distortion (R-D) with lightweight decoders. For large-scale videos, NeRV has been proposed to generate multiple video frames from frame indices via a single network, reducing storage and enabling smooth temporal interpolation [3]. Subsequent works such as Hierarchically-encoded NeRV (HiNeRV) and Neural Video Representation Compression (NVRC) improve R-D and scalability through hierarchical encoding and representation compression, while conditional decoders further exploit temporal redundancy [17, 18]. INR-based compression has also been explored for 3D imaginary [19, 16, 11] and other modalities [26, 35, 10]. For example, in [16], the INR architecture tailored to LiDAR RI has been proposed to preserve high-frequency details under rate budgets. Specifically, it improves coding efficiency by decomposing each RI into structurally distinct components, i.e., mask and depth images, and encoding them separately using dedicated INR networks.

Our work targets the same goal—accurate reconstruction at a low budget—but departs from purely classical MLPs by introducing a quantum-classical architecture to increase expressivity under tight budget constraints.

2.2 Quantum Neural Network

QNN [1, 32, 9] is an emerging paradigm that exploits quantum physics for neural network design, where classical data and weight values are embedded into a variational quantum circuit to control measurement outcomes. QNN provides universal approximation property [29] and exponentially rich expressibity [33]. In addition, it is analytically differentiable, enabling stochastic gradient optimization [31].

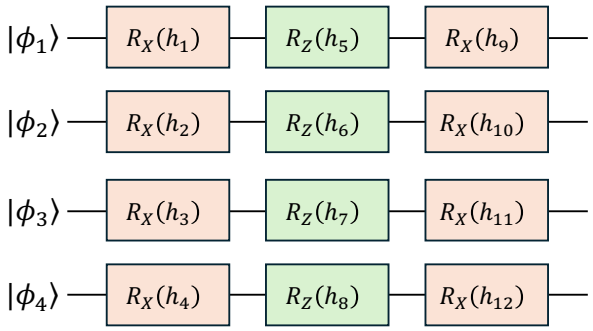
Various frameworks were migrated into a quantum domain: autoencoders [30]; graph neural networks [42]; generative adversarial networks [23, 5]; contrastive learning [2]; diffusion models [27, 38]. As QNN is extremely parameter-efficient, it was applied to fine-tuning [4, 15], learnable activation functions [12], and implicit representation [37, 40].

Our study is the first attempt to demonstrate the potential of the QNN architecture for signal compression. Specifically, we design signal encoding and decoding procedures using QNN architecture, inspired by an existing implicit representation [40]. Experiments using image and LiDAR datasets show that the proposed quINR yields better reconstruction quality at a small data size.

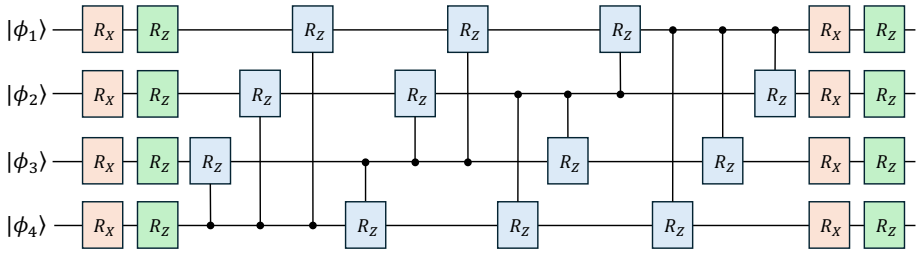
3 Quantum INR for Data Compression

3.1 Encoding and Decoding Process

Fig. 1 (a) shows the end-to-end operations of the quINR-based encoder and decoder. Given the target multimedia signal, we construct a dataset $\mathcal{D} = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}$



(a) Folded-angle embedding layer



(b) Entangling layer

Fig. 2: Exemplar architecture of QNN layer.

for training the quINR $\Phi(\boldsymbol{x}_i;\boldsymbol{\theta})$. Here, $\boldsymbol{x}_i \in \mathbb{R}^{N_{\text{in}}}$ is the ith coordinate, $\boldsymbol{y}_i \in \mathbb{R}^{N_{\text{out}}}$ is its corresponding signal value, and $\boldsymbol{\theta}$ is the learnable parameter set.

In the encoding process, the proposed quINR $\Phi(\boldsymbol{x}_i;\boldsymbol{\theta})$ is trained to obtain the optimized parameter set $\boldsymbol{\theta}$ to express the coordinate-to-value relationships contained in the dataset $\mathcal{D}$. Here, we use the mean squared error (MSE) as the loss function to obtain the optimized parameters $\boldsymbol{\theta}$:

$$\boldsymbol{\theta}^\star = \arg\min_{\boldsymbol{\theta}} \mathsf{MSE}\big(\boldsymbol{y}_i, \Phi(\boldsymbol{x}_i;\boldsymbol{\theta})\big). \tag{1}$$

This training process is coordinate-wise, i.e., the parameters are trained to obtain a mapping from each coordinate $\boldsymbol{x}_i$ to their corresponding signal values $\boldsymbol{y}_i$. The well-trained parameters $\boldsymbol{\theta}^\star$ after this encoding process are stored in storage or transmitted to the decoder as the lightweight format of the target signal.

The decoder uses the parameters $\boldsymbol{\theta}^\star$ for reproducing the target signal through the forward process of the quINR $\Phi(\boldsymbol{x}_i;\boldsymbol{\theta}^\star)$. The target signal $\hat{\boldsymbol{y}}_i$ is reconstructed by feeding the coordinates $\boldsymbol{x}_i$ to the quINR architecture $\Phi(\boldsymbol{x}_i;\boldsymbol{\theta}^\star)$. Likewise, the encoding process sequentially feeds coordinates $\boldsymbol{x}_i$ to the quINR to collect all

estimated signal values $\hat{\boldsymbol{y}}_i$, which are then reshaped to the shape of the target signal as $\hat{\boldsymbol{y}}$.

3.2 Model Architecture

Fig. 1 (b) shows the proposed INR architecture. The architecture takes the coordinates of the multimedia signals as inputs and generates the corresponding signal values as outputs. The quINR $\Phi(\boldsymbol{x}_i;\boldsymbol{\theta})$ is a hybrid quantum-classical architecture integrating QNN layers with a classical NN.

The input layer consists of a linear layer with a sinusoidal activation to obtain an embedding vector $\boldsymbol{h}_i \in \mathbb{R}^M$ from each coordinate pair $\boldsymbol{x}_i$ as follows:

$$\boldsymbol{h}_i = \sin(\omega_0 \boldsymbol{W}\boldsymbol{x}_i + \boldsymbol{b}), \tag{2}$$

where $\boldsymbol{W} \in \mathbb{R}^{M\times N_{\text{in}}}$ and $\boldsymbol{b} \in \mathbb{R}^M$ are trainable parameters of the linear layer and $\omega_0 = 30.0$ is a constant hyperparameter. The embedding vector $\boldsymbol{h}_i$ is then fed into the QNN layers. The QNN layers consist of embedding and entangling layers, as shown in Fig. 2.

For embedding layer, we propose folded-angle embedding to encode an arbitrary size of embedding vector $\boldsymbol{h}_i$ into a finite number of qubits. The conventional angle embedding has a restriction that the number of qubits must be no lower than the size of the embedding vector, while the amplitude embedding provides too small quantum space having little expressivity. To make the QNN compact yet expressive, the folded-angle embedding uses alternating R_X and R_Z gates to pack more angle parameters. Fig. 2 (a) shows an example of 3-folded embedding with four qubits to encode twelve variables.

The entangling layer is based on a parameterized quantum circuit in [33]. Specifically, the parameterized circuit sequentially carries out R_Z and R_X rotation gates for each qubit, two-qubit controlled Z-rotation (CRZ) for each two-qubit combination, and finally uses Z-rotation and X-rotation. Here, each rotation gate is controlled based on the parameter set $\boldsymbol{\theta}$. A few number of entangling layers are sequentially cascaded. These embedding and entangling layers are iterated over a few layers, with a shuffled extension of the data re-uploading trick [29].

Finally, we measure the probability value of 2^{N_q} quantum states. The output layer selects the last N_{out} state with the activation function of quantum rectified linear unit (QReLU) [28], regarded as the estimated signal value $\hat{\boldsymbol{y}}_i$. The above structure can be further iterated over layers to improve the capacity. Here, the number of the required quantum shots is approximately $O(2^{N_q})$. Even for high image resolutions, the complexity of the proposed quINR may remain practical, as it represents clean multimedia signals using a small QNN architecture, i.e., with a N_q, as demonstrated in the evaluations.

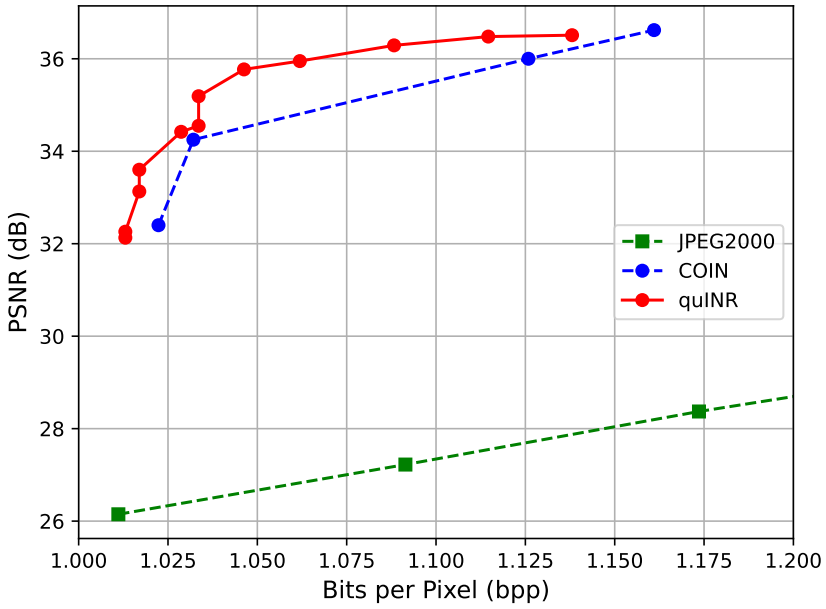


Fig. 3: PSNR vs. bpp for RI.

4 Experiments

4.1 Settings

Datasets: In this paper, we consider grayscale and color images to discuss the potential of the proposed quINR architecture. For the grayscale image, we use LiDAR RI [41] derived from the KITTI point cloud dataset [13]. RI can be mapped from three-dimensional (3D) Cartesian coordinate x-y-z to spherical coordinate ρ-ϕ-θ, and then mapped to the two-dimensional (2D) image coordinate system with the resolution of 1024×64 pixels. Here, each pixel value of RI is the distance ρ with floating-point precision. Specifically, we use LiDAR sequence 00-00 for comparison. For the color image, we perform experiments on the Kodak image dataset [8], which consists of 24 images of 768×512 pixels. We selected one image from the dataset, Kodim02.

Metric: Regarding the metrics for the decoded color and grayscale images, we use peak signal-to-noise ratio (PSNR) for comparison. Given an original image

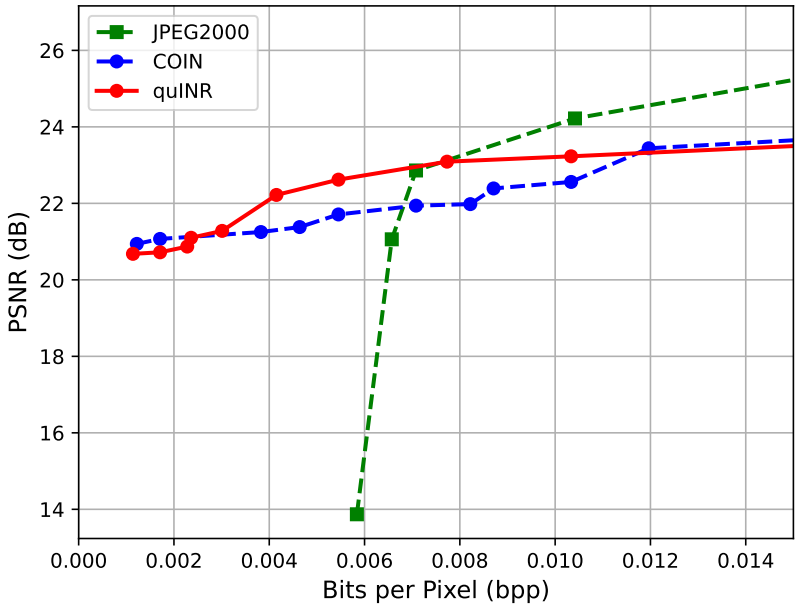


Fig. 4: PSNR vs. bpp for Kodak color image.

I and a reconstructed image $\hat{I}$, MSE can be defined as:

$$\text{MSE} = \frac{1}{WH} \sum_{i=1}^{H} \sum_{j=1}^{W} \left(I(i,j) - \hat{I}(i,j) \right)^2 . \tag{3}$$

PSNR is then obtained as:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right), \tag{4}$$

where MAX represents the maximum pixel value of the image.

Baseline: We compare with baseline methods: Joint Photographic Experts Group 2000 (JPEG2000) and COIN [6]. JPEG2000 is a typical image compression method, requiring conversion to 8-bit precision in advance for compression. COIN is an INR–based image compression baseline. The INR architecture is trained to obtain a direct mapping from the 2D pixel coordinate to the pixel value of grayscale and color images.

Implementation: NNs for COIN and our proposed quINR architectures are implemented, trained, and evaluated using PyTorch 2.0 with Python 3.9. The

quantum circuit simulations are performed by PennyLane 0.35. We use Adaptive moment estimation (Adam) with decoupled weight decay (AdamW) for optimizer with 1e-1 learning rate for both classical and quantum architectures.

4.2 Performance Comparison

Fig. 3 shows the PSNR performance for RI as a function of bit per pixel (bpp). Here, we vary hyperparameters such as embedding size M to show the Pareto frontier curves for each baseline. The results show that the proposed quINR achieves better image quality at a small bpp regime, however, the quality improvement is saturated at a large bpp regime compared with COIN architecture. It suggests that the proposed quINR may have the potential to reconstruct clean signals at band-limited and storage-limited environments.

Fig. 4 shows the PSNR performance for the Kodak color image as a function of bpp. For this case, JPEG2000 offers much better performance than RI case as the target signal is a natural image. Nevertheless, the proposed quINR architecture can be better than the other baselines in low to medium compression regimes with up to 1.2dB gain.

5 Conclusion

This paper highlights the potential of quantum techniques in advancing multimedia signal compression. The proposed quINR architecture demonstrated good PSNR performance, particularly in compressing LiDAR RI, leveraging the expressive power of QNN. Nevertheless, its rate-distortion performance for color image compression was limited, indicating the need for further improvements, e.g., with quantum network architecture search (NAS) and distillation.

Acknowledgments. T. Fujihashi's work was supported by JSPS KAKENHI Grant Number and JP22H03582.

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

References

1. Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., Lloyd, S.: Quantum machine learning. Nature **549**(7671), 195–202 (2017)
2. Chen, C.S., Tsai, A.H.W., Huang, S.C.: Quantum multimodal contrastive learning framework. arXiv preprint arXiv:2408.13919 (2024)
3. Chen, H., He, B., Wang, H., Ren, Y., Lim, S.N., Shrivastava, A.: NeRV: Neural representations for videos. In: NeurIPS (2021)
4. Chen, Z., Dangovski, R., Loh, C., Dugan, O.M., Luo, D., Soljacic, M.: QuanTA: Efficient high-rank fine-tuning of LLMs with quantum-informed tensor adaptation. In: The Thirty-eighth Annual Conference on Neural Information Processing Systems (2024), https://openreview.net/forum?id=EfpZNpkrm2

5. Dallaire-Demers, P.L., Killoran, N.: Quantum generative adversarial networks. Physical Review A **98**(1), 012324 (2018)
6. Dupont, E., Golinski, A., Alizadeh, M., Teh, Y.W., an an, A.D.: COIN: COmpression with implicit neural representations. In: ICLR Workshop Neural Compression (2021)
7. Dupont, E., Loya, H., Alizadeh, M., Goliński, A., Teh, Y.W., Doucet, A.: COIN++: neural compression across modalities. TMLR **2022**(11), 1–26 (2022)
8. Eastman Kodak Company: Kodak lossless true color image suite (1999)
9. Farhi, E., Neven, H.: Classification with quantum neural networks on near term processors. arXiv preprint arXiv:1802.06002 (2018)
10. Fujihashi, T., Kato, S., Koike-Akino, T.: Implicit neural representation for low-overhead graph-based holographic-type communications. In: ICASSP. pp. 2825–2829 (2024)
11. Fujihashi, T., Koike-Akino, T.: Fv-nerv: Neural compression for free viewpoint videos. In: Advances in Neural Information Processing Systems Workshop on Machine Learning and Compression. pp. 1–4 (2024)
12. Fujihashi, T., Kuwabara, A., Koike-Akino, T.: QKAN-GS: Quantum-empowered 3D gaussian splatting. In: ACM Int. Conf. Multimedia International Workshop on Application-driven Point Cloud Processing and 3D Vision. pp. 1–5 (2025)
13. Geiger, A., Lenz, P., Stiller, C., Urtasun, R.: Vision meets robotics: The KITTI dataset. Int. J. Robot. Res. **32**(11), 1231–1237 (2013)
14. Kim, H., Bauer, M., Theis, L., Schwarz, J.R., Dupont, E.: C3: High-performance and low-complexity neural compression from a single image or video. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9347–9358 (2024)
15. Koike-Akino, T., Tonin, F., Wu, Y., Candogan, L.N., Cevher, V.: Quantum-PEFT: Ultra parameter-efficient fine-tuning. In: Workshop on Efficient Systems for Foundation Models II@ ICML2024 (2024)
16. Kuwabara, A., Kato, S., Koike-Akino, T., Fujihashi, T.: Jscc-aided inr for high-frequency detail preservation in lidar. IEEE Open Journal of the Communications Society **6**, 6352–6367 (2025)
17. Kwan, H.M., Gao, G., Zhang, F., Gower, A., Bull, D.: HiNeRV: video compression with hierarchical encoding-based neural representation. In: Proceedings of the 37th International Conference on Neural Information Processing Systems. pp. 72692–72704 (2023)
18. Kwan, H.M., Gao, G., Zhang, F., Gower, A., Bull, D.: NVRC: Neural video representation compression. Advances in Neural Information Processing Systems **37**, 132440–132462 (2024)
19. Kwan, H.M., Zhang, F., Gower, A., Bull, D.: Immersive video compression using implicit neural representations. In: Picture Coding Symposium. pp. 1–5 (2024)
20. Ladune, T., Philippe, P., Henry, F., Clare, G., Leguay, T.: COOL-CHIC: Coordinate-based low complexity hierarchical image codec. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 13515–13522 (2023)
21. Leguay, T., Ladune, T., Philippe, P., Clare, G., Henry, F., Déforges, O.: Low-complexity overfitted neural image codec. In: IEEE 25th International Workshop on Multimedia Signal Processing. pp. 1–6 (2023)
22. Li, H., Zhang, X.: Human–machine collaborative image compression method based on implicit neural representations. IEEE Journal on Emerging and Selected Topics in Circuits and Systems **14**(2), 198–208 (2024)
23. Lloyd, S., Weedbrook, C.: Quantum generative adversarial learning. Physical review letters **121**(4), 040502 (2018)

24. Maiya, S.R., Girish, S., Ehrlich, M., Wang, H., Lee, K., Poirson, P., Wu, P., Wang, C., Shrivastava, A.: Nirvana: Neural implicit representations of videos with adaptive networks and autoregressive patch-wise modeling. In: CVPR. pp. 14378–14387 (jun 2023)
25. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: NeRF: representing scenes as neural radiance fields for view synthesis. Communications of the ACM **65**(1), 99–106 (2021)
26. Ozawa, Y., Kuwabara, A., Ogura, T., Fujihashi, T., Saruwatari, S., Watanabe, T.: Multi-channel vibrotactile signal compression based on implicit neural representation. In: IEEE World Haptics Conference. pp. 133–139 (2025)
27. Parigi, M., Martina, S., Caruso, F.: Quantum-noise-driven generative diffusion models. Advanced Quantum Technologies p. 2300401
28. Parisi, L., Neagu, D., Ma, R., Campean, F.: Quantum ReLU activation for convolutional neural networks to improve diagnosis of parkinson's disease and COVID-19. Expert Systems with Applications **187**, 1–17 (2022)
29. Pérez-Salinas, A., Cervera-Lierta, A., Gil-Fuster, E., Latorre, J.I.: Data reuploading for a universal quantum classifier. Quantum **4**, 226 (2020)
30. Romero, J., Olson, J.P., Aspuru-Guzik, A.: Quantum autoencoders for efficient compression of quantum data. Quantum Science and Technology **2**(4), 045001 (2017)
31. Schuld, M., Bergholm, V., Gogolin, C., Izaac, J., Killoran, N.: Evaluating analytic gradients on quantum hardware. Physical Review A **99**(3), 032331 (2019)
32. Schuld, M., Sinayskiy, I., Petruccione, F.: An introduction to quantum machine learning. Contemporary Physics **56**(2), 172–185 (2015)
33. Sim, S., Johnson, P.D., Aspuru-Guzik, A.: Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. Advanced Quantum Technologies **2**(12), 1–18 (2019)
34. Sitzmann, V., Martel, J.N.P., Bergman, A.W., Lindell, D.B., Wetzstein, G.: Implicit neural representations with periodic activation functions. In: NeurIPS. pp. 1–12 (2020)
35. Su, K., Chen, M., Shlizerman, E.: INRAS: Implicit neural representation for audio scenes. Advances in Neural Information Processing Systems **35**, 8144–8158 (2022)
36. Sullivan, G.J., Wiegand, T.: Rate-distortion optimization for video compression. IEEE Signal Processing Magazine **15**(6), 74–90 (1998)
37. Yang, Y., Sun, M.: A quantum-powered photorealistic rendering. arXiv preprint arXiv:2211.03418 (2022)
38. Zhang, B., Xu, P., Chen, X., Zhuang, Q.: Generative quantum machine learning via denoising diffusion probabilistic models. Physical Review Letters **132**(10), 100602 (2024)
39. Zhang, G., Zhang, X., Tang, L.: Enhanced quantified local implicit neural representation for image compression. IEEE Signal Processing Letters **30**, 1742–1746 (2023)
40. Zhao, J., Qiao, W., Zhang, P., Gao, H.: Quantum implicit neural representations. In: Proceedings of the 41st International Conference on Machine Learning. pp. 1–17 (2024)
41. Zhao, L., Ma, K.K., Liu, Z., Yin, Q., Chen, J.: Real-time scene-aware lidar point cloud compression using semantic prior representation. IEEE Trans. Circuits Syst. Video Technol. **32**(8), 5623–5637 (2022)
42. Zheng, J., Gao, Q., Lü, Y.: Quantum graph convolutional neural networks. In: 2021 40th Chinese Control Conference (CCC). pp. 6335–6340. IEEE (2021)