–Appendix–
Dynamically Masked Discriminator for GANs

Anonymous Author(s)
Affiliation
Address
email

Abstract

In this supplementary material, we first introduce the implementation details of our Dynamic Mask Discriminator (DMD), and then provide more technical details of the proposed DMD in Sec. 1. In Sec. 2.1, we elaborate on the additional details of datasets and experiment settings. In Sec. 2.3, we provide additional experimental results, where our DMD is integrated with more GANs models

We also report the error bars of our method in Sec. 3 and describe the broader impact (Sec. 4) and limitation (Sec. 5) of our method.

1 Implementation Details

In this section, we provide the implementation details of our proposed method. We have utilized the PyTorch [15] public platform to conduct our experiments. Our models were trained on a workstation equipped with 8 NVIDIA Tesla V100 GPUs with 32 GB memory capacity, a CPU of 2.8GHz, and 512GB RAM.

1.1 More details of the proposed method.

Our method is designed to automatically detect the retardation of the discriminator and force it to fast learn the new knowledge given time-varying distributions of generated samples. To show the detailed strategy of the proposed DMD method, the pseudo-code is given in Algorithm. 1. In our main paper, we have demonstrated the effectiveness of our proposed DMD method by integrating it with image generation techniques on StyleGAN-V2 [9].

Since the retardation of learning new knowledge can potentially be addressed by augmenting the input stream or the outcome of the discriminator, we also extend another two schemes to achieve dynamic discriminator adjustment. Besides the proposed dynamic feature masking scheme, we can dynamically mask the input and output of the discriminator when it is detected as retardation. The corresponding details are shown summarized in Algorithm. 2 and Algorithm. 3.

1.2 More details of model parameter difference

We detail *model parameter difference* which is used in the main paper Figure 3(a) for investigating the fixed discriminator of StyleGAN-V2. Model parameter difference metric is to measure the differences of the discriminator model weights between two adjacent training steps:

$$\mathbf{d}_{t_j} = \|W^d_{t_j} - W^d_{t_{j-1}}\|^2 \tag{1}$$

where $W^d_{t_j}$ is the parameter weight of d-th layer of the discriminator during t_j training step, and $\|\cdot\|^2$ is l2 norm. A smaller $\mathbf{d}_{t_j}$ indicates that the parameter weights of d-th layer at t_j training step are more similar to that at (t_{j-1}) training step, *i.e.*, parameter weights are updated slowly. In other words, given new arrival generated samples with new distributions, a smaller $\mathbf{d}_{t_j}$ indicates that the discriminator slows down the learning of new knowledge to some extent. In the main paper, we calculate $\mathbf{d}_i$ in the full connection layer of the discriminator.

Algorithm 1 Dynamic Mask Discriminator

Require:
Generator $\mathcal{G}_{\theta^t}$; Current Discriminator $D_\phi(\cdot)$; Non-Masked Discriminator $\mathcal{D}_\phi(\cdot)$; Dynamically Masked Discriminator $\bar{D}_\phi(\cdot)$; Training Step t;
d-th Layer of Discriminator $\mathbf{F}^{(d),t}$; Dynamic Mask $\mathbf{m}_d^t$;
d-th Masking Layer of Discriminator $\bar{\mathbf{F}}^{(d),t}$; Predefined Threshold λ; Retardation Metric R_t;
Set U_t Containing m Samples for Calculating Retardation Metric;
The number of training steps n_t; The number of images per training step n_s;
Ensure:
Initialize $R_t \leftarrow 0$ and $t \leftarrow 1$; Random $\mathbf{m}_d^t$;

```
while θ has not converged do
    for t = 1 to n_t do
        if R_t > λ then
            M^t ← m_d^t; M_T ← M^t; D_φ(·) ← D̄_φ(·)
        else
            M^t ← vector(1); M_T ← m_d^{t+1}; D_φ(·) ← 𝒟_φ(·)
        end if
        for s = 1 to n_s do
            F̄^{(d),t} ← F^{(d),t} ⊙ M^t
            L_{φ(t)} ← −E_{I,t}[log(D_φ(I))] − E_{z∼p_z,t}[log(1 − D_φ(𝒢(z, θ^t)))]
            θ_s ← Adam(∂φ(t)/∂θ_{s−1});
        end for
        ℛ_t = (1/m) Σ_{i∈U_t} (F̄_i^{(d),t} · F_i^{(d),t}) / (|F̄_i^{(d),t}||F_i^{(d),t}|)
    end for
end while
Return θ;
```

2 Additional Details on Experiments

2.1 Datasets and Experimental Settings

AFHQ-V2 [2] consists of 3 independent sub-datasets, which include around 5,000 closeups of cat, dog, and wildlife faces, respectively (denoted as **AFHQ-Cat**, **AFHQ-Dog**, and **AFHQ-Wild**). We utilized a high-quality Lanczos filter [11] to resize all images to a resolution of 256×256. We then conducted experiments on three sub-datasets while setting StyleGAN-V2 [9] as the baseline model. We maintain consistency with ADA [6], by using identical network architectures [9], weight demodulation [9], style mixing regularization [8], path length regularization, lazy regularization [9], equalized learning rate for all trainable parameters [5], non-saturating logistic loss [3] with R_1 regularization [14], and the Adam optimizer [10].

FFHQ [8] comprises 70,000 images of human faces, which we used for training after downscaling them to a resolution of 256×256. In this case, we set StyleGAN-V2 [9] as the baseline and used the same settings as those for AFHQ-V2.

LSUN-Church [19] includes 126,000 images of outdoor church. We downscale them to 256×256 as the training data. In this case, we also set StyleGAN-V2 [9] as the baseline and used the same settings as those for FFHQ.

2.2 Baselines

In accordance with previous studies [6, 4, 13], we have integrated our proposed method with StyleGAN-V2 [9]. In order to assess the effectiveness of our method, we have compared it with

Algorithm 2 Dynamic Mask Discriminator Assert in Input (***Input Masking***)

Require:
Generator $\mathcal{G}_{\theta^t}$; Discriminator $\mathcal{D}_\phi(\cdot)$; Training Step t; Dynamic Mask $\mathbf{m}^t$;
Predefined Threshold λ; Retardation Metric R_t;
Set U_t Containing m Samples for Calculating Retardation Metric;
d-th Layer of Discriminator $\mathbf{F}^{(d),t}$; After Masking Input $\bar{\mathbf{F}}^{(d),t}$;
The number of training steps n_t; The number of images per training step n_s;
Ensure:
Initialize $R_t \leftarrow 0$ and $t \leftarrow 1$; Random $\mathbf{m}^t$;

while θ has not converged **do**
  for $t = 1$ to n_t **do**
    if $R_t > \lambda$ **then**
      $\mathbf{M}^t \leftarrow \mathbf{m}^t$; $\mathbf{M}_T \leftarrow \mathbf{M}^t$;
    else
      $\mathbf{M}^t \leftarrow$ vector(1); $\mathbf{M}_T \leftarrow \mathbf{m}^{t+1}$;
    end if
    for $s = 1$ to n_s **do**
      $\bar{I} \leftarrow I \odot \mathbf{M}^t$; $\bar{\tilde{I}} \leftarrow \mathcal{G}(z, \theta^t) \odot \mathbf{M}^t$
      $L_{\phi(t)} \leftarrow -\mathbb{E}_{\bar{I},t}[log(D(\bar{I}))] - \mathbb{E}_{z\sim p_z,t}[log(1 - D(\bar{\tilde{I}})]$
      $\theta_s \leftarrow$ Adam$(\frac{\partial\phi(t)}{\partial\theta_{s-1}})$;
    end for
    $\mathcal{R}_t = \frac{1}{m}\sum_{i\in U_t} \frac{\bar{\mathbf{F}}_i^{(d),t} \cdot \mathbf{F}_i^{(d),t}}{|\bar{\mathbf{F}}_i^{(d),t}||\mathbf{F}_i^{(d),t}|}$
  end for
end while
Return θ;

state-of-the-art methods that improve discriminators through data augmentation, including ADA [6] and APA [4]. We have also compared our method with GANs that utilize regularization techniques, such as LC-Reg [16], zCR [20], InsGen [18], Adaptive Dropout [6], AdaptiveMix [13], MEE [12], and DynamicD [17].

2.3 Additional Experimental Results

Combining our method with more GAN models. To further show the effectiveness of our method, we additionally replace the discriminator of StyleGAN-V3 [7] with the proposed DMD, where StyleGAN-V3 contains two versions *i.e.*, StyleGAN-V3T and StyleGAN-V3R. Table A1 shows our method improves the FID of StyleGAN-V3R from 7.616 to 6.864 , and improves that of StyleGAN-V3T from 5.850 to 4.921. This is because our method facilitates the training of StyleGAN-V3's generator.

Table A1: Our method over StyleGAN-V3 [7] on AFHQ-Cat.

	FID ↓	IS ↑
StyleGAN-V3R	7.616	1.881
StyleGAN-V3R w/ Ours(DMD)	**6.864**	**1.915**
StyleGAN-V3T	5.850	1.916
StyleGAN-V3T w/ Ours(DMD)	**4.921**	**1.969**

Additional Generated Distribution Our main paper studies the generated distributions of StyleGAN-V2. Here, we additionally provide the distribution of the generated samples of APA [4] method on FFHQ[8] in Fig. A1. Fig. A1 also indicates that the generated distributions undergo dynamical and complex changes over time as the generator evolves during training. As a result, the generated samples are not independently and identically distributed (i.i.d) across the training progress, posing significant challenges in learning the generated distributions.

Algorithm 3 Dynamic Mask Discriminator Assert in Outcome Logits (***Dynamic Head***)

Require:
Generator $\mathcal{G}_{\theta^t}$; Discriminator $\mathcal{D}_\phi(\cdot)$; Training Step t;
Outcome Logit Number of $\mathcal{D}_\phi(\cdot)$ and $\bar{\mathcal{D}}_\phi(\cdot)$ k; Dynamic Mask $\mathbf{m}^t$;
d-th Layer of Discriminator $\mathbf{F}^{(d),t}$; After Masking Outcome Logits $\bar{\mathbf{F}}^{(d),t}$;
Predefined Threshold λ; Retardation Metric R_t;
Set U_t Containing m Samples for Calculating Retardation Metric;
The number of training steps n_t; The number of images per training step n_s;

Ensure:
Initialize $R_t \leftarrow 0$ and $t \leftarrow 1$; Random $\mathbf{m}^t$;

while θ has not converged **do**
  for $t = 1$ to n_t **do**
    if $R_t > \lambda$ **then**
      $\mathbf{M}^t \leftarrow \mathbf{m}^t$; $\mathbf{M}_T \leftarrow \mathbf{M}^t$;
    else
      $\mathbf{M}^t \leftarrow$ vector(1); $\mathbf{M}_T \leftarrow \mathbf{m}^{t+1}$;
    end if
    for $s = 1$ to n_s **do**
      $L_{\phi(t)} \leftarrow -\mathbb{E}_{I,t}[log(\sum(D(I) \odot \mathbf{M}^t)))] - \mathbb{E}_{z \sim p_z,t}[log(1 - \sum(D(\mathcal{G}(z, \theta^t)) \odot \mathbf{M}^t))$
      $\theta_s \leftarrow$ Adam($\frac{\partial \phi(t)}{\partial \theta_{s-1}}$);
    end for
    $\mathcal{R}_t = \frac{1}{m} \sum_{i \in U_t} \frac{\bar{\mathbf{F}}_i^{(d),t} \cdot \mathbf{F}_i^{(d),t}}{|\bar{\mathbf{F}}_i^{(d),t}||\mathbf{F}_i^{(d),t}|}$
  end for
end while
Return θ;

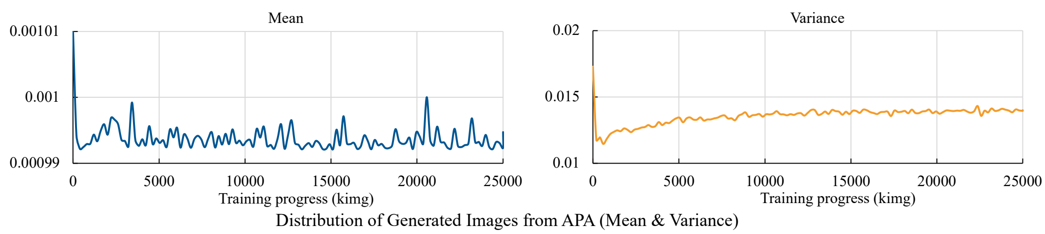

Figure A1: Illustration of time-varying distributions of generated samples in the training process of APA [4] on FFHQ [8]. The mean and variance generated samples' features show the generated distributions are dynamic and time-varying during training, as the generator evolves.

3 Error Bar

To evaluate the reproducibility of our method's results, we run our experiments three times using random seeds and the same hyper-parameters. Table A2 and Table A3 list the mean and variance of experimental results to show the error bar.

As shown in Table A2 and Table A3, our method performs stably on multiple datasets *i.e.*, AFHQ-V2, FFHQ and LSUN-Church datasets, indicating the reproducibility of our method.

4 Broader Impact

In this paper, we propose a novel method to improve the training of GANs, helping to generate high-quality images. Our method can be used for various applications such as producing training data and creating photorealistic images. On the other hand, like other generative models, our method can be misused for the application of Deepfake [1], where fake content is synthesized to deceive and mislead people, leading to a negative social impact. Nevertheless, many researchers have considered this problem, while exploring fake content detection and media forensics techniques. In addition, we

Table A2: Quantitative results of our method on AFHQ-V2 dataset [2], error bars are reported in terms of mean and variance, and Ours is StyleGAN-V2+DMD

	AFHQ-Cat		AFHQ-Dog		AFHQ-Wild	
	FID ↓	IS ↑	FID ↓	IS ↑	FID ↓	IS ↑
StyleGAN-V2	7.924	1.890	26.310	9.000	3.957	5.567
Ours (Reported in the paper)	5.879	1.988	21.240	9.698	3.471	5.647
Ours (re-run-1))	5.896	1.944	20.016	9.807	3.473	5.803
Ours (re-run-2))	6.015	1.987	20.456	10.291	3.420	5.699
Ours(Mean±Variance)	5.930±0.061	1.973±0.021	20.571±0.506	9.932±0.258	3.455±0.025	5.716±0.065

Table A3: Quantitative results of our method on FFHQ [8] and LSUN-Church [19], where error bars are reported in terms of mean and variance, and Ours is StyleGAN-V2+DMD.

	LSUN-Church (126K)		FFHQ(70K)	
	FID ↓	IS ↑	FID ↓	IS ↑
StyleGAN-V2	4.292	2.589	3.810	5.185
Ours (Reported in the paper)	3.061	2.792	3.299	5.204
Ours (re-run-1))	3.025	2.795	3.177	5.225
Ours (re-run-2))	2.993	2.787	3.285	5.200
Ours(Mean±Variance)	3.026±0.028	2.791±0.003	3.254±0.055	5.210±0.011

believe there would be regulations on fake content generation, such as forcing synthesized content to be injected with identifications that indicate it to be fake.

5 Limitations

By observing the time-varying distributions of the samples generated by the generator, we innovatively propose a method for training GANs, from the perspective of online continual learning. In this paper, we mainly show the challenges posed by the time-varying distributions, reveal that typical discriminators slow down their adaptation to the changes in the new arrival generated data, and propose a new method to address the challenges. Theoretical studies can make this work more comprehensive, however, we have not explored it in the paper, since it is beyond the scope of this paper. Moreover, while the proposed method can effectively improve the training of the CNN-based GANs models, combining our method with transformer-based ones is left to be investigated.

References

[1] Deepfake. https://en.wikipedia.org/wiki/Deepfake.

[2] Yunjey Choi, Youngjung Uh, Jaejun Yoo, and Jung-Woo Ha. Stargan v2: Diverse image synthesis for multiple domains. In *CVPR*, pages 8188–8197, 2020.

[3] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *NeurIPS*, 27:2672–2680, 2014.

[4] Liming Jiang, Bo Dai, Wayne Wu, and Chen Change Loy. Deceive d: adaptive pseudo augmentation for gan training with limited data. *NeurIPS*, 34:21655–21667, 2021.

[5] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation. *ICLR*, 2018.

[6] Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Training generative adversarial networks with limited data. *NeurIPS*, 33:12104–12114, 2020.

[7] Tero Karras, Miika Aittala, Samuli Laine, Erik Härkönen, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Alias-free generative adversarial networks. *NeurIPS*, 34:852–863, 2021.

[8] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *CVPR*, pages 4401–4410, 2019.

[9] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan. In *CVPR*, pages 8110–8119, 2020.

[10] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[11] Cornelius Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. 1950.

[12] Haozhe Liu, Bing Li, Haoqian Wu, Hanbang Liang, Yawen Huang, Yuexiang Li, Bernard Ghanem, and Yefeng Zheng. Combating mode collapse in gans via manifold entropy estimation. *arXiv preprint arXiv:2208.12055*, 2022.

[13] Haozhe Liu, Wentian Zhang, Bing Li, Haoqian Wu, Nanjun He, Yawen Huang, Yuexiang Li, Bernard Ghanem, and Yefeng Zheng. Adaptivemix: Improving gan training via feature space shrinkage. In *CVPR*, 2023.

[14] Lars Mescheder, Andreas Geiger, and Sebastian Nowozin. Which training methods for gans do actually converge? In *International conference on machine learning*, pages 3481–3490. PMLR, 2018.

[15] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. 2017.

[16] Hung-Yu Tseng, Lu Jiang, Ce Liu, Ming-Hsuan Yang, and Weilong Yang. Regularizing generative adversarial networks under limited data. In *CVPR*, pages 7921–7931, 2021.

[17] Ceyuan Yang, Yujun Shen, Yinghao Xu, Deli Zhao, Bo Dai, and Bolei Zhou. Improving gans with a dynamic discriminator. *NeurIPS*, 2022.

[18] Ceyuan Yang, Yujun Shen, Yinghao Xu, and Bolei Zhou. Data-efficient instance generation from instance discrimination. *NeurIPS*, 34:9378–9390, 2021.

[19] Fisher Yu, Ari Seff, Yinda Zhang, Shuran Song, Thomas Funkhouser, and Jianxiong Xiao. Lsun: Construction of a large-scale image dataset using deep learning with humans in the loop. *arXiv preprint arXiv:1506.03365*, 2015.

[20] Han Zhang, Zizhao Zhang, Augustus Odena, and Honglak Lee. Consistency regularization for generative adversarial networks. *arXiv preprint arXiv:1910.12027*, 2019.