Quantum Implicit Neural Compression

Fujihashi, Takuya; Koike-Akino, Toshiaki

TR2025-024 March 05, 2025

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 im- prove 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.

AAAI Conference on Artificial Intelligence 2025

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

Quantum Implicit Neural Compression

Takuya Fujihashi[1], **Toshiaki Koike-Akino**[2],

[1]Osaka University, Suita, Osaka 565-0871, Japan
[2]Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA 02139, USA
Email: tfuji@ist.osaka-u.ac.jp, 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.

Background

Representing multimedia signals (such as images and video frames) in a compact format is an important task for communicating and storing such signals. Implicit neural representation (INR) is an emerging memory-efficient format to compress data. Most INR architectures exploit a small and simple multi-layer perceptron (MLP)-based neural network (NN) architecture and train the coordinate-to-value mappings using the target signals. For example, COmpression with Implicit Neural representations (COIN) (Dupont et al. 2021, 2022) has been designed for image coding, and Neural Representations for Videos (NeRV) (Chen et al. 2021) variants have considered 3D 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 (Mildenhall et al. 2021) and activation functions (Sitzmann et al. 2020) to approximate high-frequency details even in a small NN architecture. 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) (Farhi and Neven 2018) 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 (Geiger et al. 2013) and Kodak color image dataset (Eastman Kodak Company 1999) show that the proposed quINR-based compression can provide better coding efficiency compared to the existing compression methods.

Related Work

Implicit Neural Compression

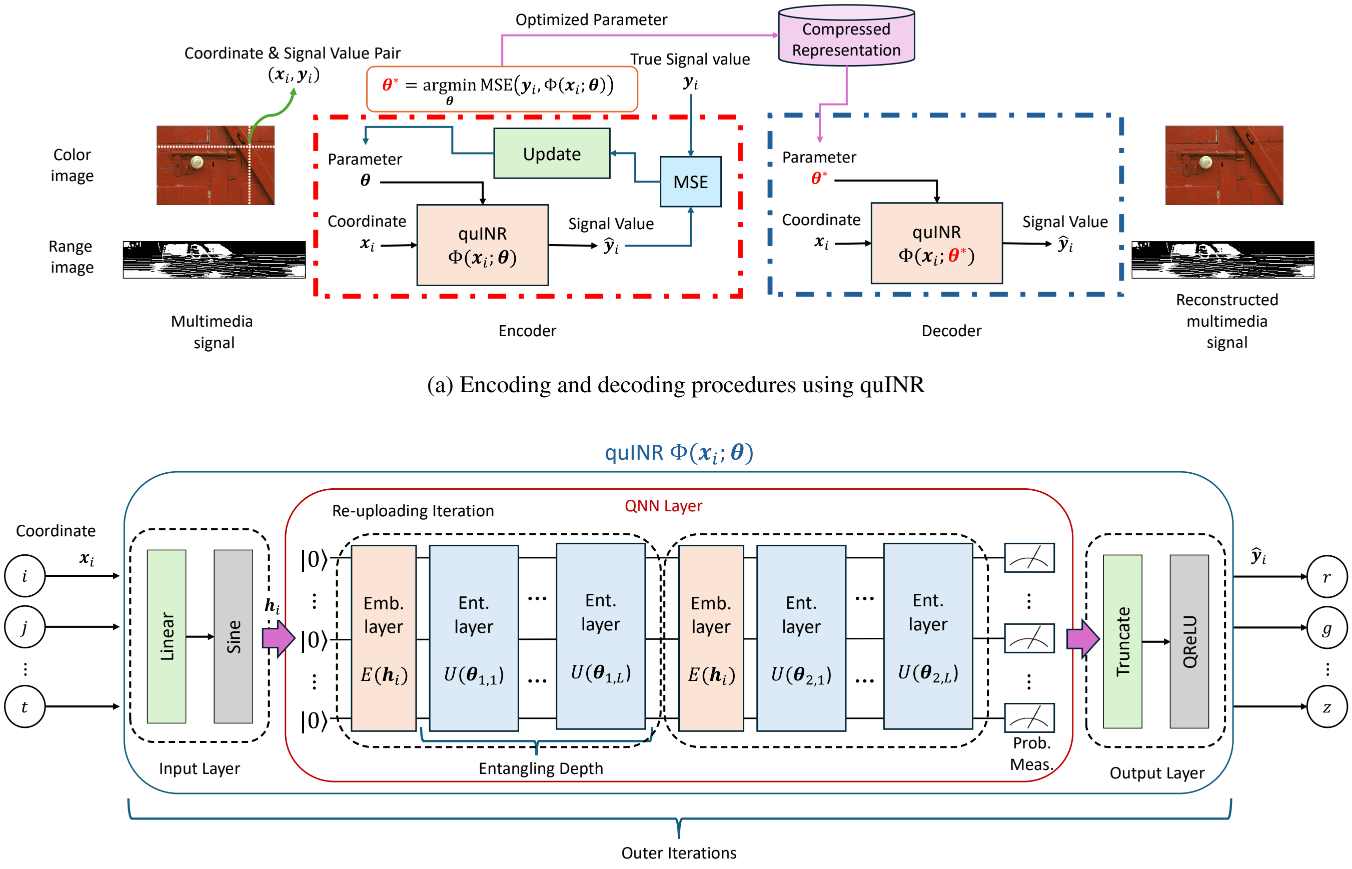
Recent studies exploit INR architectures for data compression by overfitting a small NN for a particular multi-dimensional sample. The INR architecture (Dupont et al. 2021, 2022) takes the pixel coordinate as input to reconstruct the corresponding pixel value. It was extended to video coding (Chen et al. 2021; Zhang et al. 2024b) by feeding the frame index for frame generation.

Quantum Neural Network

QNN (Biamonte et al. 2017; Schuld, Sinayskiy, and Petruccione 2015; Farhi and Neven 2018) is an emerging paradigm exploiting the quantum physics for neural network design, where classical data and weight values are embedded into a variational quantum circuit to control the measurement outcomes. QNN provides universal approximation property (Pérez-Salinas et al. 2020) and exponentially rich expressibity (Sim, Johnson, and Aspuru-Guzik 2019). In addition, it is analytically differentiable, enabling stochastic gradient optimization (Schuld et al. 2019).

Various frameworks were migrated into a quantum domain: autoencoders (Romero, Olson, and Aspuru-Guzik 2017); graph neural networks (Zheng, Gao, and Lü 2021); generative adversarial networks (Lloyd and Weedbrook 2018; Dallaire-Demers and Killoran 2018); contrastive learning (Chen, Tsai, and Huang 2024); diffusion models (Parigi, Martina, and Caruso; Zhang et al. 2024a). As QNN is extremely parameter-efficient, it was applied to fine-tuning (Chen et al. 2024; Koike-Akino et al. 2024) and implicit representation (Yang and Sun 2022; Zhao et al. 2024).

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 (Zhao et al. 2024). Experiments using image and



(a) Encoding and decoding procedures using quINR

(b) Architecture of quINR

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

LiDAR datasets show that the proposed quINR yields better reconstruction quality at a small data size.

Quantum INR for Data Compression

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 $\mathcal{D} = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}$ 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, it 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 $\hat{\boldsymbol{y}}$.

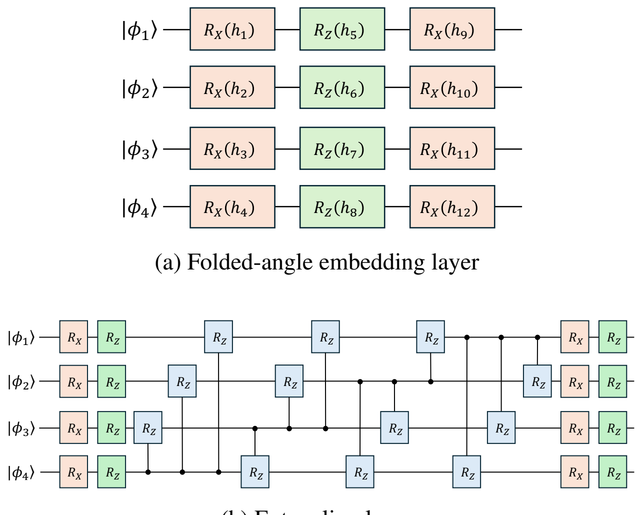
Model Architecture

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 hyperparam-



(a) Folded-angle embedding layer

(b) Entangling layer

Figure 2: Exemplar architecture of QNN layer.

eter. 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 (Sim, Johnson, and Aspuru-Guzik 2019). 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 (Pérez-Salinas et al. 2020).

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) (Parisi et al. 2022), 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.

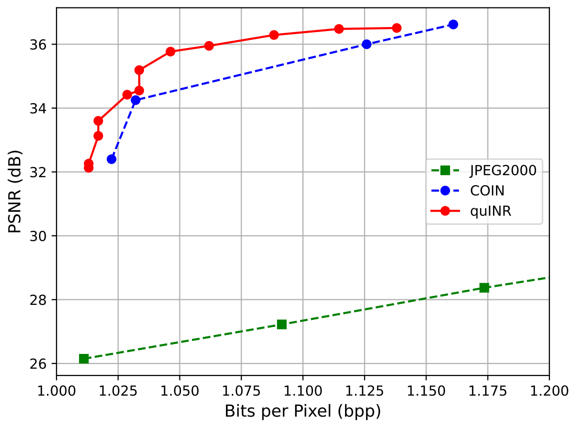


Figure 3: PSNR vs. bpp for RI.

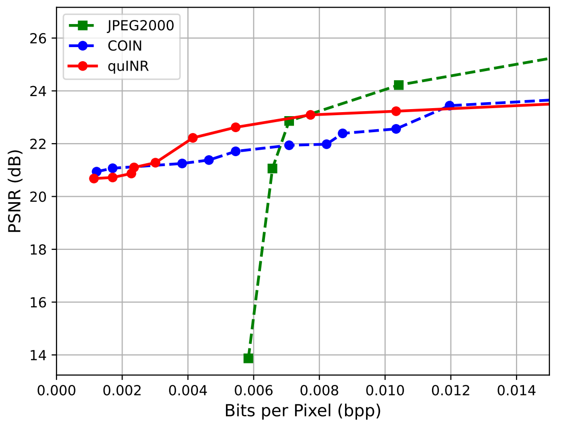


Figure 4: PSNR vs. bpp for Kodak color image.

Experiments

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 (Zhao et al. 2022) derived from the KITTI point cloud dataset (Geiger et al. 2013). 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 (Eastman Kodak Company 1999), 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 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 (Dupont et al. 2021). 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.

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.

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.

Acknowledgment

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

References

Biamonte, J.; Wittek, P.; Pancotti, N.; Rebentrost, P.; Wiebe, N.; and Lloyd, S. 2017. Quantum machine learning. *Nature*, 549(7671): 195–202.

Chen, C.-S.; Tsai, A. H.-W.; and Huang, S.-C. 2024. Quantum Multimodal Contrastive Learning Framework. *arXiv preprint arXiv:2408.13919*.

Chen, H.; He, B.; Wang, H.; Ren, Y.; Lim, S.-N.; and Shrivastava, A. 2021. NeRV: Neural Representations for Videos. In *NeurIPS*.

Chen, Z.; Dangovski, R.; Loh, C.; Dugan, O. M.; Luo, D.; and Soljacic, M. 2024. QuanTA: Efficient High-Rank Fine-Tuning of LLMs with Quantum-Informed Tensor Adaptation. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.

Dallaire-Demers, P.-L.; and Killoran, N. 2018. Quantum generative adversarial networks. *Physical Review A*, 98(1): 012324.

Dupont, E.; Golinski, A.; Alizadeh, M.; Teh, Y. W.; and an an, A. D. 2021. COIN: COmpression with Implicit Neural representations. In *ICLR Workshop Neural Compression*.

Dupont, E.; Loya, H.; Alizadeh, M.; Goliński, A.; Teh, Y. W.; and Doucet, A. 2022. COIN++: Neural compression across modalities. *TMLR*, 2022(11): 1–26.

Eastman Kodak Company. 1999. Kodak lossless true color image suite.

Farhi, E.; and Neven, H. 2018. Classification with quantum neural networks on near term processors. *arXiv preprint arXiv:1802.06002*.

Geiger, A.; Lenz, P.; Stiller, C.; and Urtasun, R. 2013. Vision meets robotics: The KITTI dataset. *Int. J. Robot. Res.*, 32(11): 1231–1237.

Koike-Akino, T.; Tonin, F.; Wu, Y.; Candogan, L. N.; and Cevher, V. 2024. Quantum-PEFT: Ultra parameter-efficient fine-tuning. In *Workshop on Efficient Systems for Foundation Models II@ ICML2024*.

Lloyd, S.; and Weedbrook, C. 2018. Quantum generative adversarial learning. *Physical review letters*, 121(4): 040502.

Mildenhall, B.; Srinivasan, P. P.; Tancik, M.; Barron, J. T.; Ramamoorthi, R.; and Ng, R. 2021. NeRF: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1): 99–106.

Parigi, M.; Martina, S.; and Caruso, F. ???? Quantum-Noise-Driven Generative Diffusion Models. *Advanced Quantum Technologies*, 2300401.

Parisi, L.; Neagu, D.; Ma, R.; and Campean, F. 2022. Quantum ReLU activation for Convolutional Neural Networks to improve diagnosis of Parkinson's disease and COVID-19. *Expert Systems with Applications*, 187: 1–17.

Pérez-Salinas, A.; Cervera-Lierta, A.; Gil-Fuster, E.; and Latorre, J. I. 2020. Data re-uploading for a universal quantum classifier. *Quantum*, 4: 226.

Romero, J.; Olson, J. P.; and Aspuru-Guzik, A. 2017. Quantum autoencoders for efficient compression of quantum data. *Quantum Science and Technology*, 2(4): 045001.

Schuld, M.; Bergholm, V.; Gogolin, C.; Izaac, J.; and Killoran, N. 2019. Evaluating analytic gradients on quantum hardware. *Physical Review A*, 99(3): 032331.

Schuld, M.; Sinayskiy, I.; and Petruccione, F. 2015. An introduction to quantum machine learning. *Contemporary Physics*, 56(2): 172–185.

Sim, S.; Johnson, P. D.; and Aspuru-Guzik, A. 2019. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, 2(12): 1–18.

Sitzmann, V.; Martel, J. N. P.; Bergman, A. W.; Lindell, D. B.; and Wetzstein, G. 2020. Implicit Neural Representations with Periodic Activation Functions. In *NeurIPS*, 1–12.

Yang, Y.; and Sun, M. 2022. A Quantum-Powered Photorealistic Rendering. *arXiv preprint arXiv:2211.03418*.

Zhang, B.; Xu, P.; Chen, X.; and Zhuang, Q. 2024a. Generative quantum machine learning via denoising diffusion probabilistic models. *Physical Review Letters*, 132(10): 100602.

Zhang, X.; Yang, R.; He, D.; Ge, X.; Xu, T.; Wang, Y.; Qin, H.; and Zhang, J. 2024b. Boosting Neural Representations for Videos with a Conditional Decoder. In *CVPR*.

Zhao, J.; Qiao, W.; Zhang, P.; and Gao, H. 2024. Quantum Implicit Neural Representations. In *Proceedings of the 41st International Conference on Machine Learning*, 1–17.

Zhao, L.; Ma, K.-K.; Liu, Z.; Yin, Q.; and Chen, J. 2022. Real-Time Scene-Aware LiDAR Point Cloud Compression Using Semantic Prior Representation. *IEEE Trans. Circuits Syst. Video Technol.*, 32(8): 5623–5637.

Zheng, J.; Gao, Q.; and Lü, Y. 2021. Quantum graph convolutional neural networks. In *2021 40th Chinese Control Conference (CCC)*, 6335–6340. IEEE.