Generative Model Perception Rectification Algorithm for Trade-Off between Diversity and Quality

Guipeng Lan*, Shuai Xiao†*, Jiachen Yang, Jiabao Wen

School of Electrical and Information Engineering, Tianjin University, Tianjin China
{lgp, xs611, yangjiachen, Wen_Jiabao}@tju.edu.cn

Abstract

How to balance the diversity and quality of results from generative models through perception rectification poses a significant challenge. Abnormal perception in generative models is typically caused by two factors: inadequate model structure and imbalanced data distribution. In response to this issue, we propose the dynamic model perception rectification algorithm (DMPRA) for generalized generative models. The core idea is to gain a comprehensive perception of the data by appropriately highlighting the low-density samples in the perception space, also known as the minor group samples. The entire process can be summarized as "search-evaluation-adjustment". To identify low-density regions in the data manifold within the perception space of generative models, we introduce a filtering method based on extended neighborhood sampling. Based on the informational value of samples from low-density regions, our proposed mechanism generates informative weights to assess the significance of these samples in correcting the models' perception. By using dynamic adjustment, DMPRA ensures simultaneous enhancement of diversity and quality in the presence of imbalanced data distribution. Experimental results indicate that the algorithm has effectively improved Generative Adversarial Nets (GANs), Normalizing Flows (Flows), Variational AutoEncoders (VAEs), and Diffusion Models (Diffusion).

Introduction

The latest advancements in generative models have led to their widespread use in image generation (Karras et al. 2020b), image editing (Wang et al. 2008), and other related tasks (Patashnik et al. 2021). These models encompass various techniques, such as Generative Adversarial Networks (GANs) (Goodfellow et al. 2014; Karras, Laine, and Aila 2019), Flows (Dinh, Krueger, and Bengio 2014; Kingma and Dhariwal 2018), Variational Autoencoders (VAEs) (Dai and Wipf 2019; Razavi, Van den Oord, and Vinyals 2019), and Diffusions (Ho, Jain, and Abbeel 2020; Nichol and Dhariwal 2021). A common characteristic of these generative models is the superior limit of diversity depends on the models' perception of data.

Data is a crucial component for generative models to be involved in training, inference, and evaluation processes. Generally, the obtained data is unbalanced. The presence of the minor group samples is a common issue in unbalanced data. The minor group samples are located in low-density areas in the data distribution. The aforementioned issue causes the model to exhibit "fuzziness" while recognizing low-density regions, resulting in abnormal perception and a decrease in the quality of the generated outcomes. To ensure the quality of generation, the prior researches can be categorized into two approaches: preprocessing-based methods (Turner et al. 2019; DeVries, Drozdzal, and Taylor 2020) and post-processing-based methods (Azadi et al. 2018; Grover et al. 2019). Preprocessing-based methods involve evaluating the dataset prior to network training and eliminating abnormal samples to maintain data balance. Post-processing methods entail evaluating the generated data from well-trained models and removing low-quality portions. While these two approaches can enhance the quality of the generated data, they inadvertently decrease diversity due to incomplete data perception. Hence, determining the balance between diversity and generation quality has emerged as significant research areas.

This paper proposes a method called DMPRA. The aim of DMPRA is to identify and explore low-density regions within the perception space of generative models, and assign informative weights to improve the models' perception of data through multiple training iterations. DMPRA is guided by the principle that samples from the minor groups possess higher information value and, thus, require greater attention. To achieve this, we employ extended neighborhood sampling to measure feature density and locate samples within low-density regions. This sampling technique helps mitigate the problem of density confusion in local areas, which arises due to the aggregation of some samples that deviate from the main distribution. Additionally, we propose a mechanism to generate informative weights that assess the fine-grained attention required for samples in low-density regions, further enhancing the generative model's perception. Finally, DMPRA dynamically adjusts the sampling weight through multiple training iterations in order to strike a balance between the diversity and the quality of the generated results.

Our contributions We summarize the contributions of this work. **Algorithm** We firstly propose a general dynamic

*These authors contributed equally.
†Corresponding Author

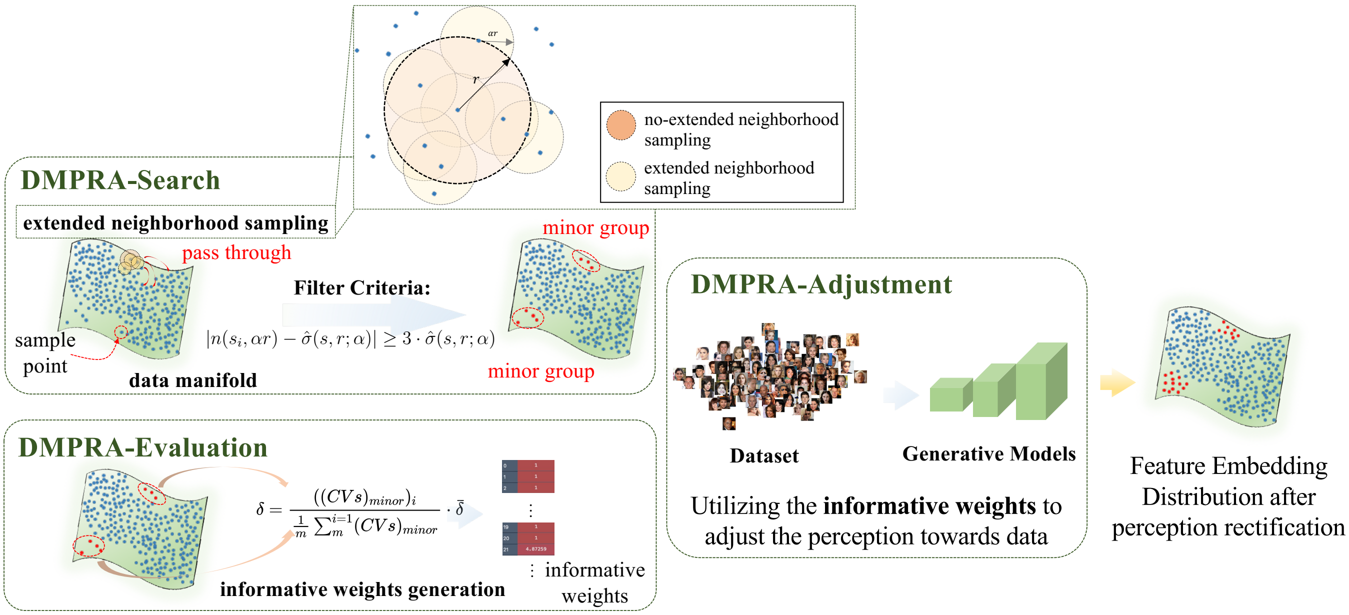

Figure 1: The schematic diagram of the DMPRA for generative models.

model perception rectification algorithm for generative models, which solves the contradiction between the diversity and the quality caused by the minor group samples. Secondly, a filtrating method for searching samples in low-density regions is proposed. We utilize extended neighborhood sampling to avoid the problem of density confusion in local areas caused by the aggregation of samples that deviate from the main distribution.This paper provides a mechanism for generating informative weights to further promote the perception ability of the generative models. **Applications** We have verified the potential of DMPRA for wide applications, such as alleviating mode collapse, completeness data augmentation and improving the data-driven downstream tasks.

Related Literatures

Model Perception Rectification Algorithms Based on Data-representation

The model perception rectification algorithms based on data-representation are to improve the overall performance by changing, deleting, re-sampling according to the feature representation or feature statistics of data. Mo et al. (Mo et al. 2019) proposed the gap of log-densities to measure the difference between the data distribution and the model distribution of the given samples, which provides an effective performance improvement for cGANs. DeVries et al. (DeVries, Drozdzal, and Taylor 2020) changed the training datasets through instance-selection before model training, and removed the samples outside the data manifold, thus improving the fidelity of the results and significantly reducing the training time. Turner et al. (Turner et al. 2019) proposed Metropolis-Hastings GAN, which applies Markov chain Monte Carlo to filter high-dimensional data to improve the effect of multiple benchmark datasets. Wu et al. (Wu et al. 2020) proposed a sample re-weighting mechanism to improve the performance of GANs by downplaying bad-quality samples.

Model Perception Rectification Algorithms Based on Model-representation

The model perception rectification algorithms based on model-representation improves the performance of the model through loss functions, optimization strategies, etc. Azadi et al. (Azadi et al. 2018) proposed the Reject Sampling Scheme (DRS), which utilizes the discriminator of GANs to correct the error in the generator distribution. DRS improves the quality of generated samples by filtering bad- quality generation results. Ding et al. (Ding, Wang, and Welch 2020) designed a novel Softplus loss for discriminator-based density ratio estimation. The algorithm does not depend on the optimization of discriminator and is applicable to multiple GANs. Grover et al. (Grover et al. 2019) used the likelihood-free importance weighting method to correct for the bias in generative models. Wu et al. (Wu et al. 2019) proposed LoGAN, which applied gradient-based optimization to improve the interaction between generator and discriminator.

Compared to the aforementioned works, DMPRA possesses three **highlights**: Firstly, it is not only applicable to GANs but also to other generative models. Secondly, DMPRA ensures both the quality and the diversity of the generated results. Lastly, our algorithm firstly addresses the issue of minor group samples through the lens of informativeness.

DMPRA

The schematic diagram of DMPRA for generative models is depicted in Fig.1.DMPRA searches for samples that are disregarded by the perception of generative models (high information samples), that is, low-density regions in the data manifold. To achieve this, DMPRA assigns informative weights to these samples, thereby encouraging generative models to pay more precise attention to their features during data perception. DMPRA represents a dynamic adjustment algorithm that enhances the data perception ability of generative models through continuous "**search-evaluation-adjustment**" (particularly reflected in the improved diversity and generation quality).

DMPRA-Search

Before searching for samples in low-density regions, it is important to obtain the feature distribution that reflects generative models' perception of the dataset. In GANs, the discriminator, which is fed both real and fake samples during joint training with the generator, is a suitable choice for extracting perception features. However, other generative models like VAEs, Flows, and Diffusions lack feature extraction networks similar to discriminators. To address this, we adopt a classification network (like discriminators) as perception feature extractors and train them using real and generated data from the generative models. Based on the extracted features, we construct the perception space of the generative models. In this paper, we have selected ResNet152 (He et al. 2016) as the perception feature extractor for the generative models. Real data and data generated from the generative models are used as inputs for the classification network. The involvement of the generated data in training the perception feature extractor, similar to the discriminators in GANs. The well-trained ResNet152 takes into account the cognition of the generative models during feature extraction. For training, ResNet152 utilizes the cross-entropy loss function. The samples' features extracted by loading the pre-trained model discriminator and the well-trained ResNet152 can reflect the models' perceptions of the dataset, i.e., constructing the perception spaces for generative models. Once the perception space is obtained, we can search for high information samples using extended neighborhood sampling, as guided by the following lemma.

lemma 1: The information value of the sample is related to its density in the distribution space (Mo et al. 2019; DeVries, Drozdzal, and Taylor 2020; Azadi et al. 2018).

In general, the higher the density in the distribution space, the greater the similarity between the features of the samples and those of the other samples in the dataset, and the lower the information value of each sample. On the contrary, samples with features in the low-density area of the distribution space are relatively rare in the overall representation of the dataset, and therefore have high information value for generative models. This paper proposes a sample information calculation method based on **Lemma 1** to quantitatively measure the information value of each sample.

Inspired by Papadimitriou et al (Papadimitriou et al. 2003), in order to avoid the problem of density confusion in local areas caused by the aggregation of samples that deviate from the main distribution, we introduce the extended neighborhood sampling. We use density-based representations of sample points in extended neighborhoods to search minor group samples within the data manifold.

Preliminaries: We describe the variables defined in this paper. The number of samples contained in the data manifold is n. $d(s_i, s_j)$ represents the feature distance between sample s_i and sample s_j. We use the Euclidean distance to measure the feature distance between samples. In addition, we also define a hyper-sphere $S(s, r)$, which represents a high-dimensional hyper-sphere with s as the center and r as the radius. Define $N(s, r)$ as the set of sample points in $S(s, r)$. Assuming that the known sample distribution density function is $f(e)$, the number of samples in $S(s, r)$ can be calculated by Eq.1.

$$n(s,r) = n \cdot \int_{S(s,r)} f(e)\mathrm{d}e, \tag{1}$$

Extended neighborhood sampling considers the mean and standard deviation of each point's performance within the set $N(s, r)$ to avoid density confusion in local areas caused by the aggregation of outliers that deviate from the main distribution. $\hat{n}(s, r; \alpha)$ is defined as the average number of samples in the αr sub hyper-sphere $S(s_i, \alpha r)$ of all sample points in $S(s, r)$. The calculation process is shown in Eq.2.

$$\hat{n}(s,r;\alpha) = \frac{\int_{S(s,r)} (n \cdot \int_{S(s_i,\alpha r)} f(e)\mathrm{d}e)\mathrm{d}e}{n(s,r)}, \tag{2}$$

$\hat{\sigma}(s, r; \alpha)$ is defined as the standard deviation of the number of samples in the αr sub hyper-sphere $S(s_i, \alpha r)$ of all sample points in $S(s, r)$. The calculation process is shown in Eq.3.

$$\hat{\sigma}(s,r;\alpha) = \sqrt{\frac{\int_{S(s,r)} (n(s_i,\alpha r) - \hat{n}(s,r;\alpha))^2}{n(s,r)}}, \tag{3}$$

where $n(s_i, \alpha r)$ is the number of samples in the sub hyper-sphere $S(s_i, \alpha r)$.

lemma 2: Chebyshev's theorem holds that in an arbitrary dataset, at least 88.9% of the data are within the range of three standard deviations of the average.

According to Chebyshev inequality:

$$\begin{aligned} P(|n(s_i,\alpha r) - \hat{\sigma}(s,r;\alpha)| \geq \varepsilon\hat{\sigma}(s,r;\alpha)) \\ \leq \frac{(D(n(s_i,\alpha r)))^2}{(\varepsilon \cdot \hat{\sigma}(s,r;\alpha))^2} = \frac{1}{\varepsilon^2}, \end{aligned} \tag{4}$$

when ε=3, 88.9% of the data meet the inequality $|n(s_i, \alpha r) - \hat{\sigma}(s, r; \alpha)| < 3 \cdot \hat{\sigma}(s, r; \alpha)$, while the remaining data are suspiciously different from other samples. This part of the sample as the minor group samples (Hawkins 1980), that is, samples with high information value. Papadimitriou et al. (Papadimitriou et al. 2003) pointed out that for the actual data distribution, the boundary is relatively loose. When ε=3, there should be less than 1% of the samples deviating. In this way, the threshold is more consistent with the understanding of the minor group samples.

In most cases, $f(e)$ does not simply obey standardized distribution, such as normal distribution, Gaussian distribution, etc. For the sample distribution under discrete conditions, Eq.2 and Eq.3 can be converted in discrete space, as shown in Eq.5.

$$\hat{n}(s,r;\alpha) = \frac{\sum_{s_i \in N(s,r)} n(s_i, \alpha r)}{n(s,r)},$$
$$\hat{\sigma}(s,r;\alpha) = \sqrt{\frac{\sum_{s_i \in N(s,r)} (n(s_i,\alpha r) - \hat{n}(s,r;\alpha))^2}{n(s,r)}}, \tag{5}$$

DMPRA-Evaluation

To improve the perception of the minor group samples, we propose an unbalanced sampling strategy. The key aspect of the unbalanced sampling strategy is determining the sampling weight of each sample. The sampling weight generation mechanism proposed in this paper is based on the discrete degree of sample information in the neighborhood.

In order to objectively measure the discrete degree of sample information, we introduce the coefficient of variation, as shown in Eq.6.

$$CVs = \hat{\sigma}(s,r;\alpha)/\hat{n}(s,r;\alpha), \tag{6}$$

According to the proposed method for evaluating sample information in section 3.1, we select the minor group samples and quantitatively determine the difference ratio of the discrete degree of sample information between the minor group samples and the main group samples, as shown in Eq.7.

$$\bar{\delta} = \frac{\frac{1}{m}\sum_{m}^{i=1}(CVs)_{minor}}{\frac{1}{n-m}\sum_{n-m}^{j=1}(CVs)_{main}}, \tag{7}$$

where m is the number of samples from the minor group samples, $(CVs)_{minor}$ is the CVs of the samples in the minor group samples, $(CVs)_{main}$ is the CVs of the samples in the main group samples.

In addition, due to the relatively rich information representation of the main group samples and in order to facilitate calculation, the sampling weights of the main group samples are set to 1.

DMPRA-Adjustment

We propose a training method called multiple training and dynamic adjustment. This method aims to enhance the models' perception of the entire dataset, leading to improved quality and diversity in the generated results. We illustrate the training method, which includes multiple training and dynamic adjustment, in Algorithm 1.

Algorithm 1: Dynamic model perception rectification algorithm

Input: Dataset that satisfies the Generative Models
Parameter: Initialize the generative models G

```
while t ← 1 to T do
    if e = 0 then
        Training without informativeness weights;
    else
        Loading the pre-training model;
        Constructing cognitive space of generative models;
        Obtaining n̂(s, r; α) and σ̂(s, r; α);
        if |n(s_i, αr) − σ̂(s, r; α)| > 3 · σ̂(s, r; α) then
            Confirm that this sample is the minor group sample with high information value;
        else
            pass;
            Generating informative weights though δ = ((CVs)_minor)_i / ((1/m) Σ_m^{i=1} (CVs)_minor) · δ̄  Utilizing the informative weights to adjust the perception towards data
        end if
    end if
end while
```

Experiments & Applications

Datasets & Quantitative Evaluation Metrics

Datasets: CelebA-HQ (Karras et al. 2017): *i.e.*, CelebFaces Attributes High Quality, is extensively used in computer vision tasks related to faces. CelebA-HQ dataset collectes 30K images with a resolution of 1024×1024. CelebA-HQ is obtained on the basis of CelebA. An autoencoder is used to remove JPEG artifacts in CelebA and upscale the images through the super-resolution network. In addition, CelebA-HQ crops the image through facial landmarks, so that each image has a normalized position and rotation.

Metrics 1: FID (Heusel et al. 2017): *i.e.*, Fréchet Inception Distance, which is applied to measure the diversity and quality of generation images. FID makes use of the mean and covariance matrix to calculate the high-dimensional distribution distance between the generation data and the original data. The smaller the FID, the similar the two distributions, indicating the better the diversity and quality of the generation images.

Metrics 2: PRDC (Naeem et al. 2020): is a quantitative metric used to measure the density and coverage of generation samples. It can measure the quality of the trained generative models from another perspective.

Metrics 3: BRISQUE (Mittal, Moorthy, and Bovik 2011): is a blind image spatial quality method based on natural scene statistic. This method measures image naturaless based on measured deviations from a natural image model through extracting the points wise statistics of local normalized luminance signals. BRISQUE performs better than the full reference structural index SSIM, which is more in line with human perception. We use BRISQUE as an evaluation metric to evaluate the naturalness of generative images.

Results and Analysis

To verify the effectiveness of the DMPRA, we select LogcoshVAE (Chen, Chen, and Zhang 2018), StyleGAN2-ada (Karras et al. 2020a), Generative Flow (Kingma and Dhariwal 2018), and Denosing diffusion models (Nichol and Dhariwal 2021) as the baseline models. We conduct five dynamic adjustments for each baseline. The results are shown in TABLE 1.

According to TABLE 1, we can obtain the following characteristics of DMPRA: (1) From the perspectives of FID, we

	LogcoshVAE				StyleGAN2-ada			
	FID	PRDC:D	PRDC:C	BRISQUE	FID	PRDC: D	PRDC: C	BRISQUE
iter 1	131.72±0.84	0.16±0.0163	0.03±0.0018	58.02±0.23	38.42±0.32	1.08±0.0214	0.77±0.0100	13.25±0.39
iter 2	106.69±0.82	0.31±0.0155	0.05±0.0009	53.73±0.27	29.18±0.31	1.32±0.0133	0.90±0.0053	11.92±0.31
iter 3	104.60±0.13	0.33±0.0098	0.06±0.0011	50.54±0.23	26.51±0.29	1.33±0.0118	0.92±0.0075	10.45±0.29
iter 4	100.24±0.25	0.33±0.0041	0.07±0.0030	49.91±0.09	23.50±0.15	1.31±0.0053	0.92±0.0075	10.32±0.21
iter 5	97.66±0.06	0.34±0.0032	0.08±0.0014	48.91±0.04	23.08±0.09	1.29±0.0025	0.94±0.0015	10.06±0.19
	Generative Flow				Denosing Diffusion Model			
	FID	PRDC:D	PRDC:C	BRISQUE	FID	PRDC: D	PRDC: C	BRISQUE
iter 1	90.22±0.83	0.76±0.0321	0.35±0.0025	18.96±0.67	98.25±0.68	0.21±0.0121	0.03±0.0023	40.39±0.54
iter 2	89.05±0.81	1.05±0.0354	0.37±0.0037	16.88±0.45	93.48±0.32	0.21±0.0045	0.04±0.0015	40.00±0.47
iter 3	88.56±0.27	1.06±0.0279	0.40±0.0021	15.97±0.39	91.99±0.34	0.22±0.0034	0.05±0.0022	39.87±0.36
iter 4	87.97±0.45	1.18±0.0249	0.41±0.0017	15.87±0.40	86.72±0.25	0.23±0.0037	0.05±0.0018	39.38±0.33
iter 5	87.13±0.19	1.28±0.0128	0.42±0.0013	14.38±0.21	85.71±0.22	0.22±0.0019	0.05±0.0009	38.99±0.32

Table 1: The results of the DMPRA.

can confirm that the generative models trained with DMPRA improve diversity and quality. (2) From the perspectives of PRDC: D and PRDC: C, as the dynamic adjustments, the feature density and coverage of the results from the generative models gradually increase, indicating that DMPRA can assist the generative models in improving their perception of data. (3) From the perspective of BRISQUE, the naturalness of the results is gradually improving, which is a reflection of the models' enhanced perception of the data.

Why do we propose the dynamic model perception rectification algorithm? To explain the contribution of DMPRA to data perception for generative models, we conduct the following experiment. We choose LogcoshVAE, StyleGAN2-ada, Generative Flow, and Denosing diffusion models as the baseline models. We compare the models with DMPRA and without DMPRA, and perform 4 adjustments, denoted as w-DMPRA and w/o-DMPRA. The quantitative metrics FID, PRDC: D, PRDC: C, and BRISQUE of the results are visualized, as shown in Fig.2.

According to Fig.2, we observe that the generated images obtained through DMPRA have significant improvements in diversity and quality. The models with DMPRA also perform better in terms of density and coverage, indicating that DMPRA can improve the models' perception of the data.

In additional, we find a phenomenon: DMPRA gains different benefits towards the different generative models. Regarding this phenomenon, our considerations are as follows: The performance improvement of generative models can be mainly divided into two aspects: data-representation and model-representation. The role of DMPRA is to assist generative models in paying more comprehensive attention to datasets from the perspective of data. However, it cannot be ignored that the improvement of the models greatly promotes the improvement of generation efficiency. Diffusion models and generative flow are the outstanding representations of generative models, they have better network structures and well-designed loss functions, which make the two models inherently have strong data perception capabilities. Therefore, compared to other generative models, the advantages of diffusion models and generative flow brought by DMPRA are not very prominent after the iteration 4. DMPRA assists generative models in comprehensively understanding data at the data level. Although the benefits of DMPRA for diffusion models and generative flow are limited after the iteration 4, from a quantitative perspective, their performances in various metrics (FID, PRDC:D, PRDC:C, and BRISQUE) are superior to that of w/o-DMPRA.

How to further demonstrate the effectiveness of DMPRA in GANs? The algorithms for improving GANs have achieved significant achievements (in section 2). To further demonstrate the effectiveness of DMPRA in GANs, we conduct the following comparative experiments. We select DRS (Azadi et al. 2018), Topk (Sinha et al. 2020), Dia-GAN (Lee et al. 2021), AdvLatGAN (qua+) (Li et al. 2022) and LatentAugment (Tronchin et al. 2023) as the baselines. Due to the fact that DRS improves generation quality by intervening in the inference process, we include Topk+DRS, Dia-GAN+DRS, AdvLatGAN+DRS and LatentAugment+DRS as baselines. The experimental results are shown in TABLE 2. We compare FID, PRDC: D, PRDC:C and BRISQUE over GAN on CelebA-HQ (5000 images). DMPRA-DRS achieves the best FID, PRDC: D and PRDC:C. This result further demonstrates the effectiveness of DMPRA in GAN.

Why do we require informative weights? To verify the effectiveness of the informative weights in the DMPRA, we conduct the following comparative experiment. We select LogcoshVAE, LogcoshVAE with DMPRA, and LogcoshVAE with DMPRA but without informative weights as the comparative models, denoted as LogcshVAE w-DMPRA, LogcshVAE w/o-DMPRA, and LogcshVAE w/o-informational_weights, respectively. Among them, the model without informative weights sets the sampling frequency of the minor group samples to $\overline{\delta}$, and the calculation

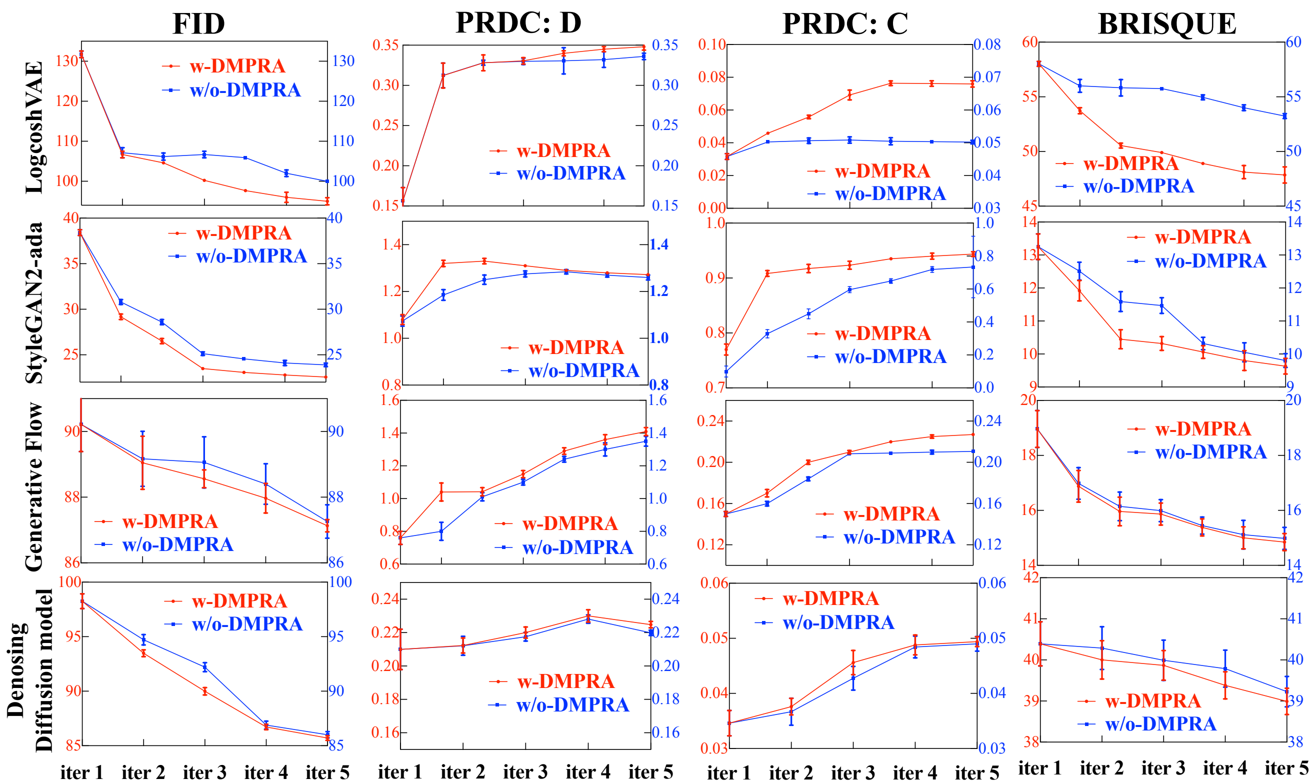

Figure 2: Comparison results between models with DMPRA and without DMPRA.

Algorithm	FID	PRDC: D	PRDC: C	BRISQUE
Vanilla	108.99	0.1472	0.1234	30.46
DRS	92.16	0.1926	0.2224	26.02
Topk	104.48	0.1824	0.1868	28.66
Tok+DRS	91.56	0.2088	0.2892	24.94
Dia-GAN	72.31	0.2305	0.2918	20.14
Dia-GAN+DRS	68.91	0.2607	0.3107	19.62
AdvLatGAN	72.54	0.2298	0.2903	21.16
AdvLatGAN+DRS	66.21	0.2816	0.3170	**19.58**
LatentAugment	75.31	0.2242	0.2875	21.68
LatentAugment+DRS	69.77	0.2694	0.3069	19.93
DMPRA	71.56	0.2552	0.3072	19.85
DMPRA+DRS	**65.83**	**0.3169**	**0.3356**	19.81

Table 2: The further comparative experiments in GANs

formula for $\bar{\delta}$ is as Eq.7. According to the above experimental mode, we conduct four dynamic adjustments and quantitatively measure the diversity and quality of the images generated by the three comparative models, as shown in Fig.3. The red ones represent the feature embeddings of the original data, while the blue ones represent the feature embeddings of the generated data.

From Fig.3, we can infer the informative weights quantify the information of the minor group samples with finer granularity, which is beneficial for the models to rectify the perception of data.

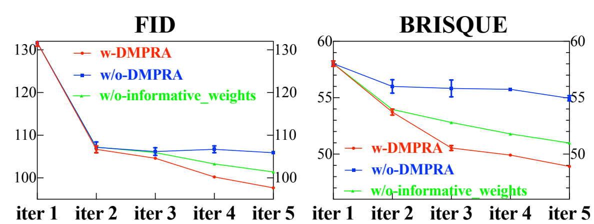

Figure 3: The effectiveness of the informative weights.

Visualizing

In order to demonstrate more clearly that our algorithm can improve the model's perception of data, we visualize the feature embeddings between the generated images and the original images. We choose StyleGAN2-ada as the baseline model. We train StyleGAN2-ada with and without DMPRA, denoted as w-DMPRA and w/o-DMPRA. We extract feature embeddings from 1000 images after each dynamic adjustment and the original dataset, and visualize the extracted feature embeddings with t-SNE (Van der Maaten and Hinton 2008), as shown in Fig.4.

From Fig.4, we find that the StyleGAN2-ada w-DMPRA gradually tends to achieve uniform distribution of the generated images. We use orange to circle the areas with high

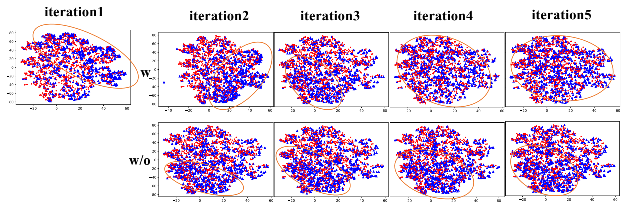

Figure 4: Visualizing the process of DMPRA for improving the StyleGAN2-ada's perception.

Figure 5: Examples of alleviating mode collapse. The left side shows the results obtained from the model without DMPRA, while the right side shows the results with DMPRA.

feature density in the generated samples. When the iteration reaches to 5, the generated images have a feature distribution that is basically similar to that of the dataset. While the StyleGAN2-ada without model perception rectification algorithm cannot achieve uniform distribution of features.

Applications

Alleviating mode collapse We find that the DMPRA proposed in this paper can to some extent alleviate the very challenging problem in generative models, i.e., mode collapse. To provide a more intuitive explanation of the effectiveness of our algorithm, we conduct a simple but interesting experiment. We treat StyleGAN2-ada as the baseline model and select 1000 images from CelebA-HQ and 50 images from AFHQ (Choi et al. 2020) as dataset. It is evident that the 50 images from AFHQ belongs to the minor group samples here. We directly train StyleGAN2-ada and use DMPRA to train StyleGAN2-ada. We filter out the images that belong to the features in AFHQ. The results are shown in Fig.5.

From the Fig.5, we can discover that in the images obtained without DMPRA, although the quality is high, the modes of the features are single and do not have good diversity. In the images with DMPRA, the generated images exhibit more modes in features such as posture, fur color.

Enhancing the ability of the data-driven downstream tasks Data augmentation is a challenging task in computer vision (Van Dyk and Meng 2001; Shorten and Khoshgoftaar 2019). It plays an important role in sensitive data substitution and downstream task data requirements (Wong et al. 2016). Utilizing DMPRA can further enhance the model's perception of the minor group samples in the dataset, achieving the augmentation of completeness data to enhance the ability of the data-driven downstream tasks. To demonstrate the ability of DMPRA to enhance downstream tasks, we conduct some practical validation.

data	accuracy
original data	81.33%
+StyleGAN2-ada	88.00%
+StyleGAN2-ada w-DMPRA	**89.33%**

Table 3: The classification results

Image classification: We choose chest X-Ray Images (Kermany et al. 2018) as the dataset, ResNet152 (He et al. 2016) as the baseline for classification, and StyleGAN2-ada as the baseline for data augmentation to achieve pneumonia screening on X-Ray images and assist doctors in making diagnoses in the background of big data. Train the classification network and complete testing on the original data, augmentation data from StyleGAN2-ada, and augmentation data from StyleGAN2-ada with DMPRA (denoted as original data,+StyleGAN2-ada, and +StyleGAN2-ada w-DMPRA). The results are shown in TABLE 3.

Object detection: The object detection algorithm based on deep learning is a data-driven task. For this study, we selected the YOLO v5 as the baseline. To assess the advantages of DMPRA in data-driven downstream tasks, we compared the performance using three distinct datasets: (1) 1000 facial images from Celeba-HQ; (2) 1000 facial images from Celeba-HQ + 1000 images from StyleGAN2-ada-w/o DMPRA; (3) 1000 facial images from Celeba-HQ + 1000 images from StyleGAN2-ada-w DMPRA. To annotate these datasets, we utilized LabelImg as an annotation tool. Subsequently, YOLO v5 was trained separately using the same hyperparameters (including training epochs, learning rate, etc.). We used a test set of 50 images to calculate the accuracy scores, which yielded 95.3%, 97.3%, and 98.0% for the three aforementioned datasets, respectively.

Conclusion

We introduce DMPRA, a dynamic model perception rectification algorithm for trade-off between diversity and quality. The entire process of DMPRA can be summarized as: "search-evaluation-adjustment". DMPRA-search exploits extended neighborhood sampling to filter out the minor group samples, which have high information. Extended neighborhood sampling can avoid the problem of density confusion in local areas caused by the aggregation of samples that deviate from the main distribution. DMPRA-evaluation put forward a mechanism for generating informative weights to provide fine-grained attention, which further promote the performance of the generative models. In DMPRA-adjustment, we have introduced dynamic adjustment to continuous rectify the generative models' perception. In addition, DMPRA has a wide range of applications, such as alleviating mode collapse, completeness data augmentation, enhancing the ability of the data-driven downstream tasks, etc.

References

Azadi, S.; Olsson, C.; Darrell, T.; Goodfellow, I.; and Odena, A. 2018. Discriminator rejection sampling. *arXiv preprint arXiv:1810.06758*.

Chen, P.; Chen, G.; and Zhang, S. 2018. Log hyperbolic cosine loss improves variational auto-encoder.

Choi, Y.; Uh, Y.; Yoo, J.; and Ha, J.-W. 2020. StarGAN v2: Diverse Image Synthesis for Multiple Domains. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Dai, B.; and Wipf, D. 2019. Diagnosing and enhancing VAE models. *arXiv preprint arXiv:1903.05789*.

DeVries, T.; Drozdzal, M.; and Taylor, G. W. 2020. Instance selection for gans. *Advances in Neural Information Processing Systems*, 33: 13285–13296.

Ding, X.; Wang, Z. J.; and Welch, W. J. 2020. Subsampling generative adversarial networks: Density ratio estimation in feature space with softplus loss. *IEEE Transactions on Signal Processing*, 68: 1910–1922.

Dinh, L.; Krueger, D.; and Bengio, Y. 2014. Nice: Non-linear independent components estimation. *arXiv preprint arXiv:1410.8516*.

Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative adversarial nets. *Advances in neural information processing systems*, 27.

Grover, A.; Song, J.; Kapoor, A.; Tran, K.; Agarwal, A.; Horvitz, E. J.; and Ermon, S. 2019. Bias correction of learned generative models using likelihood-free importance weighting. *Advances in neural information processing systems*, 32.

Hawkins, D. M. 1980. *Identification of outliers*, volume 11. Springer.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Heusel, M.; Ramsauer, H.; Unterthiner, T.; Nessler, B.; and Hochreiter, S. 2017. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30.

Ho, J.; Jain, A.; and Abbeel, P. 2020. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33: 6840–6851.

Karras, T.; Aila, T.; Laine, S.; and Lehtinen, J. 2017. Progressive growing of gans for improved quality, stability, and variation. *arXiv preprint arXiv:1710.10196*.

Karras, T.; Aittala, M.; Hellsten, J.; Laine, S.; Lehtinen, J.; and Aila, T. 2020a. Training Generative Adversarial Networks with Limited Data. In *Proc. NeurIPS*.

Karras, T.; Laine, S.; and Aila, T. 2019. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 4401–4410.

Karras, T.; Laine, S.; Aittala, M.; Hellsten, J.; Lehtinen, J.; and Aila, T. 2020b. Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 8110–8119.

Kermany, D. S.; Goldbaum, M.; Cai, W.; Valentim, C. C.; Liang, H.; Baxter, S. L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. 2018. Identifying medical diagnoses and treatable diseases by image-based deep learning. *cell*, 172(5): 1122–1131.

Kingma, D. P.; and Dhariwal, P. 2018. Glow: Generative flow with invertible 1x1 convolutions. *Advances in neural information processing systems*, 31.

Lee, J.; Kim, H.; Hong, Y.; and Chung, H. W. 2021. Self-diagnosing gan: Diagnosing underrepresented samples in generative adversarial networks. *Advances in Neural Information Processing Systems*, 34: 1925–1938.

Li, Y.; Mo, Y.; Shi, L.; and Yan, J. 2022. Improving generative adversarial networks via adversarial learning in latent space. *Advances in Neural Information Processing Systems*, 35: 8868–8881.

Mittal, A.; Moorthy, A. K.; and Bovik, A. C. 2011. Blind/referenceless image spatial quality evaluator. In *2011 conference record of the forty fifth asilomar conference on signals, systems and computers (ASILOMAR)*, 723–727. IEEE.

Mo, S.; Kim, C.; Kim, S.; Cho, M.; and Shin, J. 2019. Mining GOLD samples for conditional GANs. *Advances in Neural Information Processing Systems*, 32.

Naeem, M. F.; Oh, S. J.; Uh, Y.; Choi, Y.; and Yoo, J. 2020. Reliable fidelity and diversity metrics for generative models. In *International Conference on Machine Learning*, 7176–7185. PMLR.

Nichol, A. Q.; and Dhariwal, P. 2021. Improved denoising diffusion probabilistic models. In *International Conference on Machine Learning*, 8162–8171. PMLR.

Papadimitriou, S.; Kitagawa, H.; Gibbons, P. B.; and Faloutsos, C. 2003. Loci: Fast outlier detection using the local correlation integral. In *Proceedings 19th international conference on data engineering (Cat. No. 03CH37405)*, 315–326. IEEE.

Patashnik, O.; Wu, Z.; Shechtman, E.; Cohen-Or, D.; and Lischinski, D. 2021. Styleclip: Text-driven manipulation of stylegan imagery. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2085–2094.

Razavi, A.; Van den Oord, A.; and Vinyals, O. 2019. Generating diverse high-fidelity images with vq-vae-2. *Advances in neural information processing systems*, 32.

Shorten, C.; and Khoshgoftaar, T. M. 2019. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1): 1–48.

Sinha, S.; Zhao, Z.; ALIAS PARTH GOYAL, A. G.; Raffel, C. A.; and Odena, A. 2020. Top-k training of gans: Improving gan performance by throwing away bad samples. *Advances in Neural Information Processing Systems*, 33: 14638–14649.

Tronchin, L.; Vu, M. H.; Soda, P.; and Löfstedt, T. 2023. LatentAugment: Data Augmentation via Guided Manipulation of GAN's Latent Space. *arXiv preprint arXiv:2307.11375*.

Turner, R.; Hung, J.; Frank, E.; Saatchi, Y.; and Yosinski, J. 2019. Metropolis-hastings generative adversarial networks. In *International Conference on Machine Learning*, 6345–6353. PMLR.

Van der Maaten, L.; and Hinton, G. 2008. Visualizing data using t-SNE. *Journal of machine learning research*, 9(11).

Van Dyk, D. A.; and Meng, X.-L. 2001. The art of data augmentation. *Journal of Computational and Graphical Statistics*, 10(1): 1–50.

Wang, H.-X.; Pan, C.; Gong, H.; and Wu, H.-Y. 2008. Facial image composition based on active appearance model. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, 893–896. IEEE.

Wong, S. C.; Gatt, A.; Stamatescu, V.; and McDonnell, M. D. 2016. Understanding data augmentation for classification: when to warp? In *2016 international conference on digital image computing: techniques and applications (DICTA)*, 1–6. IEEE.

Wu, Y.; Donahue, J.; Balduzzi, D.; Simonyan, K.; and Lillicrap, T. 2019. Logan: Latent optimisation for generative adversarial networks. *arXiv preprint arXiv:1912.00953*.

Wu, Y.; Zhou, P.; Wilson, A. G.; Xing, E.; and Hu, Z. 2020. Improving gan training with probability ratio clipping and sample reweighting. *Advances in Neural Information Processing Systems*, 33: 5729–5740.