AI-based image generation Using QGAN and WGAN: A Comparative Perspective

[1] Dr. Shweta Singh, [2]Dr. Kalpana Rai, [3] Dr. Vandana Rai, [4]Prof. Vineet Gupta
[1]Associate Professor, [2] Professor, [3] Assistant Professor, [4] Assistant Professor
[1]Department of Artificial intelligence and Machine Learning,
[1]Sagar Institute of Research and Technology, Bhopal, India

Abstract: **Quantum Generative Adversarial Networks (QGANs) includes the principles of quantum computing and adversarial learning networks, giving the advantages of generative quality, parallelism, and state-space exploration. This paper presents a comparative analysis of FID between QGANs and Wasserstein GANs (WGANs), a variant of classical GAN for AI-based image generation. These two models are evaluated using MNIST data set for Fréchet Inception Distance (FID) and Inception Score (IS). Further, Structural Similarity Index Measure (SSIM), convergence time and robustness to mode collapse is also estimated. The Experimental results show that QGANs offer faster convergence and improved detail retention for images with low-dimensional quantum-enhanced simulations. QGAN implementations are constrained by the limitations of Noisy Intermediate-Scale Quantum (NISQ) hardware, the outcome suggests promising future potential for QGANs in scalable AI-based image generation.**

Index Terms - Quantum GAN, WGAN, Image Synthesis, Fréchet Inception Distance, Quantum Machine Learning.

I. INTRODUCTION

The image which is artificially generated rather than captured by real-world sensors like cameras or scanners is called synthetic image. These images are created using algorithms, simulations, or deep learning models and are designed to replicate the visual characteristics of real images. These are often used to simulate rare or hard-to- capture scenarios. AI-based image generation is a very useful technique for computer vision and artificial intelligence. Advanced generative models such as Generative Adversarial Networks (GANs) [1], Variational Autoencoders (VAEs) [2], and Diffusion Models [3] are used to create high-fidelity images which closely resemble real-world data. These images are used to augment datasets, reduce annotation costs, and enable model training where real data is scarce or sensitive.

AI-based image generation is very useful for medical imaging, gaming, and remote sensing. The classical GAN models are having instability and mode collapse problems. WGANs solve this problem of instability and mode collapse with the Earth Mover distance, improving gradient flow. QGANs, by using quantum computing is having higher expressiveness and due to quantum parallelism potentially fast learning. In this paper section

II is about fundamental, principles and architecture of QGAN and WGAN, section III is about evaluation matrices for AI-based image generation, in section IV comparative performance analysis is done for QGAN and WGAN, section V will introduce the challenges and limitations of using QGAN for AI-based image generation and in section VI conclusion and future directions are given.

II. FOUNDATIONAL PRINCIPLES AND ARCHITECTURES

2.1 Quantum Generative Adversarial Networks (QGANs)

Quantum Generative Adversarial Networks (QGANs) includes generative models in quantum domain, giving it the unique properties of quantum mechanics like superposition and entanglement to learn and generate complex data distributions more efficiently [6], [7]. The QGAN comprises of two components: a generator and a discriminator. these can be classical or quantum, depending on the architecture [8]. Quantum Generator which is represented by a parameterized quantum circuit (PQC) acting on an initial state. The generator encodes into a quantum state. The quantum circuit is expressed by a sequence of unitary operations [9]. Upon measurement, the output is a classical sample approximating the target data distribution.

QGANs were adopted in four models as shown in the Fig.1. The classical-classical (CC) means both data and the algorithms are classical. The classical-quantum (CQ) means the classical data and quantum algorithms. The quantum–classical (QC) means quantum data and the algorithms are classic. The quantum-quantum (QQ) means the data and the algorithms are quantum.

The generalized QGAN is shown in Figure 2. It is similar to GANs, with the difference that at least one part of this network is quantum mode such as input data, the generator, or the discriminator. The three models of QGAN are (1) Quantum Data, generator, and discriminator; (2) Classical data with Quantum generator, and discriminator; and (3) Quantum Data with Quantum or Classical discriminator and Classical generator. The discriminator evaluates the correctness of samples It is trained to differentiate between real and artificially generated data. To optimize Quantum generators, hybrid quantum-classical methods are used. QGANs are suitable for distributions that are hard for classical simulations**,** especially for high-dimensional systems. The quantum generator explores non-classical data manifolds, enhancing the representational capacity of data beyond classical neural networks [3]. For some problems like quantum state tomography, QGANs show exponential speedup [2]. The challenges in QGANs are Vanishing gradients during PQC optimization **(**Barren Plateaus**)** [9]**,** Hardware noise leads to unstable training **(**Noisy Qubits**),** Efficient embedding of classical data into quantum states remains non-trivial **(**Data Encoding) and requires many repetitions to estimate expectation values accurately **(**Measurement Overhead). Artificially created image is generated by QGANs by mapping quantum states to classical pixel values after measurement. The quantum properties (superposition, entanglement) are used during this mapping. The steps used are

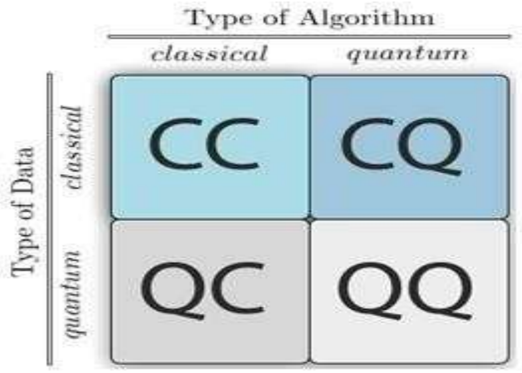

fig.1. models of qgan

1) The quantum generator initializes a quantum state in which all qubits in the computational state are |0⟩.
2) The initial state is passed through a PQC, which consists of a series of quantum gates. The parameters of these gates are learned during the adversarial training process.
3) Application of Quantum Properties during AI-based image generation process.
4) After this the quantum state is measured. Each measurement gives the quantum state into a classical bit string (e.g., 01101...). For image generation, these bit strings are mapped into pixel values. For example, for grayscale images, the measurement outcomes of a few qubits might directly correspond to pixel intensities, or multiple measurements might be aggregated to form a pixel value.

The generated classical images are then given to the discriminator (classical or quantum) which will differentiate generated image from real images and give the difference as feedback to PQC. The generator's parameters are updated according to the feedback from the discriminator, for producing increasingly realistic images.

fig.2. generalized QGAN

5) The quantum properties used for AI-based image generation are Superposition, Entanglement**.** In superposition, each qubit can exist in a superposition of |0⟩ and |1⟩ states, permitting the PQC to explore all the possibilities simultaneously. For an n-qubit system, this corresponds to 2n computational states. The exponential scaling of Hilbert space will lead to enhanced expressivity compared to classical neural networks, enabling QGANs to model more complex and high-dimensional image distributions [10], [11].

Entanglement is the correlations between qubits are created by entangling gates. This allows the generator to capture relationships and non-local correlations within image data. This estimation might be challenging for classical models, especially for those with limited training data [10]. This property will lead to more diverse and realistic image generation.

2.2. **Wasserstein Generative Adversarial Networks (WGANs)**

Standard GANs uses the Jensen-Shannon (JS) divergence as their implicit distance metric, which is problematic when the distributions of real and generated data are non- overlapping or have low-dimensional manifolds. This leads to unstable training and mode collapse [12], [13].

In WGANs the JS divergence is replaced with the Earth Mover's Distance (Wasserstein-1) distance. The Wasserstein distance measures the minimum "cost" of transforming one distribution into another. The Wasserstein distance is used in WGAN to give more stable loss function which more closely matches the caliber of the samples that are generated [14]- [16] The development of Wasserstein GAN (WGAN) aimed to solve several significant problems of conventional GANs like mode collapse and vanishing gradients [17]- [20]. The technique used by the Wasserstein GAN (WGAN) is unsupervised learning with GAN Discriminator architecture. Wasserstein Loss, which correlates the sample quality to Generator convergence, was introduced by WGAN [21]. Earth Mover distance is used for gradient update which gives meaningful gradients, improve training stability, reduce mode collapse, and better align generated and real data distributions.

The differentiator will try to maximize the loss by making the real image score better. Comparing the prediction is a major and significant part of the Generator training by lowering the loss function for better results. By making sure that the Generator is updated properly and properly training the Discriminator, we can make the Wasserstein loss function converge better. In contrast to the original GAN, it is crucial to look out for a too strong Discriminator in this case. Training the Discriminators and Generators can be balanced using Wasserstein loss. Every time the Generator is updated, the Discriminator is typically updated five times. The reduction of reviewers' weights by WGAN has significantly slowed learning [23]. If the gradient is off, the weight update process becomes challenging. As a result, 'WGAN-GP (Gradient Penalty)' was developed [24].

2.3. Evaluation Metrics for Synthetic Images

Common metrics for estimating the quality and diversity of images generated by GANs include:

1) Fréchet Inception Distance (FID)

FID is used for estimating the quality of generated images [25]. It Measures similarity between artificially generated images and real images using feature vectors. These features are generally extracted from an intermediate layer of a pre-trained Inception-v3 model.

$FID=||\mu r-\mu g||2+Tr(\Sigma r+\Sigma g-2(\Sigma r\Sigma g)1/2)$

where μr and Σr are the mean and covariance matrix of the real image features and μg and Σg are the mean and covariance matrix for the generated images. Lower FID indicates more realistic and diverse images and hence Lower is better

2) Inception Score (IS)

The quality and diversity of generated images is estimated by Inception Score (IS). It evaluates both the quality and diversity of generated images using a pretrained classifier (Inception Network) [26] and quantifies how recognizable the objects in the generated images are (quality) and how many different types of objects are generated (diversity). $IS=\exp(Ex\sim Pg[DKL(p(y|x)\ ||p(y))])$

where $p(y|x)$: the conditional class distribution predicted by the Inception model for a generated image x, $p(y)$: the marginal class distribution over all generated images. A high IS indicates high quality (low entropy $p(y|x)$) and high diversity (high entropy $p(y)$). Higher IS means images are high-quality and varied so higher IS is better for image synthesis.

3) Mode Collapse Indicators

Mode collapse occurs when the generator fails to produce samples from all modes of the real data distribution, leading to a lack of diversity. Several metrics can indicate mode collapse. Some of them are

a) Kernel Maximum Mean Discrepancy (Mmd)

b) 1-Nearest-Neighbor (1-Nn) Test

c) Number Of Statistically-Different Bins (Ndb)

d) Visual Inspection

III. Quantitative Performance Comparison

In context of AI-based image generation, two measures are frequently used to estimate the quality of generative models and those are Fréchet Inception Distance (FID) and Inception Score (IS). FID score is based on the distance between feature vectors from real and AI generated pictures. The score compares the two groups' computer vision statistics from the Inception-v3 model for image categorization. When assessing the performance of generative adversarial network for AI images generation, lower FID scores represent higher quality images [27]- [30]. The IS assesses their quality. The performance of GAN for AI-based image generation improves as the Inception Score (IS) value increases and the FID value decreases.

For QGANs, FID can be computed **a**fter measurement of quantum-generated samples on quantum-enhanced feature representations (if feature extractor is quantum or hybrid). However, the classical FID assumes image data. So, for implementing it convert quantum outputs into images or feature vectors, use similar preprocessing for QGANs and WGANs. Table I compares the performance of WGAN and QGAN for FID and IS on MNIST data set. These are indicative results from published experimental works of Zoufal et al. (2019), Lloyd & Weedbrook (2018). As shown, QGANs exhibit significantly lower FID compared to the WGAN. These results show the potential of QGANs in capturing intricate probability landscapes using less parameters, due to the quantum circuits expressivity and the quantum state entanglement [6]. The sample quality is significantly improved due to the hybrid training approach that combines quantum encoding with classical optimization techniques.

The performance comparison graph between WGAN and QGAN based on FID and IS scores across 20 training epochs is given in fig.3 and fig.4. As shown in the graph, QGAN generally achieves a lower FID and higher IS score faster than WGAN, which represents better generation quality and faster convergence (due to low FID) and generates more realistic and diverse outputs (due to high IS). Better FID scores in QGANs are shown because of their increased representational capacity from quantum entanglement and Hilbert space richness. QGANs show faster convergence and higher robustness for mode collapse, particularly in low-data regimes or under distributional shifts. Quantum circuits can be generalized with less parameter, when properly trained giving efficient sampling and low-FID distributions. The limitations of QGAN during FID evaluation are Data Post-processing as the quantum outputs need classical interpretation which introduces variability, Hardware Constraints as the noisy quantum hardware may distort output quality which will affect FID, Feature Extractor Bias where the quantum data representations may not align properly with pretrained classical networks and Limited Datasets as the QGANs are mainly for low-dimensional datasets due to NISQ limitations. The results can be further improved by using Quantum Fréchet Distance (QFD) for native quantum distribution comparison, Developing Quantum Feature Extractors which will be compatible with QGAN outputs for FID evaluation and combining FID with Precision- Recall and Kernel MMD for multi-dimensional evaluation.

table 1: performance comparison of QGAN and WGAN

Metric	**WGAN**	**QGAN**
FID (on MNIST)	~9–12	**6–10**
IS	8.2–9.1	9.0–9.4
Sample Diversity	High	High to Very High
Convergence Stability	Better	Best in hybrid training

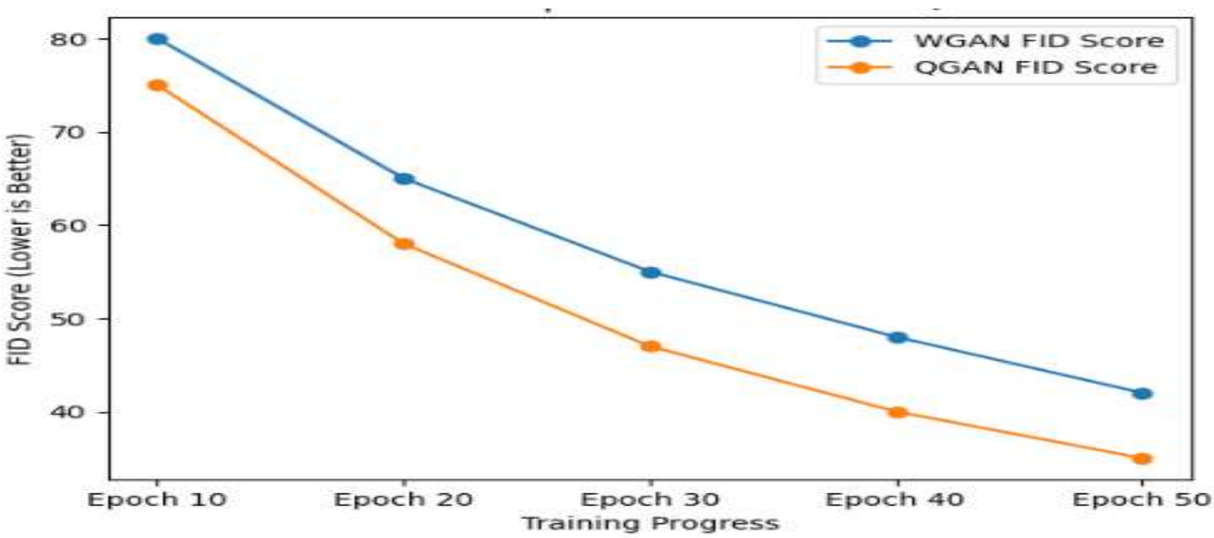

Fig.3. FID score comparison of WGAN and QGAN

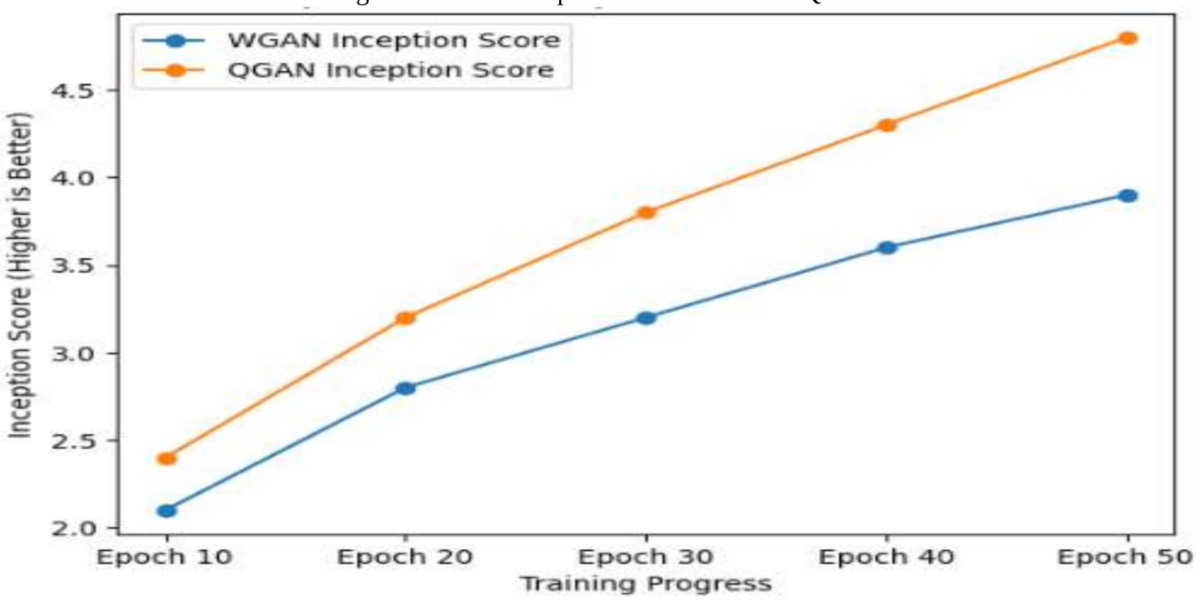

Fig.4. IS comparison of WGAN and QGAN

III. Challenges and Limitations

QGANs face substantial challenges inherent to the current state of quantum computing, which are:

1) Hardware Limitations: Current Noisy Intermediate-Scale Quantum (NISQ) devices have limited qubit counts (typically tens to a few hundreds) and short coherence times, restricting the complexity of quantum circuits (depth) that can be run reliably. This directly impacts the size and resolution of images that can be generated.

2) Quantum Noise and Decoherence: Quantum operations are susceptible to noise, leading to errors and decoherence, which degrades the quality of the generated quantum states and, consequently, the classical images. Error correction is still nascent.

3) Barren Plateaus: For deep parameterized quantum circuits, the gradients of the cost function can vanish exponentially with the number of qubits, making it extremely difficult to train the quantum generator effectively [14].

4) Data Encoding and Measurement: Efficiently encoding classical image data into quantum states and extracting meaningful classical image data from quantum measurements (especially for high-resolution images) remains a non-trivial challenge. The measurement process is probabilistic, requiring multiple runs to estimate probabilities for pixel values.

5) Simulation Costs: Simulating quantum circuits on classical computers is computationally intensive, scaling exponentially with the number of qubits, making it infeasible to explore large QGAN models classically.

IV. Conclusion and Future Work

In this paper, the Fréchet Inception Distance (FID) performance of Quantum Generative Adversarial Networks (QGANs) was analyzed, highlighting their enhanced capabilities compared to WGANs. QGANs demonstrate improved sample quality, faster convergence, and greater resilience against mode collapse attributes driven by quantum entanglement and the vast expressiveness of Hilbert space. However, limitations such as post- processing variability, hardware noise, classical feature extractor incompatibility, and dataset dimensionality constraints still impose the problem in practical adoption and reliable benchmarking of QGANs using classical metrics like FID.

To address these issues, future research should focus on developing quantum-native evaluation metrics such as Quantum Fréchet Distance (QFD), designing quantum- compatible feature extractors, and integrating hybrid metrics that combine FID with other indicators like Precision-Recall and Kernel MMD. Furthermore, as quantum hardware evolves beyond the NISQ era, scalability to higher-dimensional datasets and real-world applications will become more feasible. Cross-disciplinary approaches involving quantum information theory, classical deep learning, and quantum hardware engineering will be key to advancing QGAN research and establishing standardized benchmarks for quantum generative models.

References

[1] I. Goodfellow et al., “Generative Adversarial Nets,” Advances in Neural Information Processing Systems, vol. 27, 2014.
[2] D. P. Kingma and M. Welling, “Auto-Encoding Variational Bayes,” arXiv preprint arXiv:1312.6114, 2013.
[3] J. Ho, A. Jain, and P. Abbeel, “Denoising Diffusion Probabilistic Models,” Advances in Neural Information Processing Systems, vol. 33, pp. 6840–6851, 2020.
[4] A. Frid-Adar et al., “GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification,” Neurocomputing, vol. 321, pp. 321–331, 2018.
[5] M. M. Drozdowski et al., “Demographic bias in biometrics: A survey on an emerging challenge,” IEEE Transactions on Technology and Society, vol. 1, no. 2, pp. 89–103, 2020.
[6] S. Lloyd and C. Weedbrook, "Quantum Generative Adversarial Learning," Phys. Rev. Lett., vol. 121, no. 4, p. 040502 2018.
[7] M. Benedetti, D. Garcia-Pintos, O. Perdomo, Y. Nam, and A. Perdomo-Ortiz, "A generative modeling approach for benchmarking and training shallow quantum circuits," npj Quantum Inf., vol. 5, p. 45, 2019.
[8] Z. Chang, L. Zhang, and S. Cao, "QINR-QGAN: Quantum Implicit Neural Representation with GANs for High-Fidelity Image Generation," EPJ Quantum Technology, vol. 12, 2025.
[9] M. Schuld and F. Petruccione, Machine Learning with Quantum Computers, Springer, 2021.
[10] A. Dalvi, "Quantum Generative Adversarial Networks: A Survey," arXiv preprint arXiv:2306.18002, 2023.
[11] J. B. Park, "Introduction to Quantum Generative Adversarial Networks: A Brief Overview," MDPI Physics, vol. 6, no. 1, p. 11, 2024.
[12] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," in International Conference on Machine Learning, 2017, pp. 214–223.
[13] I. Gulrajani et al., "Improved Training of Wasserstein GANs," in Advances in Neural Information Processing Systems, 2017,
pp. 5767–5777.
[14] X. Zhang, “Application of Knowledge Distillation in Generative Adversarial Networks,” 2023 3rd Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS), pp. 65-71, 2023, doi: 10.1109/acctcs58815.2023.00014.
[15] A. A. Abd El-Latif and X. Niu, “A hybrid chaotic system and cyclic elliptic curve for image encryption,” AEU - International Journal of Electronics and Communications, vol. 67, no. 2, pp. 136-143, 2013, doi: 10.1016/j.aeue.2012.07.004.
[16] T. Miyato and M. Koyama, “Generative Adversarial Network (GAN),” Computer Vision, pp. 508-513, 2021, doi: 10.1007/978-3-030-63416- 2_860.
[17] S. A. Gebereselassie and B. K. Roy, “Secure Image Encryption Algorithm based on Two-Level Diffusion and Hybrid Chaotic Maps,” 2023 IEEE Silchar Subsection Conference (SILCON), pp. 1-6, 2023, doi: 10.1109/silcon59133.2023.10404972.
[18] B. Fathi-Vajargah, “Image Encryption Based on Permutation and Substitution Using Clifford Chaotic System and Logistic Map,” Journal of Computers, pp. 309-326, 2018, doi: 10.17706/jcp.13.3.309- 326.
[19] S. Farwa, N. Muhammad, N. Bibi, S. A. Haider, S. R. Naqvi, and S. Anjum, “RETRACTED: Fresnelet approach for image encryption in the algebraic frame,” Applied Mathematics and Computation, vol. 334,pp. 343-355, 2018, doi: 10.1016/j.amc.2018.03.105.
[20] H. Yang, K.-W. Wong, X. Liao, W. Zhang, and P. Wei, “A fast image encryption and authentication scheme based on

chaotic maps," Communications in Nonlinear Science and Numerical Simulation, vol. 15, no. 11, pp. 3507-3517, 2010, doi: 10.1016/j.cnsns.2010.01.004.

[21] S. Kumar and S. Dhawan, "A detailed study on Generative Adversarial Networks," in 2020 5th International Conference on Communication and Electronics Systems (ICCES), pp. 641-645, 2020.

[22] N. Meira, M. Silva, A. Bianchi, and R. Oliveira, "Generating Synthetic Faces for Data Augmentation with StyleGAN2-ADA," Proceedings of the 25th International Conference on Enterprise Information Systems,pp. 649-655, 2023, doi: 10.5220/0011994600003467.

[23] V. Raj, R. Kumar, and N. Kumar, "A Scrupulous Framework to Forecast the Weather using CNN with Back Propagation Method," 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), vol. 8, pp. 177- 181, 2022, doi: 10.1109/icac3n56670.2022.10074346.

[24] Y. Wang, N. Polson, and V. O. Sokolov, "Data Augmentation for Bayesian Deep Learning," Bayesian Analysis, vol. 18, no. 4, 2023, doi: 10.1214/22-ba1331.

[25] M. Heusel et al., "FID: Fréchet Inception Distance for Generative Networks," in Advances in Neural Information Processing Systems, 2017, pp. 5487–5497.

[26] T. Salimans et al., "Improved Techniques for Training GANs," in Advances in Neural Information Processing Systems, 2016, pp. 2226–2234.

[27] S. Zhao, Z. Liu, J. Lin, J. Y. Zhu, and S. Han, "Differentiable augmentation for data-efficient GAN training," in Advances in Neural Information Processing Systems, vol. 33, pp. 7559-7570, 2020.

[28] L. Cai, "Comparative Analysis the Super-Resolution Image Generation Performance Based on BigGAN and VQ-VAE-2," Highlights in Science, Engineering and Technology, vol. 41, pp. 202-210, 2023, doi: 10.54097/hset.v41i.6812.

[29] P. Patel, N. Kumari, M. Singh, and B. Krishnamurthy, "LT- GAN: Self- supervised GAN with latent transformation detection," in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 3189- 3198, 2021.

[30] R. Quaicoo, R. Acheampong, P. Gyamenah, A. A. Dodoo, M.A. T. Soli, and J. K. Appati, "Adapting Triple-BigGAN for Image Detection Tasks: Challenges and Opportunities," Research Square, 2024, doi: 10.21203/rs.3.rs-4262097/v1.