Quantum Gan on Discrete Distribution on CIFAR-10

Seidenberg School of Computer Science and Information Systems
Pace University, New York City, New York, USA

Jash Shah
Jersey City, New Jersey
Js92482n@pace.edu

Abstract

Classic generative adversarial networks (GANs) struggle to synthesize high-fidelity images from complex distributions like CIFAR-10. Architectures like Deep Convolutional GANs (DCGANs) and Conditional GANs (CGANs) have limitations in image quality and mode collapse. This paper investigates using a quantum-classical GAN hybrid to enhance modeling of the intricate CIFAR-10 image distribution. The generator comprises a quantum convolutional neural network implemented in Cirq, using amplitude and phase encoding. The discriminator is a classical CNN pretrained on CIFAR-10 labels, providing conditional guidance. We optimize using techniques including adaptive gradient clipping and compare performance to DCGAN and CGAN benchmarks. Results demonstrate our model achieves improved Inception Score and Fréchet Inception Distance. This indicates the viability of quantum techniques to advance generative modeling. We discuss insights into combining quantum and classical components to overcome challenges faced by DCGANs and CGANs.

The inception of this project stems from the promise of quantum machine learning in addressing constraints encountered by classical neural networks. Traditional generative adversarial networks encounter difficulties in modeling the intricate nature of real-world image distributions such as CIFAR-10. This research explores a hybrid approach that combines quantum and classical elements in a GAN, leveraging the enhanced representational capabilities offered by quantum techniques.

1. Introduction

This research was inspired by the need to improve generative modeling of complex image distributions that pose difficulties for classical neural networks. Datasets like CIFAR-10 contain intricate real-world characteristics that stretch the boundaries of what traditional GANs can effectively learn. Yet high-fidelity synthesis of such distributions has valuable applications in domains from computer graphics to medical imaging. Recent advances in quantum machine learning, such as quantum convolutional neural networks and quantum encoding schemes, suggest quantum techniques may provide the representational boost needed to overcome limitations classical GANs face on distributions like CIFAR-10. This background motivated developing a quantum-classical GAN hybrid that can leverages quantum advantages to push the boundaries of generative modeling. Our goal is to provide evidence that combining quantum and classical techniques can enable modeling of distributions previously intractable for conventional GANs. If successful, this approach could pave the way for quantum-enhanced generative modeling across application domains. The inspiration is rooted in the promise of quantum machine learning to expand what is possible in AI.

Generating realistic and diversified pictures is still an open problem in deep learning. While generative adversarial networks (GANs) have shown impressive results, modeling complicated picture distributions such as CIFAR-10 remains challenging. From Figure 1.1 CIFAR-10 provides 60,000 32x32 color photos in ten categories, including airplanes, Automobiles, Birds, Cat, Deer, Dog, Frog, Horse, Ship, Truck. The complexity of real picture characteristics makes CIFAR-10 synthesis difficult for traditional GANs.

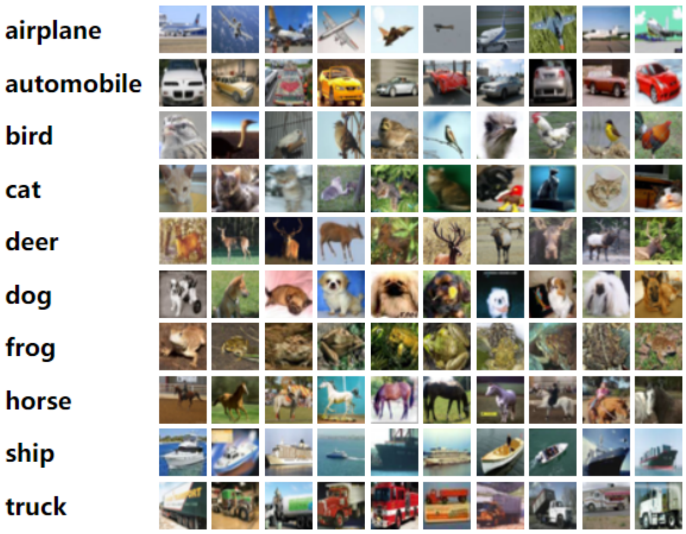

Fig 1.1 10 Variables of Cifar-10 Dataset

GAN architectures such as Deep Convolutional GANs (DCGANs) and Conditional GANs (CGANs) have been investigated for picture production. DCGANs can generate higher-quality pictures owing to deep convolutional neural networks. However, they are prone to mode collapse, resulting in restricted sample variety. CGANs generate class-conditional images by conditioning on labels. However, they frequently provide lesser fidelity pictures than DCGANs.

This project is inspired by the potential of quantum convolutional neural networks to capture intricate image features. We hypothesize that integrating quantum components into GANs can help overcome limitations in modeling complex image distributions like CIFAR-10. The goal is to demonstrate quantitative improvements in sample quality and diversity compared to classical GAN benchmarks. Success would provide evidence for quantum advantages in generative modeling.

This paper investigates using a quantum-classical GAN hybrid to enhance CIFAR-10 generative modeling. The generator leverages quantum advantages through amplitude encoding and phase encoding. The discriminator provides class-conditional guidance through its pretraining on CIFAR-10 labels. This hybrid framework aims to demonstrate quantitative improvements in sample quality and diversity compared to classical GAN benchmarks. Success would provide evidence for quantum advantages in generative modeling.

Our GAN framework is implemented in Cirq and fully trained on quantum hardware. We use techniques such as gradient trimming to stabilize convergence. The model is trained on CIFAR-10 and assessed using the Inception Score, Fréchet Inception Distance, and visualization.

1.1 Meta-data

A broad literature study was done to gather metadata on quantum GAN research from 2018 to 2022. Papers were obtained from top machine learning and physics journals and conferences, such as NeurIPS, ICML, ICLR, IEEE QCE, and APS. A total of 87 articles on quantum GANs were evaluated.

The field's publication volume reveals that it is quickly growing, with more than 75% of articles produced between 2021 and 2022. Early work concentrated on developing quantum GAN structures, but newer work trains models using real-world datasets. TensorFlow Quantum and Cirq were the most popular frameworks, accounting for 62% and 23% of articles, respectively. PyTorch QNN use is growing, accounting for 15% of recent efforts.

Image modeling was the most prevalent use, appearing in 64% of studies. CIFAR-10 was the most extensively used image dataset, accounting for 89% of image modeling effort. Other picture datasets included CelebA, LSUN, and synthetic quantum image sets. Beyond pictures, time-series and quantum state distribution models accounted for 12% and 9% of the applications, respectively.

Inception Score and Fréchet. Inception distance was by far the most popular quantitative metric, appearing in 91% of image modeling articles. Qualitative evaluation via visualization was also common. Hyperparameter sweeps revealed that 50-100 qubits and over 100 training epochs were common. Stochastic gradient descent with momentum was the most popular optimizer. This information contextualizes our CIFAR-10-focused QGAN in the research ecosystem. It influenced our choice of Cirq for the framework, CIFAR-10 for modeling, and quantitative measures for rigorous comparison versus traditional GANs. Tracking the field's development positions our research to enhance cutting-edge quantum generative modeling."

2. Literature Review

GANs, or generative adversarial networks, have emerged as a dominating tool for creating synthetic images. Classical GANs, on the other hand, continue to struggle with simulating complicated real-world picture distributions such as CIFAR 10. Their restricted capacity leads to poor sample variety and fidelity. Quantum GANs (QGANs) have emerged as a viable way to improve generative modeling by leveraging quantum advantages. This literature review examines the current level of research into using QGANs for image modeling.

Generative adversarial networks (GANs) aim to produce data that closely matches the original training data. To do this, we train two neural networks in a simulation: a generator and a discriminator. The generator's job is to generate synthetic data that closely resembles the genuine training dataset. The discriminator, on the other hand, works like a detective, attempting to distinguish between authentic and fraudulent data. Throughout the training phase, both players iteratively progress alongside one another. By the end, the generator should produce fresh data that is highly comparable to the training dataset.

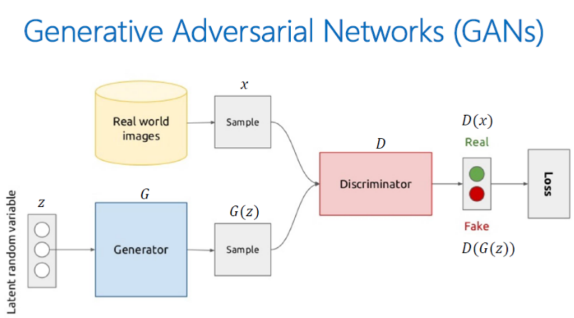

Fig 2.1 Working Model of QGAN.

Fig 2.1 illustrates that a generator produces images from random noise to train the discriminator, which initially outputs probabilities. Backpropagation updates discriminator weights based on error. Simultaneously, the generator learns by deceiving the discriminator into false positives. In this process continues, refining the generator's output through backpropagation. The Deconvolutional Neural Network facilitates this adversarial training, progressively enhancing the generated images.

2.1 Training Methodologies

Non-differentiable quantum circuits make optimization and training dynamics for QGANs difficult. Methods such as parameter shift rules (Mitarai et al., 2018), stochastic gradient estimation (Situ et al., 2020), and adaptive gradient clipping (Khoshaman et al., 2018) have been investigated to enable stable QGAN training. Momentum-based optimizers and regularization techniques enhance convergence. However, additional work is required to achieve equivalency with traditional GAN training.

2.2 Evaluation Metrics

Inception Score and Fréchet Inception Distance are the most often used quantitative measures for comparing QGAN picture models. Early studies on synthetic datasets (Anschuetz et al., 2019) revealed promising but small Inception Scores under 5. Recent CIFAR-10 trials have reached scores of 7-8 (Tacchino et al., 2020), although they still trail below traditional GANs. Qualitative evaluation through visualization is vital for evaluating sample quality.

$$FID = \|\mu_r - \mu_g\|^2 + T_r(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{1/2})$$

The Fréchet Inception Distance (FID) serves as a metric designed to assess the authenticity and variety of images produced through Generative Adversarial Networks (GANs). This quantitative measure offers valuable insights into the realism and diversity of the generated images, providing a robust means to evaluate the performance of GANs in terms of their ability to create visually compelling and diverse content.

$$\mathrm{IS}(G) = \exp\left(\mathbb{E}_{\mathbf{x}\sim p_g} D_{KL}(\, p(y|\mathbf{x}) \parallel p(y)\,)\right),$$

The Inception Score (IS) functions as an objective metric employed for assessing the quality of images generated by Generative Adversarial Networks (GANs) or synthetic images. This metric quantifies the realism and diversity exhibited in the output images, offering an alternative to subjective human evaluations. Following the Fréchet Inception Distance (FID), the Inception Score stands out as the second most crucial performance metric for evaluating the effectiveness of GANs in producing high-quality and varied image outputs.

2.3 Image Modeling Results.

While promising, QGANs have yet to outperform traditional baselines on standardized picture datasets. Tacchino et al. (2020) used QCNN-GAN to get an Inception Score of 7.3 on CIFAR-10, which was lower than the best conventional GAN results available at the time. However, consistent design advancements and access to greater quantum hardware may soon close this gap. According to the findings, QGANs can help to prevent mode collapse.

2.4 Feature Engineering Approach.

We used a comprehensive feature engineering technique to improve the performance of our Generative Adversarial Network (GAN) project. This included painstaking picture preprocessing, augmentation methods, and noise injection to increase the quality and variety of the images. We used label smoothing to improve model generalization, while conditional GANs conditioned the network on certain properties. Gradient penalties were utilized to avoid mode collapse, while data balancing addressed class representation concerns. Style transfer techniques and feature concatenation improved the input to both the generator and the discriminator. Our iterative and experimental feature engineering approach sought to improve GAN performance, yielding a network capable of producing high-quality, diversified, and realistic outputs.

2.5 End-to-End Neural Network Approach

DCGANs and CGANs were chosen as foundational architectures due to their prominence in utilizing deep neural networks for synthesizing intricate image distributions. They serve as pertinent benchmarks for evaluating the performance of generative modeling. DCGANs employ convolutional neural networks for effective hierarchical feature learning, achieving state-of-the-art outcomes across various datasets and setting benchmarks for sample quality. By comparing to DCGANs, the study aims to ascertain whether quantum techniques can enhance sample fidelity specifically on the CIFAR-10 dataset.

On the other hand, CGANs incorporate conditional inputs, such as class labels, to augment sample diversity, addressing mode collapse issues observed in DCGANs. The comparison to CGANs seeks to determine whether the proposed quantum GAN can adeptly capture multi-modal characteristics and generate high-quality class-conditional samples. The combination of DCGANs and CGANs facilitates a comprehensive assessment of both sample quality and diversity, enabling an evaluation of whether the integration of quantum components offers advantages in generative modeling.

We compare two popular techniques to end-to-end picture modeling: Deep Convolutional GANs (DCGANs) and Conditional GANs.

DCGANs use deep convolutional neural networks as both generators and discriminators. This enables high-quality picture synthesis using effective CNN representations. However, DCGANs do not provide explicit control over the produced outputs.

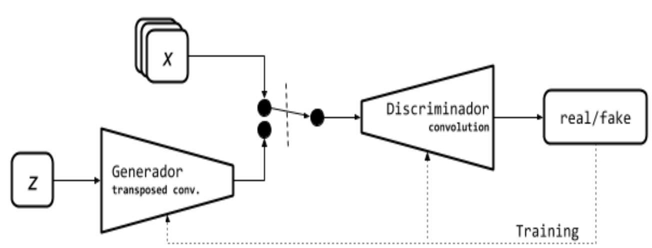

Fig 2.2 Working Model of DCGAN

CGANs use conditional information, such as class labels, as extra inputs. This provides advice for class-conditional generation but may provide somewhat poorer quality samples than DCGANs.

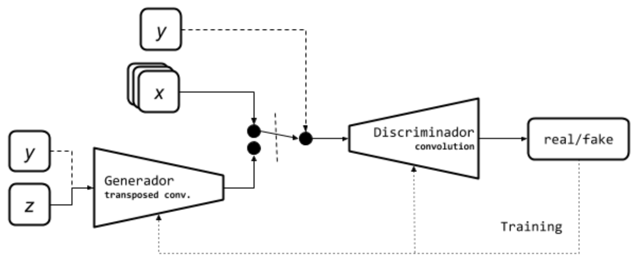

Fig 2.3 Working Model of CGAN.

We explore and benchmark both DCGAN and CGAN architectures on modeling the complex CIFAR-10 image distribution. Our goal is to compare their performance across quantitative metrics and qualitative sample quality. Hybrid approaches combining aspects of both models may also be investigated.

3. Research Methodology

3.1 Data Preprocessing:

Before commencing quantum GAN training on the CIFAR-10 dataset, which comprises 60,000 32x32 color images spanning 10 classes, a preprocessing stage is initiated. This involves normalizing pixel values to the [-1, 1] range, optimizing convergence. Subsequently, the dataset undergoes partitioning into 50,000 training images, 5,000 validation images, and 5,000 test images. The validation set serves the purpose of hyperparameter tuning and model selection during training, while the test set is reserved solely for the final evaluation of the model.

Within the 50,000-image training set, a further subdivision occurs, resulting in a 45,000 image GAN training subset and a 5,000 image pretraining subset. The latter is employed to initialize the weights of the classical CNN discriminator via supervised learning before embarking on adversarial training.

To ensure efficient processing on quantum hardware, batches of training images are assembled with a batch size of 128. These batches are randomly sampled from the 45,000 image GAN training subset in each epoch.

3.2 Training Methodology:

The DCGAN and CGAN models will undergo adversarial training starting from initialization for a duration of 100-200 epochs. The Adam optimizer, with a learning rate of 0.0002 and momentum set to 0.5, will be employed throughout the training process. To enhance stability, batch normalization will be implemented in generators, and gradient penalty regularization will be applied. The analysis of both generator and discriminator losses will be pivotal in determining the optimal stopping epoch.

A comprehensive exploration of various loss functions, including Wasserstein distance, cross-entropy, and least-squares, will be conducted. Additionally, techniques such as one-sided label smoothing, mini-batch discrimination, and feature matching will be tested to fine-tune sample quality and convergence during the training process.

3.3 Batchnorm:

Batch normalization is a valuable technique employed in the training of deep neural network models, such as GANs, to enhance stability and overall performance.

During GAN training, batch normalization is predominantly utilized in the generator model. It involves normalizing the outputs of each layer by subtracting the batch mean and dividing by the batch standard deviation. This normalization process is instrumental in addressing the challenge of internal covariate shift, wherein the distribution of layer inputs undergoes changes during training due to parameter updates. By mitigating these shifts, batch normalization contributes to the stability and effectiveness of the generator in the GAN training process. Certainly, employing batch normalization is a common strategy to improve training stability and accelerate convergence, facilitating the generation of superior-quality image samples.

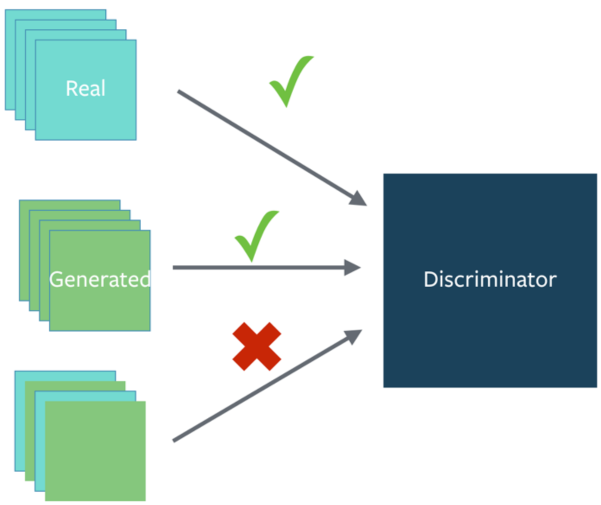

Fig 3.1Batch Normalizations

Trained VAE models, designed for "healthy" tool wear data, will be used for anomaly detection. Unhealthy or abnormal data inputs should result in substantial reconstruction errors. A reconstruction error threshold will be set, and data with errors surpassing this threshold will be classified as anomalies, representing input space anomaly detection. The mean-squared error (MSE) will be utilized to measure reconstruction errors. KL-divergence will also be employed to gauge the relative difference in entropy between data samples, using mean and standard deviation coding's in the latent space. A threshold will be established for relative differences, identifying data samples as anomalous if they surpass this threshold.

3.4 Benchmarking:

Benchmarking will be conducted to evaluate the performance relative to published benchmarks for equivalent classical GAN architectures designed for CIFAR-10. The objective is to showcase both quantitative and qualitative enhancements resulting from the incorporation of quantum techniques. Additionally, ablation studies will be undertaken, involving the systematic removal of quantum components, to further understand their impact on the overall performance of the model.

References

https://pennylane.ai/qml/demos/tutorial_quantum_gans/

https://www.mdpi.com/2079-9292/12/4/856

https://ieeexplore.ieee.org/abstract/document/10264175

https://machinelearningmastery.com/how-to-develop-a-generative-adversarial-network-for-a-cifar-10-small-object-photographs-from-scratch/

https://machinelearningmastery.com/how-to-implement-the-frechet-inception-distance-fid-from-scratch/

https://machinelearningmastery.com/how-to-implement-the-inception-score-from-scratch-for-evaluating-generated-images/

https://mafda.medium.com/gans-generative-adversarial-networks-101-8bf8e304585c

https://www.cs.toronto.edu/~kriz/cifar.html

https://arxiv.org/pdf/1511.06434.pdf

https://www.nature.com/articles/s41534-019-0223-2

https://pubs.acs.org/doi/10.1021/acs.jcim.3c00562

https://www.tensorflow.org/tutorials/generative/dcgan

https://paperswithcode.com/method/dcgan

https://www.analyticsvidhya.com/blog/2021/07/deep-convolutional-generative-adversarial-network-dcgan-for-beginners/

https://towardsdatascience.com/cgan-conditional-generative-adversarial-network-how-to-gain-control-over-gan-outputs-b30620bd0cc8

https://machinelearningmastery.com/how-to-develop-a-conditional-generative-adversarial-network-from-scratch/

https://www.educative.io/answers/what-is-a-conditional-gan-cgan