Image Fidelity in Generative Adversarial Network: Challenges and Solutions for Small Datasets

Zhengyang Dong[1*]

[1]University of International Business and Economics, No.10 Huixin Dongjie, Chaoyang District, Beijing, China

Abstract. This paper systematically summarizes the challenges and countermeasures in image generation when the training samples of the generated countermeasure network (GAN) are very few (less than 100 – 1000). The research proposes a cross method comparison framework, which classifies the existing strategies into three categories: data enhancement (such as DiffAug and ADA), regularization and structure improvement (such as ContraGAN and FastGAN), and cross domain adaptation (such as FreezeD and SSR). Experiments show that DiffAug performs better under very low sample size, and ADA is suitable for slightly large-scale data; Regular methods improve diversity and structural integrity, and migration strategies enhance cross domain performance. The study emphasized that the generation quality should be evaluated by combining multiple indicators such as KID and precision recall rate, and pointed out that in the future, attention should be paid to enhanced regular joint scheduling, generalized migration paradigm and more reliable evaluation system.

1 Introduction

1.1 Potential and application scenarios of GAN in image generation

In recent years, Generic Adversary Networks (GANs) have demonstrated high fidelity and diversity generation capabilities in medical imaging, style migration, virtual content generation and other fields [1-2]. Its adversarial training mechanism, especially in high-resolution image generation and synthetic data enhancement, significantly improves the visual quality and the performance of downstream tasks (such as classification and segmentation) [1, 3].

1.2 Image fidelity under small sample training

When the training data is extremely limited (the number of samples is<100 – 1000), GAN will encounter the following challenges: Discriminator overfitting will lead to memorizing

* Corresponding author: 202215026@uibe.edu.cn

training samples, resulting in training feedback deviation[4], training instability or convergence; Pattern collapse shows that different potential vectors generate highly similar samples, which seriously lacks diversity [2]; structural distortion and detail blur make it difficult for the generator to learn high-frequency details and microstructure, which may lead to image blur or structural distortion [5]. These challenges have been confirmed in the research of DiffAugment and other methods [1, 4].

1.3 Objectives and contributions of this paper

The purpose of this review is to systematically classify the core challenges that affect image fidelity under the condition of small samples, and build a cross method comparative analysis framework, covering data enhancement, regularization architecture [1-4], self-monitoring mechanism and cross domain adaptation [6-7]. Different from the research focusing on the evolution of technical details (such as StyleGAN-ADA series), this paper emphasizes the perspective of "cross method comparison and comprehensive strategy evaluation".

2 GAN core challenges under small sample conditions

2.1 Discriminator over fitting

When the training samples are small (for example,<1k), the discriminator is easy to remember the training set, provide the generator with a non generalized gradient, and train oscillations or even divergences; At the same time, too strong discriminator suppression will induce the generator to degenerate to "shortcut" (copy samples or shrink to narrow distribution) in order to score higher. Experience and theory show that in low data scenarios, adaptive/consistency enhancement and regularization can significantly alleviate this phenomenon [1, 4]. Typical symptoms include rapid improvement of FID in the early training period, but then rebound; The generated results are highly similar to a few images in the training set; The discriminator training loss continues to decline while the verification distribution statistics deteriorate.

2.2 Colour illustrations

The generator maps a large number of different latent variables to repeated or highly similar samples, resulting in insufficient diversity. Small samples magnify this risk: insufficient data coverage+discriminator overfitting jointly result in narrow feasible region. It is proved that adversarial contrastive learning, small batch relationship modelling and consistency regularization can improve the separation between inside and outside the category and alleviate the collapse [2]. Typical symptoms include high MS-SSIM, low PR curve recall, and visual appearance of "same face with different posture" or texture repetition.

2.3 Structural distortion and fuzzy details

Data scarcity makes it difficult for the generator to learn high-frequency texture and geometric structure; Excessive data enhancement or improper consistency constraints may also "smooth out" details. In the case of small samples, the feature quality and lightweight structure promotion gradient of the self supervised characterization intensifier discriminator are effectively used, which is proved to be able to maintain the structure fidelity when there are less than 100 samples [3]. Typical signs include edge collapse, "wax feeling" of material, and local configuration dislocation; FID improved but perception sharpness was insufficient.

2.4 Interdependence of challenges

Overfitting, collapse and distortion are not isolated: discriminator overfitting will lead to generator shrinkage, further leading to mode collapse; The strong data enhancement introduced to suppress overfitting may sacrifice details [1, 4]. Therefore, a single means cannot actually solve the problem, and the combination strategy needs to be balanced between "generalization diversity detail".

The core crux of small sample training lies in the imbalance of discriminator generator strength and insufficient statistical coverage. Therefore, the evaluation of subsequent schemes should focus on three main axes: (1) whether the discriminator overfitting is relieved; (2) Whether to improve diversity without losing details; (3) Whether it can be combined with other means.

3 Scheme classification and discussion of small sample GAN fidelity

3.1 Data enhancement driving method

ADA (adaptive discriminator enhancement) dynamically adjusts the enhancement intensity according to the discriminator overfitting signal; Training from zero or fine tuning is available; Significantly stable small sample training. Its advantages are plug and play and architecture friendly; The risk is that if the enhancement is too strong, the high-frequency details will be weakened, and intensity scheduling is required [8]. DiffAug (slightly enhanced) is used symmetrically on true/false samples, and end-to-end back-propagation; On multiple benchmarks, FID/IS is significantly improved with a small amount of data. The advantages are high efficiency and versatility; Pay attention to the balance with discriminator capacity/learning rate to avoid over smoothing [1]. To quantitatively compare the performance differences between ADA and DiffAug, the following table 1 summarizes the key indicators reported in relevant literature:

Table 1. Performance comparison of data enhancement methods in small sample scenarios [1]

Method	Data set	Sample size	FID	IS	Applicable scenario
StyleGAN2 baseline	FFHQ	1k	62.16	-	no enhanced baseline, high FID
StyleGAN2 + ADA	FFHQ	1k	21.29	-	more stable with slightly more samples (about 1k), suitable for balanced scenarios
StyleGAN2 + DiffAug	FFHQ	1k	25.66	-	more advantages for small samples ($\leqslant$ 1k) of high fidelity tasks
StyleGAN2 baseline	LSUN-Cat	1k	182.85	-	FID is extremely high and cannot be used at all
StyleGAN2 + ADA	LSUN-Cat	1k	43.25	-	good diversity, suitable for scenes with minor details
StyleGAN2 + DiffAug	LSUN-Cat	1k	42.26	-	stable and slightly superior, suitable for applications requiring detail preservation
BigGAN + DiffAug	ImageNet-128	-	6.80	100.8	High resolution large-scale task with strong portability of DiffAug

According to the data in the table above, DiffAug and ADA both significantly reduce FID, with similar effects, and are key means to improve fidelity with few samples (<1k).

DiffAug usually gets a lower FID value, indicating that its generated image is closer to the real distribution, especially in high-resolution tasks with high requirements for detail fidelity (such as ImageNet-128). Therefore, if the sample size is extremely limited (<1k) and high fidelity is pursued, DiffAug is more applicable; If the sample size is slightly larger and diversity needs to be considered, ADA may be a better choice.

3.2 Model architecture and regularization innovation

Consistency/LeCam class regularization improves the generalization and gradient stability of discriminators by imposing smoothing/consistency constraints on discriminators, and complements enhancement methods [4]. ContraGAN explicitly models the sample sample and sample class relationship in the discriminator to alleviate the collapse within the class and improve the discrimination [2]. The lightweight structure+self supervised discriminator (FastGAN) enhances the discriminator feature quality with self-monitoring characterization, and together with the lightweight generator module, it can still converge stably and retain structural details when less than 100 pieces of data are available [3]. The following table 2 compares the performance of different regularization and structural methods in small sample scenarios:

Table 2. Performance comparison between regularization and structural methods [2-4]

Method	Data set and setting	Improvement mode	FID	Application scenario analysis
ContraGAN	Tiny-ImageNet	Comparative Learning + 2C Loss	37.56(ProjGAN) → 32.72	enhanced classification, suitable for multiple types of small samples
ContraGAN	ImageNet-128	Comparative Learning + 2C Loss	21.07(BigGAN) → 19.44	large scale multi class generation task
FastGAN	AnimalFace-Cat	Self supervisionD + SLE module	71.76(BigGAN) → 35.11	very low samples (<100), suitable for small sample fields such as medicine/art
BigGAN + RLC	CIFAR-100 (20% data)	LeCam regularization	32.99 → 25.51	Multi category small sample scenarios, stable training dynamics
BigGAN + RLC	ImageNet (10% data)	data enhancement + LeCam regularization	37.71 → 24.38	complements data enhancement methods, significantly improving performance

LeCam regularization further significantly reduces FID on the basis of enhancement, and has obvious advantages in training stability. ContraGAN uses the contrast loss to achieve about 7 – 8% FID improvement on Tiny ImageNet. Although no specific value is given, the description is clear, and the discriminator overfitting is reduced. FastGAN can still maintain low FID and high stability under extremely low sample size (100 pieces), which is suitable for medical imaging and other sample scarce fields.

3.3 Cross domain adaptation/migration strategy

FreezeD fine-tuning by freezing the lower layer of the discriminator to retain common features, only fine-tuning the upper layer and the generator, it can inhibit overfitting and

accelerate convergence when small sample domains migrate. Small sample GAN adaptation starts from large-scale source domain pre training GAN, and uses similarity/smooth regularization to maintain one-to-one latent variable correspondence on the target small sample set, so as to improve the stability and diversity of style conversion [6]. The SSR proposed by ICCV 2023 is effective under very few samples. Reverse contrast/inter class constraints such as RCL use "reverse contrast learning" to improve the quality and diversity under a small number of samples [7]. The following table 3 compares the performance of different migration strategies:

Table 3. Performance comparison of cross domain adaptation methods [6-7]

Method	Transfer type	Sample size	FID	Applicable scenario
FreezeD vs. Fine-tuning	FFHQ → Animal Face	389	64.28 → 61.46	the domain gap is small, the structure is similar, and it is suitable for rapid fine-tuning
FreezeD vs. Fine-tuning	ImageNet → Oxford Flower	~100/class	27.05 → 24.80	condition generation task to maintain category consistency
SSR	Face → Anime	10-shot	97.3	domain gap is large, suitable for cross style migration
SSR	Church → Shells	10-shot	140.5	very low sample size, large structural differences
SSR	ImageNet → Flowers	10-shot	106.4	applies to unconditional adaptation of conditional GAN

FreezeD clearly quantifies that the migration efficiency is excellent and suitable for fast adaptation tasks. Although SSR constrains potential spatial smoothness through regularization, showing significant advantages under extremely low sample size and huge domain differences, providing a new technical path for small sample migration.

3.4 Re examination of evaluation indicators

FID is biased estimation and the deviation is related to the model, so it is easy to be distorted when encountering small samples; KID is more robust to sample size under the framework of nuclear MMD [9]. Recent research further points out the statistical uncertainty and estimation bias of FID/KID when the sample is insufficient. Precision recall for GANs can decompose fidelity (precision) and coverage (recall), and can also diagnose mode collapse; In recent years, it has been recommended as a supplement to FID in the overview/benchmark [4, 10]. Small sample evaluation should pay more attention to multi-dimensional indicators. If only a single FID is used to make conclusions, the risk is high. The introduction of KID/PR and visual review can significantly reduce the probability of miscarriage of justice.

4 Conclusion

This research systematically analyzes the core challenges faced by the generation of countermeasures network (GAN) under the condition of small samples in terms of image fidelity, including discriminator over fitting, pattern collapse, structural distortion and other problems, and classifies and quantifies the current mainstream solutions. The research results show that data enhancement methods (such as DiffAug and ADA) can effectively alleviate over fitting, especially in very low sample size scenarios, DiffAug shows higher fidelity;

Regularization and structural innovation methods (such as ContraGAN and FastGAN) have their own advantages in diversity and detail maintenance; Cross domain migration strategies (such as FreezeD and SSR) provide a stable and efficient way to generate domain adaptation tasks. The comparison of multiple sets of quantitative indicators shows that different methods are suitable for different data sizes and task objectives, and the combination strategy may be the key way to break through the bottleneck of small sample generation in the future.

The significance of this study is to provide a clear framework for challenge induction and method classification for small sample GAN training, and help researchers and engineering practitioners to select appropriate combination schemes under different task requirements. This framework can not only improve the generation quality under low sample conditions, but also provide technical reference for practical applications (such as medical image synthesis, remote sensing image enhancement and artistic style generation).

Future research should further explore three directions: first, develop more efficient data enhancement and regularization collaboration mechanisms to achieve automatic scheduling to avoid excessive detail smoothing; Second, build a general small sample migration paradigm to make cross domain adaptation more flexible, without a large number of super parameter adjustments; Third, improve the analysis method of multi-dimensional evaluation indicators to ensure more reliable comprehensive diagnosis of generation quality. Through these efforts, small sample GAN training is expected to achieve higher fidelity, stronger generalization ability, and play a greater role in various data scarcity scenarios.

References

1. S. Zhao, Z. Liu, J. Lin, J.-Y. Zhu, S. Han, Differentiable Augmentation for Data-Efficient GAN Training. arXiv preprint arXiv:2006.10738 (2020). https://doi.org/10.48550/arXiv.2006.10738
2. M. Kang, J. Park, ContraGAN: Contrastive Learning for Conditional Image Generation. arXiv preprint arXiv:2006.12681 (2020). https://doi.org/10.48550/arXiv.2006.12681
3. B. Liu, Y. Zhu, K. Song, A. Elgammal, Towards Faster and Stabilized GAN Training for High-Fidelity Few-Shot Image Synthesis. arXiv preprint arXiv:2101.04775 (2021). https://doi.org/10.48550/arXiv.2101.04775
4. H.-Y. Tseng, L. Jiang, C. Liu, M.-H. Yang, W. Yang, Regularizing Generative Adversarial Networks under Limited Data. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) 2021, pp. 7917-7927 (2021). https://doi.org/10.1109/CVPR46437.2021.00783
5. T. Shaham, T. Dekel, T. Michaeli, SinGAN: Learning a Generative Model from a Single Natural Image. ICCV (2019): 4568–4577. https://doi.org/10.1109/ICCV.2019.00479
6. Mo, Freeze discriminator: A simple baseline for finetuning GANs. arXiv preprint arXiv:2002.10964 (2020). https://doi.org/10.48550/arXiv.2002.10964
7. V. Sushko, R. Wang, J. Gall, Smoothness Similarity Regularization for Few-Shot GAN Adaptation. Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV) 2023 (2023). https://doi.org/10.48550/arXiv.2308.09717
8. T. Karras, M. Aittala, J. Hellsten et al., Training Generative Adversarial Networks with Limited Data. NeurIPS (2020). https://doi.org/10.48550/arXiv.2006.06676
9. Y. Yang, W. Wang, Image Synthesis under Limited Data: A Survey and Taxonomy. arXiv preprint arXiv:2303.13202 (2023). https://doi.org/10.48550/arXiv.2303.13202
10. M. S. M. Sajjadi, O. Bachem, M. Lucic, O. Bousquet, S. Gelly, Precision-Recall Curves for Generative Models. arXiv preprint arXiv:1806.00035 (2018). https://doi.org/10.48550/arXiv.1806.00035