HQSP-GAN: A hybrid quantum–classical architecture for high-quality image generation

Minghua Pan(潘明华)[1,2,3], Binggui Huang(黄炳贵)[1], Shaohua Sun(孙绍华)[2,4], Yichao Xu(徐益超)[1], Zhitao Chen(陈智涛)[1], and Shenggen Zheng(郑盛根)[5,†]

[1] *School of Computer and Information Security, Guilin University of Electronic Technology, Guilin* 541000, *China*
[2] *Guangxi Engineering Research Center of Industrial Internet Security and Blockchain, Guilin University of Electronic Technology, Guilin* 541004, *China*
[3] *Guangxi Key Laboratory of Cryptography and Information Security, Guilin University of Electronic Technology, Guilin* 541000, *China*
[4] *School of Artificial Intelligence, Guilin University of Electronic Technology, Guilin* 541000, *China*
[5] *Guangdong–Hong Kong–Macau Greater Bay Area (Guangdong) Quantum Science Center, Shenzhen* 518045, *China*

(Received 16 September 2025; revised manuscript received 12 November 2025; accepted manuscript online 17 November 2025)

Quantum machine learning has attracted considerable attention due to its potential in classification and recognition tasks. As a significant branch of this field, quantum generative adversarial networks (QGANs) have achieved notable progress in image generation. Although existing models have improved in terms of generation quality and efficiency, their performance remains suboptimal. To overcome these limitations, we propose HQSP-GAN, a hybrid quantum–classical generative adversarial model integrating a shared parameter mechanism. This model innovatively designs a quantum parameter sharing scheme to reduce the number of trainable parameters in the quantum module. Additionally, by incorporating attention mechanisms, the Wasserstein distance, and gradient penalty techniques, HQSP-GAN enhances training stability and generation performance. Extensive experiments on standard datasets such as MNIST and Fashion MNIST demonstrate that HQSP-GAN can generate high-quality grayscale images without additional dimensionality reduction or preprocessing. Moreover, it outperforms SOTA models including QINR-GAN and PQWGAN in metrics such as SSIM and PSNR, highlighting the efficiency and practicality of its architectural design.

Keywords: quantum generative adversarial network, quantum–classical hybrid architecture, parameter sharing mechanism, attention mechanism, image generation

PACS: 03.67.Ac, 03.67.Lx, 84.35.+i, 42.50.–p, 07.05.Mh

DOI: 10.1088/1674-1056/ae1fe6 **CSTR:** 32038.14.CPB.ae1fe6

1. Introduction

In recent decades, generative models have become a crucial research branch in the field of machine learning (ML), widely applied in tasks such as image synthesis,[1] text modeling,[2] and speech generation.[3] Their core objective is to learn the underlying probability distribution of data and generate samples statistically consistent with real data. Among various generative models, generative adversarial networks (GANs) have attracted extensive attention due to their intuitive architecture and realistic generation results.[4,5] A GAN consists of a generator and a discriminator that engage in a minimax game: the generator aims to learn the latent data distribution to produce indistinguishable fake samples, while the discriminator attempts to differentiate real samples from generated ones. Through this adversarial process, GANs continuously improve the realism and quality of generated data.

Despite the theoretical strength of GANs in generation tasks, their training process often suffers from instability issues, including vanishing gradients and mode collapse.[6] To tackle these problems, various improvements have been proposed. In 2017, Arjovsky *et al*. proposed the Wasserstein GAN (WGAN),[7] which replaces the original Jensen–Shannon divergence with the Earth Mover's Distance (also known as the Wasserstein distance). This effectively alleviates the vanishing gradient problem and leads to a more stable training process. Subsequently, Gulrajani *et al*. proposed WGAN-GP,[8] which incorporates a gradient penalty term instead of weight clipping, better satisfying the Lipschitz continuity constraint and enhancing the model's generalization and stability. However, performance limitations remain when handling high-dimensional and complex image data.

With the rapid development of quantum computing technologies, quantum machine learning (QML) has emerged as a promising research frontier.[9] Leveraging unique quantum physical properties such as superposition, entanglement, and interference, quantum models demonstrate inherent advantages in modeling complex probability distributions.[10] Within this context, quantum generative adversarial networks (QGANs) integrate the core ideas of GANs with the high-dimensional linear transformation capabilities of quantum circuits.[11] By replacing either the generator, the discrimina-

†Corresponding author. E-mail: zhengshenggen@quantumsc.cn

© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

CPB Online IP Press

tor, or both with quantum modules, QGANs aim to maintain expressive power while significantly reducing the number of parameters and improving computational efficiency.[12]

Existing studies have demonstrated the advantages of QGANs in various generation tasks. For example, QuGAN introduces quantum state fidelity as a loss function, significantly improving image quality on MNIST generation tasks while drastically reducing model parameters.[13] Moreover, QuGAN has been successfully implemented on superconducting quantum circuits, achieving an average reconstruction accuracy of quantum data distributions as high as 98.8%. To balance expressiveness and training efficiency, Tsang *et al*. proposed the parameterized quantum Wasserstein GAN (PQWGAN) model,[14] which constructs a hybrid quantum–classical architecture combining Wasserstein distance and gradient penalty. Without relying on any dimensionality reduction or classical preprocessing, this method successfully generated 28×28 grayscale images on the MNIST and Fashion MNIST datasets, reducing trainable parameters by three orders of magnitude while maintaining classical generation quality. Furthermore, Wu *et al*. proposed the QINR-GAN (quantum-inspired neural representation GAN) model,[15] which combines implicit neural representation (INR) with quantum circuits to encode spatial image information quantum mechanically, achieving high-fidelity facial image synthesis. Experiments on the CelebA dataset showed that QINR-GAN outperforms several classical baselines in structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR), achieving comparable results to state-of-the-art (SOTA) models with fewer quantum parameters.

Despite the great theoretical potential and preliminary experimental success, QGANs face multiple challenges in practical applications. On one hand, current quantum hardware suffers from limited qubit counts and poor circuit stability, making it difficult to support the large-scale parameter training required for high-dimensional image generation. This limitation restricts the expressive capability of models under constrained quantum resources. On the other hand, pure quantum models are highly susceptible to environmental noise and lack robust optimization mechanisms, leading to unstable training. Therefore, constructing a hybrid architecture that combines the stable training ability of classical neural networks with the strong expressive power of quantum models is a key breakthrough to improve the practical performance of QGANs.

Motivated by these challenges, this paper proposes HQSP-GAN (hybrid quantum shared-parameter GAN), a hybrid quantum–classical generative adversarial network based on a shared parameter mechanism. The model introduces hybrid quantum shared blocks within the quantum generator to enhance expressiveness under limited qubit conditions, thereby reducing the number of training parameters while maintaining effective training performance. At the same time, it uses the powerful image modeling capability of classical convolutional neural networks to efficiently learn complex image data distributions and generate high-fidelity samples. The main contributions of this work are as follows:

(i) We propose a hybrid generative adversarial network architecture integrating quantum shared-parameter blocks with classical neural networks, capable of generating high-quality images with reduced parameters.

(ii) An attention mechanism is incorporated to enhance the model's capability in capturing critical information, thereby improving the quality and details of the generated images.

(iii) Multiple evaluation metrics are employed to demonstrate algorithm performance, including classical image generation metrics and quantum resource consumption, with detailed theoretical and experimental comparisons against SOTA algorithms.

The structure of this paper is organized as follows: Section 2 introduces the fundamentals of qubits, quantum gates, and parameterized quantum circuits. Section 3 details the HQSP-GAN model architecture, focusing on the design of quantum shared-parameter blocks, the construction of the hybrid quantum generator, and the implementation of the discriminator module along with related training strategies. Section 4 presents experimental results and performance analysis, demonstrating the model's performance and comparative studies across multiple datasets. Finally, Section 5 concludes our work, discusses limitations of the current model, and outlines future research directions.

2. Preliminaries

As the fundamental unit of quantum computation, qubits differ essentially from classical bits. Although a classical bit can exist only in one of two distinct states, 0 or 1, a qubit can exist in a superposition of both states simultaneously. The general form of a qubit can be represented as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{1}$$

where $|0\rangle$ and $|1\rangle$ are orthogonal computational basis states, with complex coefficients α and β denoting their probability amplitudes. The squared magnitudes $|\alpha|^2$ and $|\beta|^2$ give the probabilities of measuring the qubit in $|0\rangle$ and $|1\rangle$, respectively. Multiple qubits together form quantum states in a high-dimensional Hilbert space, enabling the representation of more complex information structures.[16]

Quantum gates are reversible operations on qubits, represented by unitary matrices. Common single-qubit gates include the Pauli gates (X, Y, Z) and the corresponding rotation gates around different axes. The matrix forms of the rotation

CPB Online In-Press

gates $R_x(\theta)$, $R_y(\theta)$, and $R_z(\theta)$ are as follows:

$$R_x(\theta) = \begin{pmatrix} \cos\dfrac{\theta}{2} & -\mathrm{i}\sin\dfrac{\theta}{2} \\ -\mathrm{i}\sin\dfrac{\theta}{2} & \cos\dfrac{\theta}{2} \end{pmatrix},$$

$$R_y(\theta) = \begin{pmatrix} \cos\dfrac{\theta}{2} & -\sin\dfrac{\theta}{2} \\ \sin\dfrac{\theta}{2} & \cos\dfrac{\theta}{2} \end{pmatrix},$$

$$R_z(\theta) = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\theta/2} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\theta/2} \end{pmatrix}. \tag{2}$$

Two-qubit gates can establish entanglement between two qubits. Among them, the controlled-NOT (CNOT) gate is one of the most commonly used two-qubit gates. It operates on two input qubits: a control qubit and a target qubit. The matrix representation of the CNOT gate is given by

$$\mathrm{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{3}$$

Parameterized quantum circuits (PQCs)[17] are trainable quantum circuit architectures that incorporate classical optimization strategies. The central idea is to incorporate tunable parameters into the design of quantum gates, allowing the circuit to be trained via optimization algorithms such as gradient descent,[18,19] thereby enabling tasks such as classification, generation, and regression. A typical PQC starts from the ground state $|0\rangle^{\otimes n}$, followed by a sequence of rotation gates that encode classical data into the quantum state space, an approach known as data encoding[20] or data re-uploading.[21] After encoding, the qubits are passed through several parameterized rotation gates (*e.g.*, $R_x(\theta)$, $R_y(\theta)$, $R_z(\theta)$) and entangling gates such as CNOT. The rotation angles act as trainable parameters and are optimized during training.

3. Hybrid quantum–classical generative adversarial network with shared parameters

This section introduces the proposed hybrid quantum–classical generative adversarial network with shared parameters, referred to as HQSP-GAN, and provides a detailed explanation of its architecture, components, and optimization strategies. As shown in Fig. 1, the HQSP-GAN model integrates the advantages of quantum computing with the powerful capabilities of classical neural networks to generate high-quality image data. The architecture of HQSP-GAN mainly consists of two components: a quantum–classical hybrid generator and a classical discriminator. The generator adopts a hybrid design that incorporates both quantum and classical modules. The quantum module is built using a quantum neural network that includes a data re-uploading encoding layer and a parameterized processing layer, and it is augmented by multiple classical neural network layers to enhance its expressiveness. The discriminator is entirely composed of classical neural networks and is responsible for determining the authenticity of the input images.

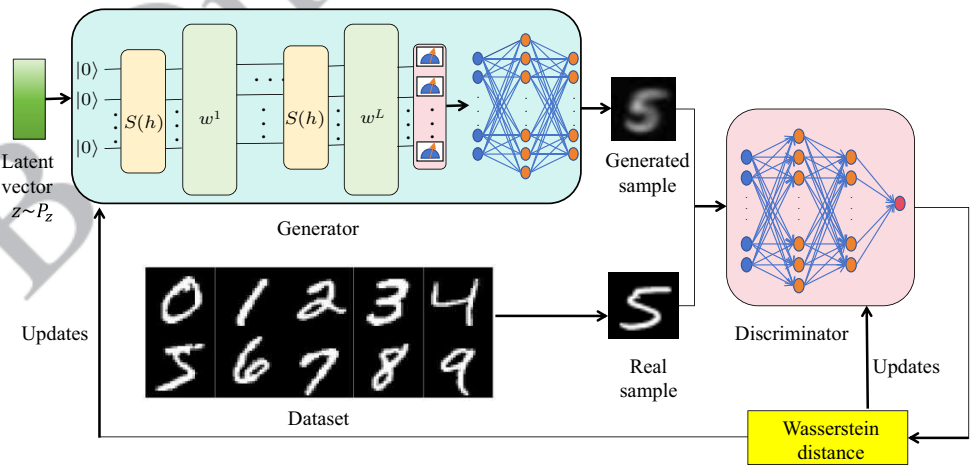


Fig. 1. Overall architecture of the HQSP-GAN model.

3.1. Design of shared-parameter mechanism

In the HQSP-GAN model, the core objective of the shared parameter mechanism is to enhance training efficiency and generation performance while effectively reducing the total number of parameters and network complexity. This mechanism is applied to the quantum parameter layers of the generator, where multiple quantum neural network modules share a common set of key quantum gate parameters to achieve feature transfer and information compression. Specifically, the quantum parameter layer in the generator consists of multiple repeated shared parameter blocks U_{SP}, whose detailed structure is illustrated in Fig. 2. This module employs a two-qubit quantum circuit comprising three rotation gates and two CNOT gates, forming a basic quantum neural network unit. Param-

CPB Online Ir-Press

eterized rotation gates $R_Y\left(\alpha^{l_1}\right)$ and $R_Y\left(\alpha^{l_2}\right)$ are applied respectively to the upper and lower qubits to initialize the directions of the quantum states, where α^{l_1} and α^{l_2} denote the first and second parameters of the l-th layer. Subsequently, a CNOT gate is used to introduce entanglement between the qubits. A rotation gate is then applied on the target qubit for phase adjustment, with $R_Y\left(\alpha^{l_3}\right)$ representing the third parameter of the α^{l_3}-th layer. Finally, another CNOT gate is applied to enhance nonlinear interference between quantum states. Each quantum parameter layer is composed of shared parameter blocks (U_{SP}) that share parameters, meaning all submodules within the same layer use the same set of rotation gate angle parameters α^l. Through the shared parameter mechanism, the generalizability of the network and the utilization efficiency of parameters are significantly improved.

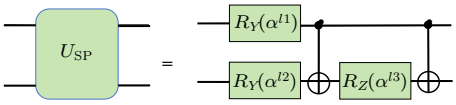


Fig. 2. Shared parameter block.

The shared parameter block can be represented as

$$U_{\mathrm{SP}}(j,k) = \mathrm{CNOT}_{j,k} \cdot \left(I \otimes R_{Z_k}(\alpha^{l_3})\right) \cdot \mathrm{CNOT}_{j,k} \cdot \left(R_{Y_j}(\alpha^{l_1}) \otimes R_{Y_k}(\alpha^{l_2})\right). \tag{4}$$

Here, (j,k) denotes a pair of qubits, where j serves as the control qubit and k as the target qubit.

3.2. Structure of the generator

The generator of HQSP-GAN adopts a quantum–classical hybrid architecture, as illustrated in Fig. 3. This architecture has been validated in Ref. [22] for its practical effectiveness and performance advantages on specific tasks. Furthermore, the authors in Ref. [15] provided additional theoretical and empirical support from multiple perspectives — such as Fourier approximation theory and high performance with few parameters — demonstrating the critical role of quantum modules in generative tasks. The quantum module consists of alternating encoding layers $S(h)$ and parameter layers $W(l)$. The input prior noise vector $z \in R^d \sim \mathcal{N}(0,I)$ is embedded in the quantum module through a data reuploading mechanism, where input features are injected across multiple layers of the quantum circuit. This approach not only enhances the model's expressive power but also alleviates the representation bottleneck caused by limited qubits. In each layer of the quantum module, the parameter layer is composed of multiple shared parameter blocks U_{SP}, aiming to achieve information compression and feature transfer. In the l-th layer, the encoding layer and the parameter layer can be expressed as follows:

$$\begin{aligned} &U^l = W^l \cdot S(h), \quad S(h) = R_X(h_1) \otimes R_X(h_2) \otimes \cdots \otimes R_X(h_n), \\ &W^l = \bigotimes_{i=1}^{n} R_Z^{(i)}(\theta_i^{l_2}) \cdot \bigotimes_{i=1}^{n} R_Y^{(i)}(\theta_i^{l_1}) \cdot U_{\mathrm{SP}}^{(l)}, \quad U_{\mathrm{SP}}^{(l)} = \bigotimes_{i=1}^{n/2} U_{\mathrm{SP}}(2i,2i+1) \cdot \bigotimes_{i=1}^{n/2} U_{\mathrm{SP}}(2i-1,2i). \end{aligned} \tag{5}$$

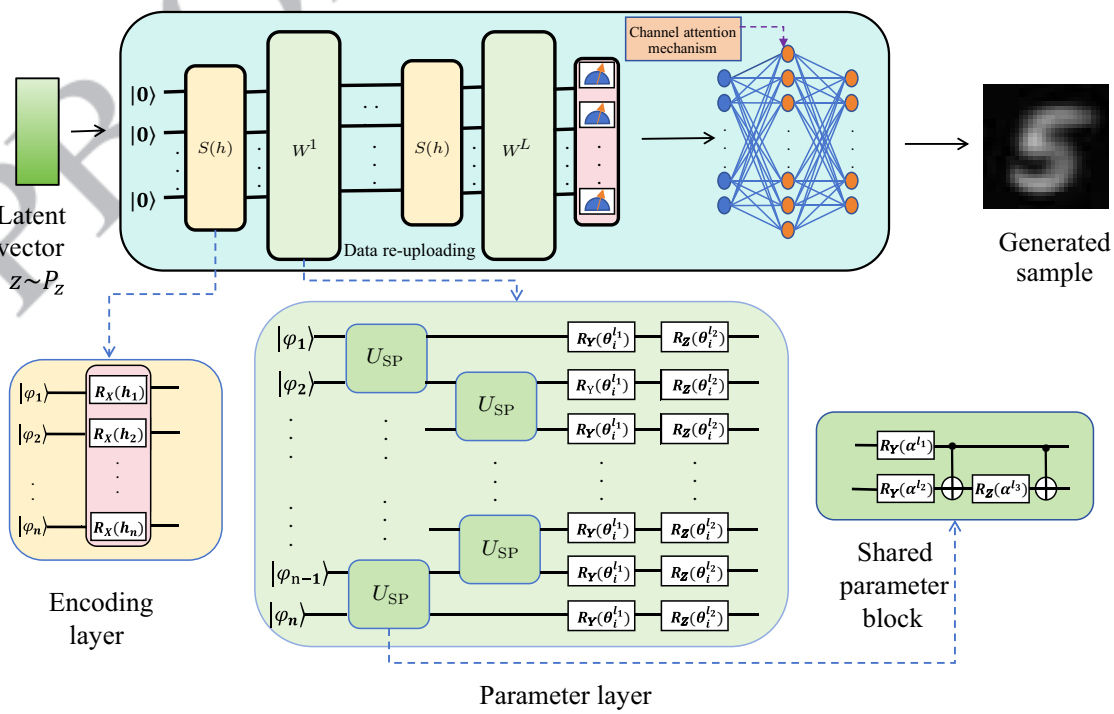


Fig. 3. Architecture of the generator.

CPC Online In-Press

Here, $R_X(h_i)$ represents a fixed input encoding based on the parameter h_i, while $\bigotimes_{i=1}^{n} R_Z^{(i)}(\theta_i^{l_2})$ and $\bigotimes_{i=1}^{n} R_Y^{(i)}(\theta_i^{l_1})$ correspond to the second and first rotation layers, respectively. After quantum feature extraction is completed, the quantum state is measured and mapped to a classical vector representation, which is then fed into the subsequent classical neural network module. The entire process can be expressed as follows:

$$|\psi_\Theta(z)\rangle = \left(\prod_{l=1}^{L} U^l\right)|0\rangle^{\otimes n}. \tag{6}$$

Here,

$$\Theta = \left\{\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(L)}; \theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(L)}\right\}$$

denotes the set of trainable parameters across all layers, where the parameters of each layer are trained independently. The state $|0\rangle^{\otimes n}$ represents the initial state of an n-qubit quantum system.

To further enhance the network's ability to selectively represent inter-channel features in image data, a channel attention mechanism is introduced into the classical network structure. This mechanism adaptively adjusts the weights of different feature channels, enabling the network to focus more on semantically significant or structurally critical channels, thereby improving the clarity and discriminability of the generated images. Unlike traditional fully quantum generators, the design of HQSP-GAN not only improves the expressive capacity of quantum circuits for efficient encoding and information processing through a parameter-sharing mechanism, but also enhances the controllability and fine-grained modeling capability of image generation via the incorporation of channel attention in the classical neural network.

3.3. Structure of the discriminator

Given the limitations of current quantum hardware in terms of qubit count and circuit depth, the expressive capacity of the generator remains constrained. In this context, employing an overly complex discriminator architecture may result in the discriminator significantly outperforming the generator during training, thereby disrupting the balance of adversarial learning. To address this issue, HQSP-GAN adopts the discriminator architecture proposed in the WGAN-GP framework[23,24] as its baseline design, as illustrated in Fig. 4. This discriminator consists of two fully connected hidden layers, each containing a certain number of neurons, followed by a single-neuron output layer for final discrimination. The architecture is simple yet provides strong training stability, effectively supporting the quantum generator in adversarial optimization.

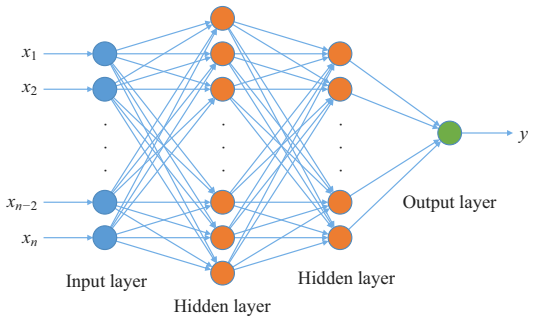


Fig. 4. Architecture of the discriminator.

3.4. Optimization strategies

In the HQSP-GAN model, to more effectively enhance the generator's image synthesis quality and the discriminator's distinguishing capability, an adversarial training strategy based on the Wasserstein distance (Wasserstein GAN, WGAN) is adopted as the core optimization framework. This is further combined with a parameter sharing mechanism and a stepwise update principle to ensure training stability and continuous improvement of generation performance. The training process consists of the following key components.

3.4.1. Loss function

To avoid the training instability and gradient vanishing problems commonly encountered in traditional GANs, the Wasserstein distance (Earth-Mover distance) is adopted as the metric to measure the difference between the generated image distribution and the real image distribution. This is complemented by a parameter sharing mechanism and a stepwise update strategy to ensure training stability and continuous improvement of generation performance. The Wasserstein distance possesses favorable smoothness properties, which facilitate the generator in learning the support of the true data distribution. Specifically, the Wasserstein loss is defined as

$$\mathcal{L}_{\mathrm{w}} = E_{x\sim P_{\mathrm{real}}}[D(x)] - E_{z\sim P_z}[D(G(z))]. \tag{7}$$

Here, $D(x)$ denotes the discriminator's output for real samples, and $D(G(z))$ represents the discriminator's score for generated samples. The original WGAN relies on weight clipping to enforce the 1-Lipschitz constraint on the discriminator; however, this limits the network capacity and makes training difficult. Therefore, the WGAN-GP training scheme is adopted, which introduces a gradient penalty term. Its form is as follows:

$$\mathcal{L}_{\mathrm{GP}} = \lambda \cdot E_{\hat{x}\sim P_{\hat{x}}}\left[\left(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1\right)^2\right], \tag{8}$$

where $\hat{x}$ is obtained by interpolating between real and generated samples, and λ is the penalty coefficient, commonly set to 10. This term encourages the discriminator's gradient norm with respect to its input to be close to 1, thereby promoting smoother behavior and enforcing the 1-Lipschitz condition. Consequently, the adversarial training discriminator loss function is defined as follows:

$$\mathcal{L} = E_{x\sim P_{\mathrm{real}}}[D(x)] - E_{z\sim P_z}[D(G(z))]$$

CPB Online In-Press

$$+\lambda \cdot E_{\hat{x}\sim P_{\hat{x}}}\left[\left(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1\right)^2\right]. \tag{9}$$

3.4.2. Synchronization of training under parameter sharing mechanism

During training, the shared parameter module U_{SP} is invoked by multiple generation paths simultaneously. To avoid gradient conflicts and redundant updates, a synchronized update strategy is adopted: within a single training batch, the quantum parameters shared by all submodules in each layer are updated collectively only after each complete generator update, ensuring consistency in the gradient direction.

3.4.3. K-step discriminator update strategy

Following the WGAN training framework, the discriminator is updated K times (typically $K = 5$) prior to each generator parameter update to strengthen the discriminator's distinguishing capability. This design effectively suppresses the "deceptive" gradients encountered by the generator during the early training stages, thereby improving the quality of the generated samples.

3.5. Algorithm implementation

Based on the design principles and training strategies of HQSP-GAN introduced above, the overall training procedure is summarized in Algorithm 1. The algorithm takes random noise as input, initializes the generator and discriminator parameters, and alternately updates them using the specified optimization strategy. For detailed steps, including the specific update rules and gradient computations, see Appendix A.

4. Experiments and analysis

To comprehensively evaluate the performance advantages of HQSP-GAN, comparative experiments were conducted against several representative generative adversarial network models, including recently proposed hybrid quantum–classical architectures such as PQWGAN[14] and QINR-GAN.[15] PQWGAN, which integrates parameterized quantum circuits with the Wasserstein distance, has been widely applied in quantum generation tasks. Meanwhile, QINR-GAN enhances the generator's modeling capability by simulating quantum superposition states and interference properties. These two models represent the current development directions of quantum–classical generative adversarial networks. We performed comparisons with the same datasets and training conditions, allowing for a more accurate assessment of the improvement of HQSP-GAN in generation quality, model stability, and convergence speed.

4.1. Datasets and experimental environment

This study uses the publicly available MNIST[25] and Fashion MNIST[26] datasets as benchmark data sets to evaluate model performance. To ensure that the training process can be completed within a reasonable time frame, the first 500 samples from each class in the training set are selected for model training. Although generating MNIST and Fashion MNIST data is generally considered a relatively simple task in GAN research, prior studies have pointed out that generating high-quality 28×28 resolution images remains a significant challenge for QGANs in the NISQ era. Therefore, validating the proposed model's image generation capabilities under this setting holds practical significance.[27,28] We evaluated HQSP-GAN via simulations using PyTorch 1.12.0, PennyLane 0.20.1, and Python 3.9.13.

4.2. Evaluation metrics

To comprehensively evaluate the performance and effectiveness of HQSP-GAN in image generation tasks, four representative image quality evaluation metrics are adopted: Fréchet inception distance (FID),[29] structural similarity index measure (SSIM),[30] cosine similarity (CS),[31] and peak signal-to-noise ratio (PSNR).[32] These metrics quantify the similarity between generated and real images from multiple perspectives. For detailed definitions and formulas of these metrics, please refer to Appendix B.

4.3. HQSP-GAN model parameter settings

During the training of HQSP-GAN, a set of optimized hyperparameters was configured to ensure a good balance between model stability and generation quality. Specifically, the learning rate of the generator was set to 0.00001, and its quantum module was constructed using 6 qubits. The generator's feedforward neural network consists of three linear layers with 256, 512, and 256 neurons, respectively. The discriminator's learning rate was set to 0.00012. Besides these customized parameters, other training strategies follow the default settings of WGAN-GP, including a gradient penalty coefficient $\lambda = 10$ and $n_c = 5$ discriminator updates before each generator update. The Adam optimizer was employed with hyperparameters $\beta_1 = 0$ and $\beta_2 = 0.9$. For the prior noise design, the input was restricted to the interval $[0, 1)$ rather than the more common $[-\pi, \pi]$ range, better adapting to the input characteristics of the hybrid quantum–classical architecture. The model was trained for 2000 epochs with a batch size of 1, aiming to capture fine-grained changes during the training process.

4.4. Different numbers of generator layers

To evaluate the impact of different numbers of hybrid quantum–classical layers on the performance of the HQSP-GAN model, we compared its generation quality under configurations with 2, 4, 6, and 8 hybrid layers. The model was evaluated using four metrics: FID, CS, SSIM, and PSNR, considering both qualitative and quantitative aspects. For the computation of these metrics, each converged model was used to

CPB Online In-Press

generate 500 images, which were then compared against the original dataset, and the results were averaged.

Table 1. Performance comparison under different generator layer configurations.

Layer depth	FID	CS	SSIM	PSNR	Data	Training time (s)
2-layers	181.16	0.9297	**0.3118**	12.33	5	**2753.57**
4-layers	180.23	0.9297	0.3059	12.33		4052.07
6-layers	**180.17**	0.9297	0.3081	12.33		5033.08
8-layers	181.40	0.9297	0.3061	12.32		5970.22

Table 1 presents the performance of the HQSP-GAN model on the MNIST dataset for the image generation task of digit "5" under different generator layer configurations. As shown in the table, increasing the number of generator layers does not lead to significant improvements in image quality metrics (FID, CS, SSIM, and PSNR), as the variations in values are minimal. However, the training time increases substantially with deeper architectures. This suggests that simply increasing the number of layers in the generator does not effectively enhance image quality but instead results in higher training costs. Therefore, the depth of the model should be reasonably controlled to balance generation performance and computational efficiency in this task.

4.5. Experiments on the MNIST dataset

To evaluate the image generation capability of HQSP-GAN, experiments were conducted on four digit classes (1, 3, 5, 7) from the MNIST dataset. Its performance was compared with PQWGAN and QINR-GAN. As shown in Fig. 5, real MNIST samples and generated images from the three quantum generative models are presented. Subjectively, HQSP-GAN clearly outperforms the others, producing sharper images and better preserving digit structural features. These results validate the effectiveness of the shared parameter block design, attention mechanisms, and hybrid quantum–classical strategy in HQSP-GAN for image modeling.

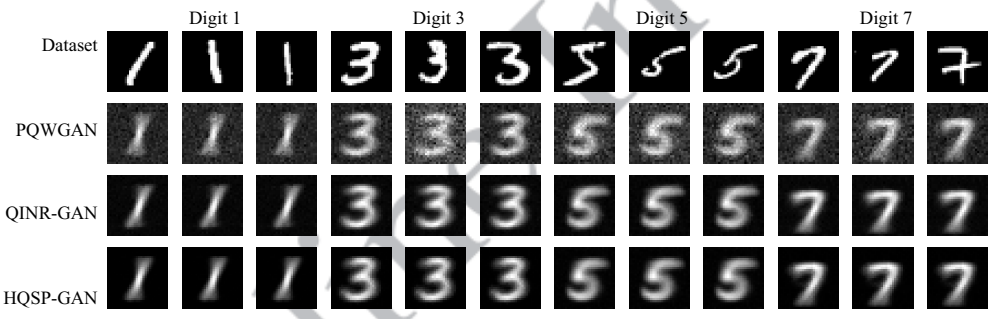


Fig. 5. Comparison of digit image generation by different quantum generative models.

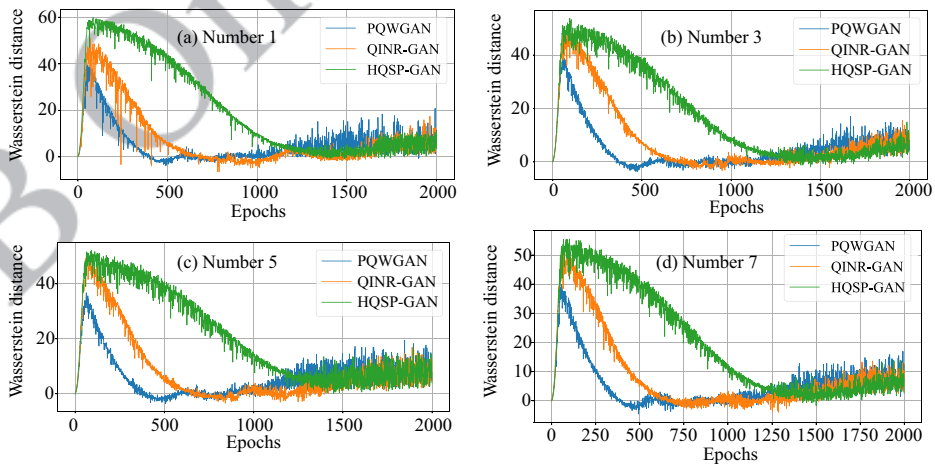


Fig. 6. Comparisons of Wasserstein distance generation using different quantum generative models.

In addition, Fig. 6 illustrates the evolution of the Wasserstein distance during training for each model. It can be observed that the Wasserstein distance steadily decreases as training progresses and eventually stabilizes, indicating that the generator gradually learns the target data distribution and that the models demonstrate strong effectiveness in the image generation task. The continuous decline in the Wasserstein distance suggests that the distribution gap between generated and real samples is progressively narrowing. It is worth noting that although PQWGAN and QINR-GAN achieve convergence of the Wasserstein distance to near zero within fewer training epochs, this is primarily due to the larger parameter scales of their models, which typically come at the cost of higher per-epoch computational overhead.

CPB Online-First

Figure 7 presents a comparison of the total training time required by each model to complete 2000 epochs on the digit "5". For PQWGAN, the total training time was 9425.32 seconds, with the Wasserstein distance converging around the 500th epoch, corresponding to approximately 2350 seconds. In contrast, HQSP-GAN completed 2000 training epochs in just 2967.67 seconds, with its Wasserstein distance converging around the 1250th epoch, taking approximately 1875 seconds. Therefore, when evaluating training efficiency, it is important not only to consider the number of epochs required for convergence but also to take into account the time consumed per epoch to accurately assess the model's actual computational cost and efficiency. Overall, HQSP-GAN demonstrates significantly higher training efficiency compared to the other models.

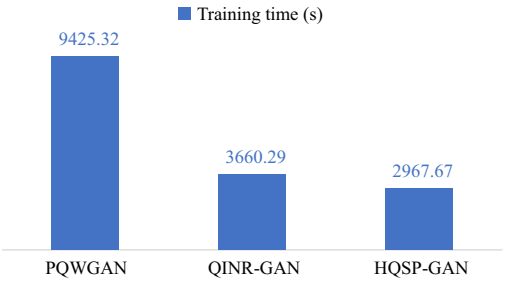


Fig. 7. Comparison of total training for generating digit 5 by different quantum generative models.

A comparison of image quality generated by the HQSP-GAN model, PQWGAN, and QINR-GAN is provided in Appendix C. Each model is evaluated using four metrics: FID, SSIM, CS, and PSNR. For the FID metric, lower values indicate better performance; thus, the HQSP-GAN model performs comparably to QINR-GAN, and both outperform PQWGAN. The second metric, CS, favors higher values, with values close to 1 indicating higher similarity. The third metric, SSIM, shows that the HQSP-GAN model achieves higher scores than both QINR-GAN and PQWGAN. Compared to other metrics, SSIM better highlights the superior performance of QINR-GAN. The fourth metric is PSNR, where higher values represent better image quality. Overall, across all evaluation metrics, the QINR-GAN model demonstrates better performance than the other two models.

4.6. Experiments on the fashion MNIST dataset

To evaluate the transferability and generalization ability of the HQSP-GAN model under different data distributions, this section selects the Fashion MNIST dataset, which has higher structural complexity, to assess its generation performance. HQSP-GAN achieves clearer contours and higher image quality compared with PQWGAN and performs on par with QINR-GAN. For a detailed discussion and extended analysis of these results, please refer to Appendix D.

4.7. Ablation study

To evaluate the actual contribution of each module in the HQSP-GAN model, an ablation study was conducted. In this experiment, HQSP-GAN (w/o CA) denotes the model with the channel attention mechanism removed, while HQSP-GAN (w/o SP & CA) denotes the model with both the channel attention mechanism and the shared parameter mechanism removed. The experiments were carried out on the "Number 5" class of the MNIST dataset and the "sneaker" class of the Fashion-MNIST dataset. The results are summarized in Table 2.

Table 2. Performance evaluation of the HQSP-GAN ablation experiments.

Model	FID	CS	SSIM	PSNR	Data	Training time (s)
HQSP-GAN	181.16	0.9297	0.3118	12.33	Number 5	2967.67
HQSP-GAN (w/o CA)	184.30	0.9294	0.2915	12.29		2921.73
HQSP-GAN (w/o SP & CA)	185.21	0.9291	0.2893	12.28		3273.96
HQSP-GAN	180.07	0.9573	0.4632	15.48	Sneaker	3039.92
HQSP-GAN (w/o CA)	183.52	0.9565	0.4600	15.38		3194.87
HQSP-GAN (w/o SP & CA)	184.89	0.9567	0.4610	15.44		3377.49

From the comparisons, it can be observed that removing the shared parameter mechanism (HQSP-GAN w/o SP & CA) leads to degraded generation performance and increased training time. This indicates that the shared parameter mechanism not only enhances performance but also effectively reduces computational overhead during training. Furthermore, comparing HQSP-GAN with HQSP-GAN (w/o CA) shows that removing the attention mechanism results in a significant increase in the FID score and a noticeable decrease in SSIM, while training time remains largely unchanged. This demonstrates that the channel attention mechanism can further improve image generation quality without adding a significant training burden.

In conclusion, the experimental results validate the effectiveness and synergy of the shared parameter mechanism and the channel attention mechanism in improving both the performance and efficiency of the model.

CPB Online In Press

4.8. Quantum resource analysis

The quantum resources consumed by a model have gradually become one of the key metrics for evaluating its quality. Unlike external factors such as experimental apparatus, the evaluation of quantum resources is more intrinsic and helps to fundamentally analyze the complexity and practicality of a model. To this end, we compare the quantum resource requirements of HQSP-GAN with those of QINR-GAN and PQWGAN under the same image generation task, as summarized in Table 3.

Table 3. Comparison of quantum resources usage among different models. SQG: single-qubit gates, TP: trainable parameters, TQG: two-qubit gates.

Model	Qubits	Generators	Layers	SQG	TP	TQG
PQWGAN[14]	6	28	4	1148	2016	560
QINR-GAN[15]	6	1	2	96	72	28
HQSP-GAN(Ours)	6	1	2	66	54	20

As shown in Table 3, all three models employ 6 qubits in their generators. However, the PQWGAN model demonstrates significantly higher resource consumption: it requires 28 generators, each with a 4-layer structure, and involves more single-qubit gates and trainable parameters. In contrast, QINR-GAN and HQSP-GAN consume noticeably fewer resources. A further comparison reveals that, given the same number of generator qubits, generators, and layers, HQSP-GAN requires significantly fewer single-qubit gates, trainable parameters, and two-qubit gates than QINR-GAN, exhibiting higher resource efficiency.

These results indicate that HQSP-GAN not only achieves superior performance but also significantly reduces dependence on quantum resources, demonstrating strong application potential in the context of the NISQ era.

5. Conclusion and future work

In this paper, we propose a hybrid adversarial generative network architecture, HQSP-GAN, that integrates a quantum shared-parameter structure with attention mechanisms. The model is designed to ensure high-quality image generation while significantly reducing the demand for quantum resources. By introducing the quantum shared parameter mechanism, HQSP-GAN effectively improves resource utilization while maintaining both generative capacity and training stability.

Experimental results demonstrate that HQSP-GAN outperforms state-of-the-art quantum generative models, such as PQWGAN and QINR-GAN, across various image generation tasks. It achieves superior performance in image quality metrics (*e.g.* FID, SSIM, PSNR) and exhibits clear advantages in training efficiency. Moreover, a detailed comparison of quantum resource requirements shows that HQSP-GAN consistently requires fewer qubits, quantum gates, and trainable parameters than the baseline models, highlighting its strong practical potential in the NISQ era. Furthermore, this study performed generation experiments on the CelebA dataset. The current work focused on low-resolution 28×28 pixel images, with results provided in Appendix E. Future work will extend the QINR-GAN model for higher-dimensional data.

Appendix A: Pseudocode for HQSP-GAN

Algorithm 1 HQSP-GAN

Input: Require Adam hyperparameters lr_D, lr_G, β_1, β_2; gradient penalty coefficient λ; discriminator update steps per generator step K; total epochs *epoch*; batch size m

Output: High-quality generated images

Initialize discriminator parameters ω and generator parameters θ
for each epoch **do**
  for $k = 1$ to K **do**
    for $i = 1$ to m **do**
      Sample real data $x \sim P_{\text{data}}$, noise $z \sim P_z$, and $\varepsilon \sim \mathcal{U}[0,1]$
      Generate fake sample $\bar{x} = G_\theta(z)$
      Interpolate $\hat{x} = \varepsilon x + (1-\varepsilon)\bar{x}$
      Compute discriminator loss:
      $$Loss_D = D(\bar{x}) - D(x) + \lambda(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2$$
    end for
    Update discriminator parameters ω
  end for
  for $i = 1$ to m **do**
    Sample noise $z \sim P_z$
    Generate fake sample $\bar{x} = G_\theta(z)$
    Compute generator loss: $Loss_G = -D(\bar{x})$
  end for
  Update generator parameters θ
end for

Appendix B: Evaluation metrics

B1. Fréchet inception distance (FID)

FID is one of the most widely used evaluation metrics for image generative models. It extracts feature representations of generated and real images from the Inception network and assumes that these representations follow a multivariate Gaussian distribution. The FID between the two distributions is then computed. A lower FID value indicates that the distribution of generated images is closer to that of real images, implying higher image quality. Compared to traditional pixel-level metrics, FID better reflects the perceptual fidelity and diversity as perceived by humans. The formula for FID is given as follows:

$$\text{FID} = \|\mu_{\text{r}} - \mu_{\text{g}}\|^2 + Tr\left(\Sigma_{\text{r}} + \Sigma_{\text{g}} - 2\left(\Sigma_{\text{r}}\Sigma_{\text{g}}\right)^{1/2}\right). \quad \text{(B1)}$$

Here, μ_{r} and Σ_{r} denote the mean and covariance of the real images, while μ_{g} and Σ_{g} represent the mean and covariance of the generated images.

B2. Structural similarity index measure (SSIM)

SSIM is a metric that measures the similarity between two images in terms of structure, luminance, and contrast, making

CPB Online In Press

it particularly suitable for evaluating image fidelity. Its value ranges from 0 to 1, with values closer to 1 indicating higher structural similarity to the reference image. Compared to traditional mean squared error (MSE), SSIM better aligns with human visual perception. The SSIM is defined as follows:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}. \tag{B2}$$

Here, μ and σ represent the mean and standard deviation of the images, respectively; σ_{xy} denotes the covariance between the two images; and c_1, c_2 are small constants used for numerical stability.

B3. Cosine similarity (CS)

Cosine similarity is a method to measure the similarity in direction between two image vectors in a high-dimensional space, thereby assessing the "directional consistency" between the images. It is defined as follows:

$$CS = \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|}. \tag{B3}$$

Here, $\boldsymbol{x}$ and $\boldsymbol{y}$ represent the vectorized forms of the images. A value closer to 1 indicates higher directional consistency between the two images in the feature space, implying that the generated image is more semantically similar to the original image.

B4. Peak signal-to-noise ratio (PSNR)

PSNR is a classical image quality evaluation metric commonly used to measure the clarity of images after compression or reconstruction. It is expressed in decibels (dB), where a higher value indicates less distortion and better image quality. The calculation formula is as follows:

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}_I^2}{\mathrm{MSE}}\right). \tag{B4}$$

Here, MAX_I denotes the maximum possible pixel value of the image, and MSE stands for the mean squared error between the generated and reference images.

Appendix C: Additional experimental results on the MNIST dataset

The experimental results of average FID, SSIM, cosine similarity, and PSNR metrics on the MNIST dataset are shown in Fig. C1. The relevant analysis is described in Subsection 4.5.



	0	1	2	3	4	5	6	7	8	9
PQWGAN	367.43	253.93	337.65	326.01	298.13	354.03	327.26	289.26	348.83	304.76
QINR-GAN	251.48	139.66	206.12	205.82	179.74	187.22	203.87	178.71	210.39	189.28
HQSP-GAN	249.18	137.88	198.82	198.82	170.30	181.16	195.13	173.26	196.52	183.36

	0	1	2	3	4	5	6	7	8	9
PQWGAN	0.8941	0.9466	0.8904	0.9009	0.9055	0.8869	0.9084	0.9181	0.8925	0.9128
QINR-GAN	0.9249	0.9686	0.9223	0.9313	0.9386	0.9288	0.9377	0.9437	0.9299	0.9415
HQSP-GAN	0.9260	0.9692	0.9239	0.9332	0.9401	0.9297	0.9391	0.9451	0.9315	0.9434

	0	1	2	3	4	5	6	7	8	9
PQWGAN	0.2610	0.3574	0.2031	0.2401	0.1754	0.1806	0.2541	0.2845	0.2074	0.2555
QINR-GAN	0.3516	0.4677	0.2457	0.3060	0.2587	0.2664	0.3242	0.3285	0.2780	0.3119
HQSP-GAN	0.3760	0.5988	0.2812	0.3460	0.3235	0.3118	0.3947	0.4185	0.3335	0.3998

	0	1	2	3	4	5	6	7	8	9
PQWGAN	10.67	13.29	10.43	10.88	11.00	10.27	11.22	11.60	10.54	11.42
QINR-GAN	12.14	15.67	11.91	12.47	12.88	12.29	12.90	13.26	13.26	13.16
HQSP-GAN	12.19	15.76	11.98	12.58	13.00	12.33	12.98	13.37	12.48	13.30

Fig. C1. Average FID, SSIM, cosine similarity, and PSNR metrics.

Appendix D: Experiments on the fashion MNIST dataset

The image generation results of various models on the Fashion MNIST dataset are shown in Fig. D1. As observed from the figure, images generated by PQWGAN exhibit noticeable noise interference, unstable output quality, and blurred image details. This may be attributed to the greater structural complexity and category diversity of Fashion MNIST compared to MNIST, which challenges the model's generalization capability. In contrast, both QINR-GAN and HQSP-GAN demonstrate better generation performance on this dataset. The images they produce have clearer contours, fewer noise artifacts, and overall higher quality.

The Wasserstein distance during training, which reflects the convergence behavior of each model, is shown in Fig. D2. It can be observed that the Wasserstein distances of all models

CPB Online First Press

converge to values close to zero. In the early stages of training, the proposed HQSP-GAN converges relatively slowly. However, it exhibits rapid convergence in the later stages, and its final value is closer to the ideal value of zero compared to the other two methods.

To more comprehensively evaluate the image generation performance of the HQSP-GAN model on the Fashion MNIST dataset, four quantitative metrics are introduced for analysis, as shown in Fig. D3. In terms of SSIM, HQSP-GAN achieves significantly higher scores than both QINR-GAN and PQWGAN, indicating superior performance in structural fidelity. For the other metrics, HQSP-GAN clearly outperforms PQWGAN and performs on par with QINR-GAN, demonstrating overall high-quality image generation. These results indicate that HQSP-GAN possesses better adaptability and stability.

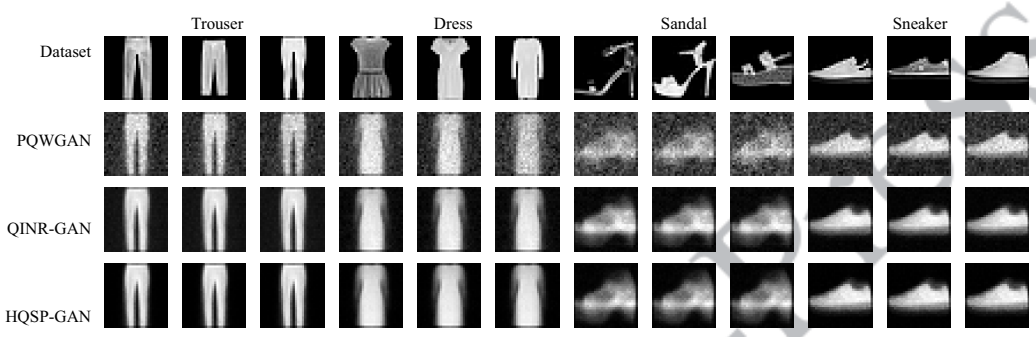


Fig. D1. Comparisons of digit image generation by different quantum generative models.

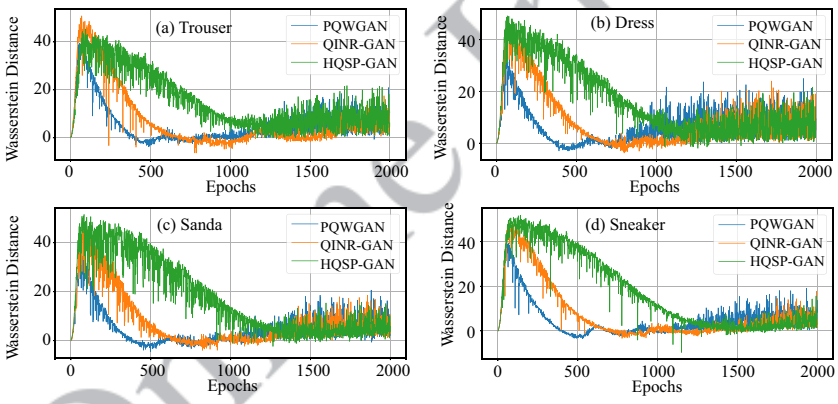
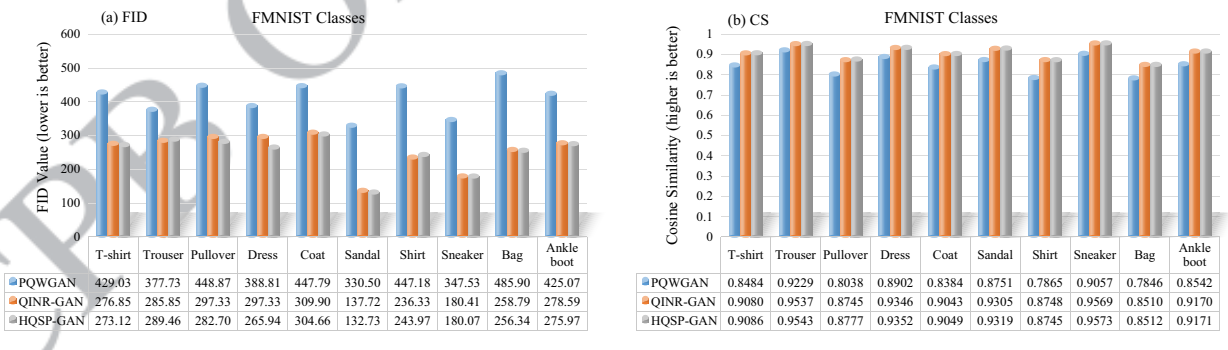


Fig. D2. Comparisons of Wasserstein distance using different quantum generative models.



	T-shirt	Trouser	Pullover	Dress	Coat	Sandal	Shirt	Sneaker	Bag	Ankle boot
PQWGAN	429.03	377.73	448.87	388.81	447.79	330.50	447.18	347.53	485.90	425.07
QINR-GAN	276.85	285.85	297.33	297.33	309.90	137.72	236.33	180.41	258.79	278.59
HQSP-GAN	273.12	289.46	282.70	265.94	304.66	132.73	243.97	180.07	256.34	275.97

	T-shirt	Trouser	Pullover	Dress	Coat	Sandal	Shirt	Sneaker	Bag	Ankle boot
PQWGAN	0.8484	0.9229	0.8038	0.8902	0.8384	0.8751	0.7865	0.9057	0.7846	0.8542
QINR-GAN	0.9080	0.9537	0.8745	0.9346	0.9043	0.9305	0.8748	0.9569	0.8510	0.9170
HQSP-GAN	0.9086	0.9543	0.8777	0.9352	0.9049	0.9319	0.8745	0.9573	0.8512	0.9171

	T-shirt	Trouser	Pullover	Dress	Coat	Sandal	Shirt	Sneaker	Bag	Ankle boot
PQWGAN	0.1655	0.3030	0.1190	0.1800	0.1475	0.0954	0.0884	0.2056	0.0419	0.1432
QINR-GAN	0.2309	0.4282	0.1804	0.2564	0.2255	0.2067	0.1425	0.3961	0.0901	0.2935
HQSP-GAN	0.2614	0.5176	0.2001	0.3135	0.2606	0.2955	0.1568	0.4632	0.1109	0.3530

	T-shirt	Trouser	Pullover	Dress	Coat	Sandal	Shirt	Sneaker	Bag	Ankle boot
PQWGAN	10.86	12.90	10.19	11.53	10.71	10.47	9.72	11.62	8.90	10.46
QINR-GAN	13.36	15.47	12.49	13.84	13.17	13.44	12.34	15.40	11.22	13.12
HQSP-GAN	13.38	15.55	12.55	13.90	13.19	13.55	12.30	15.40	11.21	13.15

Fig. D3. Average FID, SSIM, cosine similarity, and PSNR metrics.

Appendix E. Experiments on the CelebA dataset

A comparative analysis of the generation capabilities of QINR-GAN and HQSP-GAN is illustrated in Fig. E1, based on their outputs from the CelebA dataset. From the perspective of subjective image quality assessment, the HQSP-GAN model proposed in this paper significantly outperforms the QINR-GAN model in terms of generation quality. Specifically, the facial images generated by HQSP-GAN exhibit superior performance in detail clarity, texture realism, and visual comfort. It effectively suppresses noise and accurately reconstructs facial features, demonstrating its considerable capability for complex face data generation tasks. However, when compared to the original images, the output of HQSP-GAN still exhibits a certain degree of over-smoothing. The restoration of high-frequency details, particularly in skin texture and hair strands, remains insufficient, leading to reduced edge acuity in the reconstructed facial features and a generally slightly blurred appearance in the overall image.

Based on the current experimental results, we attribute the performance limitations of the HQSP-GAN model to three main factors: (i) Information loss during the data preprocessing stage: The dimensionality reduction method employed may be overly aggressive, resulting in the loss of critical details from the original images prior to model input, thereby impairing the model's ability to extract and reconstruct high-quality features. (ii) Dimensionality constraints of the training data: The low-resolution images (28×28 pixels) used in the experiments contain limited visual information, which is insufficient for the model to learn rich semantic features and textural details, thus restricting its generative capability. (iii) The design of the quantum circuit (*e.g.*, the number of quantum layers and qubits) and the classical network components (*e.g.*, the depth and width of fully connected layers) may not be well-suited for the generation of this CelebA dataset.

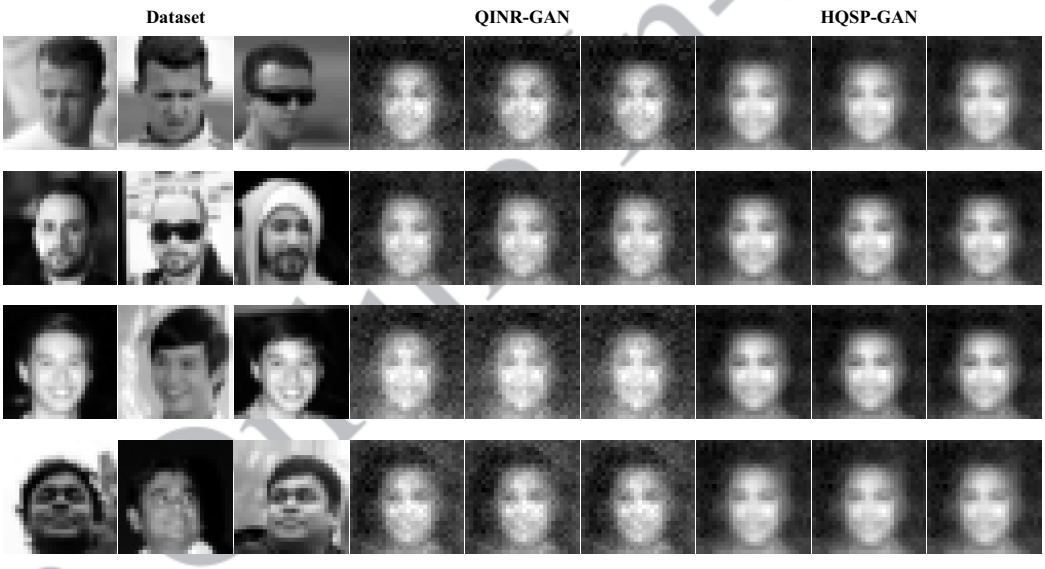


Fig. E1. Comparisons of CelebA image generation by different quantum generative models.

Acknowledgements

Project supported in part by the Natural Science Foundation of Guangxi Zhuang Autonomous Region, China (Grant No. 2025GXNSFAA069491), the National Natural Science Foundation of China (Grant No. 62361021), the Innovation Project of the Guilin University of Electronic Technology (Grant No. 2025YCXS057), the Guangdong Provincial Quantum Science Strategic Initiative (Grants No. GDZX2403001), and the Innovation Program for Quantum Science and Technology–National Science and Technology Major Project (Grant No. 2021ZD0302901).

References

[1] Elasri M, Elharrouss O, Al-Maadeed S and Bouridane A 2022 *Neural Processing Letters* **54** 4609
[2] Yu L, Zhang W, Wang J and Yu Y 2017 *Proceedings of the AAAI Conference on Artificial Intelligence* **31** 1
[3] van den Oord A, Dieleman S, Zen H, *et al*. 2016 arXiv:1609.03499
[4] Goodfellow I, Pouget-Abadie J, Mirza M, *et al*. 2020 *Communications of the ACM* **63** 139
[5] Wang K, Gou C, Duan Y, Lin Y, Zheng X, Wang F and Yang L 2017 *IEEECAA Journal of Automatica Sinica* **4** 588
[6] Creswell A, White T, Dumoulin V, *et al*. 2018 *IEEE Signal Processing Magazine* **35** 53
[7] Arjovsky M, Chintala S and Bottou L 2017 arXiv: 1701.07875
[8] Zhang Z, Luo C and Yu J 2019 *Proceedings of the 28th ACM International Conference on Information and Knowledge Management* pp. 2377–2380
[9] Silver D, Patel T, Cutler W, *et al*. 2023 *Proceedings of the IEEE/CVF International Conference on Computer Vision* pp. 7030–7039
[10] Chang S Y, Thanasilp S, Saux B L, *et al*. 2024 arXiv: 2406.02668
[11] Thomas A M, Youel H and Jose S T 2025 *Quantum Mach. Intell.* **7** 91
[12] Yang X, Zhou R, Jia S Z, *et al*. 2026 *Expert Systems with Applications* **296** 128865
[13] Hu L, Wu S H, Cai W, *et al*. 2019 *Sci. Adv.* **5** eaav2761
[14] Tsang S L, West M T, Erfani S M, *et al*. 2023 *IEEE Transactions on Quantum Engineering* **4** 1

[15] Ma Q G, Hao C L, Si N W, *et al*. 2025 *EPJ Quantum Technology* **12** 3
[16] Steane A 1998 *Reports on Progress in Physics* **61** 117
[17] Gilyén A, Arunachalam S and Wiebe N 2019 *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms* pp. 1425–1444
[18] Rebentrost P, Schuld M, Wossnig L, *et al*. 2019 *New J. Phys.* **21** 073023
[19] Sim S, Johnson P D and Aspuru-Guzik A 2019 *Advanced Quantum Technologies* **2** 1900070
[20] Schuld M, Sweke R and Meyer J J 2021 *Phys. Rev. A* **103** 032430
[21] Pérez-Salinas A, Cervera-Lierta E, Gil-Fuster E, *et al*. 2020 *Quantum* **4** 226
[22] Li J, Topaloglu R O and Ghosh S 2021 *IEEE Transactions on Quantum Engineering* **2** 1
[23] Gulrajani I, Ahmed F, Arjovsky M, *et al*. 2017 *Advances in Neural Information Processing Systems* **30** 5767
[24] Pan M, Wang B, Tao X, *et al*. 2022 arXiv: 2208.09834
[25] LeCun Y, Bottou L, Bengio Y, *et al*. 1998 *Proc. IEEE* **86** 2278
[26] Xiao H, Rasul K and Vollgraf R 2017 arXiv: 1708.07747
[27] Shu R, Xu X, Yung M H, *et al*. 2024 arXiv: 2402.01791
[28] Huang H L, Du Y, Gong M, *et al*. 2021 *Phys. Rev. Appl.* **16** 024051
[29] Heusel M, Ramsauer H, Unterthiner T, *et al.* 2017 *Proceeding of the Advances in Neural Information Processing Systems* **30** 6626
[30] Brunet D, Vrscay E R and Wang Z 2011 *IEEE Transactions on Image Processing* **21** 1488
[31] Xia P, Zhang L and Li F 2015 *Information Sciences* **307** 39
[32] Korhonen J and You J 2012 *2012 Fourth International Workshop on Quality of Multimedia Experience* (IEEE) pp. 37–38

CPB Online In-Press