Regularizing Generative Adversarial Networks under Limited Data

Hung-Yu Tseng*2, Lu Jiang1, Ce Liu1, Ming-Hsuan Yang1,2,4, Weilong Yang3

1Google Research 2University of California, Merced 3Waymo 4Yonsei University

Abstract

Recent years have witnessed the rapid progress of generative adversarial networks (GANs). However, the success of the GAN models hinges on a large amount of training data. This work proposes a regularization approach for training robust GAN models on limited data. We theoretically show a connection between the regularized loss and an $f$-divergence called LeCam-divergence, which we find is more robust under limited training data. Extensive experiments on several benchmark datasets demonstrate that the proposed regularization scheme 1) improves the generalization performance and stabilizes the learning dynamics of GAN models under limited training data, and 2) complements the recent data augmentation methods. These properties facilitate training GAN models to achieve state-of-the-art performance when only limited training data of the ImageNet benchmark is available. The source code is available at https://github.com/google/lecam-gan.

1. Introduction

Generative adversarial networks (GANs) [2, 7, 13, 44] have made significant progress in recent years on synthesizing high-fidelity images. The GAN models are the cornerstone techniques for numerous vision applications, such as data augmentation [11, 12], domain adaptation [18, 19], image extrapolation [60], image-to-image translation [20, 34, 75], and image editing [1, 4, 21, 63].

The success of the GAN methods heavily relies on a large amount of diverse training data which is often labor-expensive or cumbersome to collect [65]. As the example of the BigGAN [7] model presented in Figure 1, the performance significantly deteriorates under the limited training data. Consequently, several very recent approaches [28, 71, 1]...
have been developed to address the data insufficiency issue. A representative task in this emerging research direction aims to learn a robust class-conditional GAN model when only a small proportion of the ImageNet data [54] are available for the training. Generally, existing methods exploit data augmentation, either conventional or differentiable augmentation, to increase the diversity of the limited training data. These data augmentation approaches have shown promising results on several standard benchmarks.

In this paper, we address the GAN training task on limited data from a different perspective: model regularization. Although there are numerous regularization techniques for the GAN models in the literature [14, 45, 47, 57, 74], none of them aim to improve the generalization of the GAN models trained on limited data. In contrast, our goal is to learn robust GAN models on limited training data that can generalize well on out-of-sample data. To this end, we introduce a novel regularization scheme to modulate the discriminator’s prediction for learning a robust GAN model. Specifically, we impose an $\ell_2$ norm between the current prediction of the real image and a moving average variable that tracks the historical predictions of the generated image, and vice versa. We theoretically show that, under mild assumptions, the regularization transforms the WGAN [2] formulation towards minimizing an $f$-divergence called LeCam-divergence [33]. We find that the LeCam-divergence is more robust under the limited training data setting.

We conduct extensive experiments to demonstrate the three merits of the proposed regularization scheme. First, it improves the generalization performance of various GAN approaches, such as BigGAN [7] and StyleGAN2 [29]. Second, it stabilizes the training dynamics of the GAN models under the limited training data setting. Finally, our regularization approach is empirically complementary to the data augmentation methods [28, 71]. As presented in Figure 1, we obtain state-of-the-art performance on the limited (e.g., 10%) ImageNet dataset by combining our regularization (i.e., $R_{LC}$) and the data augment method [71].

### 2. Related Work

**Generative adversarial networks.** Generative adversarial networks (GANs) [2, 7, 13, 25, 29, 44, 68] aim to model the target distribution using adversarial learning. Various adversarial losses have been proposed to stabilize the training or improve the convergence of the GAN models, mainly based on the idea of minimizing the $f$-divergence between the real and generated data distributions [50]. For example, Goodfellow et al. [13] propose the saturated loss that minimizes the JS-divergence between the two distributions. Similarly, the LSGAN [44] formulation leads to minimizing the $\chi^2$-divergence [51], and the EBGAN [70] approach optimizes the total variation distance [2]. On the other hand, some models are designed to minimize the integral probability metrics (IPM) [48, 58], such as the WGAN [2, 14] frameworks. In this work, we design a new regularization scheme that can be applied to different GAN loss functions for training the GAN models on the limited data.

**Learning GANs on limited training data.** With the objective of reducing the data collection effort, several studies [15, 65] raise the concern of insufficient data for training the GAN models. Training the GAN models on limited data is challenging because the data scarcity leads to the problems such as unstable training dynamics, degraded fidelity of the generated images, and memorization of the training examples. To address these issues, recent methods [28, 62, 69, 71, 72, 73] exploit data augmentation as a mean to increase data diversity, hence preventing the GAN models from overfitting the training data. For example, Zhang et al. [69] augment the real images and introduce a consistency loss for training the discriminator. The DA [71] and ADA [28] approaches share a similar idea of applying differential data augmentation on both real and generated images, in which ADA further develops an adaptive strategy to adjust the probability of augmentation. In contrast to prior work, we tackle this problem from a different perspective of model regularization. We show that our method is conceptually and empirically complementary to the existing data augmentation approaches.

**Regularization for GANs.** Most existing regularization methods for GAN models aim to accomplish two goals: 1) stabilizing the training to ensure the convergence [45, 46], and 2) mitigating the mode-collapse issue [55]. As the GAN frameworks are known for unstable training dynamics, numerous efforts have been made to address the issue using noise [22, 57], gradient penalty [14, 30, 45, 53], spectral normalization [47], adversarial defense [74], etc. On the other hand, a variety of regularization approaches [5, 8, 41, 43, 59, 67] are proposed to alleviate the model-collapse issue, thus increasing the diversity of the generated images. Compared with these methods, our work targets a different goal: improving the generalization of the GAN models trained on the limited training data.

**Robust Deep Learning.** Robust deep learning aims to prevent the deep neural networks from overfitting or memorizing the training data. Recent methods have shown successes in overcoming training data bias such as label noise [3, 16, 23, 24, 52, 40, 49, 66] and biased data distributions [56, 10, 35, 38, 9]. Recently, few approaches [6, 26, 27, 61] have been proposed for learning the robust GAN model. While these approaches are designed to overcome label or image noise in a corrupted training set, we improve the generalization of the GAN models trained on the limited uncorrupted training data.
3. Methodology

We first review the GAN models, then detail our regularization scheme. Finally, we discuss the connection between the proposed method and the LeCam-divergence along with the effect on robust learning under the limited data setting.

3.1. Generative Adversarial Networks

A GAN model consists of a discriminator $D$ and a generator $G$. Let $V_D$ and $L_G$ denote the training objectives of the discriminator $D$ and generator $G$, respectively. The training of the GAN frameworks can be generally illustrated as:

$$
\begin{align*}
\max_D V_D, V_D &= \mathbb{E}_{x \sim \mathcal{T}}[f_D(D(x))] + \mathbb{E}_{z \sim p_z}[f_G(D(G(z)))] & (1) \\
\min_G L_G, L_G &= \mathbb{E}_{z \sim p_z}[g_G(D(G(z)))] & (2)
\end{align*}
$$

where $p_z$ is the prior distribution (e.g., $\mathcal{N}(0, I)$) and $\mathcal{T}$ is the training (observed) image set used to approximate the data distribution. The notations $f_D$, $f_G$, and $g_G$ in Eq. (2) represent the mapping functions from which various GAN losses can be derived (cf. [39]).

3.2. Regularizing GANs under Limited Data

Our goal is to improve the performance of the GAN models when the training set $\mathcal{T}$ merely contains a limited amount of data, as the example shown in Figure 1. Different from the existing data augmentation methods [28, 71], we approach this problem by incorporating the regularization on the discriminator. We present the overview of the proposed method in Figure 2. The core idea is to regulate the discriminator predictions during the training phase. Specifically, we introduce two exponential moving average [32] variables $\alpha_R$ and $\alpha_F$, called anchors, to track the discriminator’s predictions of the real and generated images. The computation of the anchors $\alpha_R$ and $\alpha_F$ is provided in the supplementary document. We then use the identical objective $L_G$ described in Eq. (2) for training the generator, and minimize the regularized objective $L_D$ for the discriminator:

$$\min_D L_D, L_D = -V_D + \lambda R_{LC}(D),$$

where $R_{LC}$ is the proposed regularization term:

$$R_{LC} = \mathbb{E}_{x \sim \mathcal{T}}[\|D(x) - \alpha_F\|^2] + \mathbb{E}_{z \sim p_z}[\|D(G(z)) - \alpha_R\|^2].$$

At first glance, the objective in Eq. (3) appears counterintuitive since the regularization term $R_{LC}$ pushes the discriminator to mix the predictions of real and generated images, as opposed to differentiating them. However, we show in Section 3.3 that $R_{LC}$ offers meaningful constraints for optimizing a more robust objective. Moreover, we empirically demonstrate in Section 4 that with the appropriate weight $\lambda$, this simple regularization scheme 1) improves the generalization under limited training data, and 2) complements the existing data augmentation methods.

Why moving averages? Tracking the moving average of the prediction reduces the variance across mini-batches and stabilizes the regularization term described in Eq. (4). Intuitively, the moving average becomes stable while the discriminator’s prediction gradually converges to the stationary point. We find this holds for the GAN models used in our experiments (e.g., Figure 8). We illustrate a general case of using two moving average variables $\alpha_R$ and $\alpha_F$ in Figure 2. In some cases, e.g., in theoretical analysis, we may use a single moving average variable to track the predictions of either real or generated images.

3.3. Connection to LeCam Divergence

We show the connection of the proposed regularization to the WGAN [2] model and an $f$-divergence called LeCam (LC)-divergence [33] or triangular discrimination [64]. Under mild assumptions, our regularization method can enforce WGANs to minimize the weighted LC-divergence. We show that the LC-divergence 1) can be used for training GAN models robustly under limited training data, and 2) has a close relationship with the $f$-divergences used in other GAN models [13, 44, 70].
Proposition 2 (Properties of LeCam-divergence). \textit{LC-divergence} $\Delta$ is an $f$-divergence with following properties:

\begin{itemize}
  \item $\Delta$ is non-negative and symmetric.
  \item $\Delta(p_d\|p_q)$ is bounded, with the minimum 0 when $p_d = p_q$ and the maximum 2 when $p_d$ and $p_q$ are disjoint.
  \item $\Delta$-divergence is a symmetric version of $\chi^2$-divergence, i.e., $\Delta(P||Q) = \chi^2(P||M) + \chi^2(Q||M)$, where $M = \frac{1}{2}(P + Q)$.
  \item The following inequalities hold [42]: $\frac{1}{2}\Delta(P, Q) \leq JS(P, Q) \leq \frac{1}{2}\Delta(P, Q) \leq \frac{1}{2}\text{TV}(P, Q)$, where $JS$ and $TV$ represent JS-divergence and Total Variation.
\end{itemize}

Proposition 2 shows that the LC-divergence is closely related to the $f$-divergences used in other GAN methods. For example, it is a symmetric and smoothed $\chi^2$-divergence used in the LSGAN [44]. The weighted $\Delta$ lower bounds the JS-divergence used in the saturated GAN [13] and the Total Variation distance used in the EBGAN [70] approaches.
Table 1: Quantitative results on the CIFAR dataset. We report the average FID scores ($\downarrow$) of three evaluation runs. The best performance is in **bold** and the second best is **underlined**.

<table>
<thead>
<tr>
<th>Methods</th>
<th>CIFAR-10</th>
<th></th>
<th>CIFAR-100</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full data 20% data</td>
<td>10% data</td>
<td>Full data 20% data</td>
<td>10% data</td>
</tr>
<tr>
<td>Non-saturated GAN [13]</td>
<td>9.83 ± 0.06</td>
<td>18.59 ± 0.15</td>
<td>41.99 ± 0.18</td>
<td>13.87 ± 0.08</td>
</tr>
<tr>
<td>LS-GAN [44]</td>
<td>9.07 ± 0.01</td>
<td>21.60 ± 0.11</td>
<td>41.68 ± 0.18</td>
<td>12.43 ± 0.11</td>
</tr>
<tr>
<td>RaHinge GAN [25]</td>
<td>11.31 ± 0.04</td>
<td>23.90 ± 0.22</td>
<td>48.13 ± 0.33</td>
<td>14.61 ± 0.21</td>
</tr>
<tr>
<td>BigGAN [7]</td>
<td>9.74 ± 0.06</td>
<td>21.80 ± 0.29</td>
<td>48.08 ± 0.10</td>
<td>13.60 ± 0.07</td>
</tr>
<tr>
<td>BigGAN + $R_{LC}$ (Ours)</td>
<td><strong>8.31 ± 0.05</strong></td>
<td><strong>15.27 ± 0.10</strong></td>
<td><strong>35.23 ± 0.14</strong></td>
<td><strong>11.88 ± 0.12</strong></td>
</tr>
</tbody>
</table>

Figure 4: FID curves during the training on the CIFAR-10 dataset. The proposed method 1) improves the best performance, and 2) stabilizes the training dynamic of the BigGAN model under the limited (e.g., 20%) data setting.

4. Experimental Results

We conduct extensive experiments on several benchmarks to validate the efficacy of our method on training the leading class-conditional BigGAN [7] and unconditional StyleGAN2 [29] models on the limited data.

Datasets. The CIFAR 10/100 [31] and ImageNet [54] datasets are standard benchmarks for the image generation models. The resolutions of the images in the CIFAR, ImageNet datasets are 32x32, and 128x128, respectively.

Evaluation metrics. We use two common metrics: Inception Score (IS) [55] and Fréchet Inception Distance (FID) [17]. Unless specified otherwise, we follow the evaluation protocol in the DA paper [71] that reports the average and standard deviation values over three evaluation trials.

Setups. We conduct the CIFAR experiments using the BigGAN [7] framework implemented by Zhao et al. [71].\(^1\) We train the BigGAN model on TPU for the ImageNet experiments.\(^2\) Finally, the StyleGAN2 [29] framework is trained and evaluated using the implementation from Zhao et al. [71] and Karras et al. [28].\(^3\) As for the hyperparameter settings, we use the decay factor of 0.99 for the exponential moving average variables. We set the regularization weight $\lambda$ to 0.3, 0.01 for the CIFAR, ImageNet experiments, respectively.

Baselines. We compare three types of baseline methods on the CIFAR datasets. The first group are GAN models that optimize various loss functions including non-saturated [13], LS [44], and RaHinge [25]. Second, we compare with three regularization methods: instance noise [57], zero-centered gradient penalty (GP-0) [45] and consistency regularization (CR) [69]. Finally, we compare with two recent differentiable data augmentation methods DA [71] and ADA [28] that address the limited data issue for GANs. For the experiments on other datasets, we focus on comparing with the state-of-the-art methods. For a fair comparison, we compare the baseline methods under the same GAN backbone using their official implementation on each dataset, except Table 5 in which we cite the numbers of [71] reported in the original paper.

<table>
<thead>
<tr>
<th>Method</th>
<th>CIFAR-10</th>
<th></th>
<th>CIFAR-100</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full data 20% data</td>
<td>10% data</td>
<td>Full data 20% data</td>
<td>10% data</td>
</tr>
<tr>
<td>BigGAN [7]</td>
<td>9.74 ± 0.06</td>
<td>21.86 ± 0.29</td>
<td>13.60 ± 0.07</td>
<td>32.99 ± 0.24</td>
</tr>
<tr>
<td>+ noise [57]</td>
<td>9.64 ± 0.06</td>
<td>21.87 ± 0.11</td>
<td>13.85 ± 0.07</td>
<td>32.98 ± 0.01</td>
</tr>
<tr>
<td>+ CR [69]</td>
<td>8.96 ± 0.10</td>
<td>20.62 ± 0.10</td>
<td><strong>11.59 ± 0.05</strong></td>
<td>36.91 ± 1.12</td>
</tr>
<tr>
<td>+ GP-0 [45]</td>
<td>10.30 ± 0.16</td>
<td>19.10 ± 0.08</td>
<td>14.67 ± 0.08</td>
<td>29.85 ± 0.04</td>
</tr>
<tr>
<td>+ $R_{LC}$ (Ours)</td>
<td><strong>8.31 ± 0.05</strong></td>
<td><strong>15.27 ± 0.10</strong></td>
<td><strong>11.88 ± 0.12</strong></td>
<td><strong>25.51 ± 0.19</strong></td>
</tr>
</tbody>
</table>

4.1. Results on CIFAR-10 and CIFAR-100

As shown in Table 1, the proposed method improves the generalization performance of the BigGAN model. The comparison between other GAN models shows the competitive performance of the proposed method, especially under limited training data. These results substantiate that our regularization method minimizes a sensible divergence on limited training data. To further understand the impact on the training dynamics, we plot the FID scores during the training stage in Figure 4. The proposed method stabilizes the training process on limited data (i.e., FID scores deteriorate in a later stage) and achieves the lowest FID score at the final iteration (100K). This result suggests that our method can stabilize the GAN training process on limited data.

---

\(^1\)https://github.com/mit-han-lab/data-efficient-gans

\(^2\)https://github.com/google/compare_gan

\(^3\)https://github.com/NVlabs/stylegan2-ada
Table 3: **Quantitative comparisons to data augmentation.** We report the average FID (↓) scores of three evaluation runs.

<table>
<thead>
<tr>
<th>Methods</th>
<th>CIFAR-10 Full</th>
<th>10%</th>
<th>CIFAR-100 Full</th>
<th>10%</th>
<th>StyleGAN Full</th>
<th>1K</th>
</tr>
</thead>
<tbody>
<tr>
<td>BigGAN [7] + DA [71]</td>
<td>8.75 ± 0.05</td>
<td>23.34 ± 0.28</td>
<td>11.99 ± 0.10</td>
<td>35.39 ± 0.16</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BigGAN + DA + $R_L$ (Ours)</td>
<td>8.46 ± 0.06</td>
<td>16.69 ± 0.02</td>
<td>11.20 ± 0.09</td>
<td>27.28 ± 0.05</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>StyleGAN2 [29] + ADA [28]</td>
<td>2.68 ± 0.02</td>
<td>6.72 ± 0.03</td>
<td>3.04 ± 0.02</td>
<td>14.06 ± 0.07</td>
<td>3.82 ± 0.01</td>
<td>23.27 ± 0.14</td>
</tr>
<tr>
<td>StyleGAN2 + ADA + $R_L$ (Ours)</td>
<td>2.47 ± 0.01</td>
<td>6.65 ± 0.02</td>
<td>2.99 ± 0.01</td>
<td>13.01 ± 0.02</td>
<td>3.49 ± 0.04</td>
<td>21.70 ± 0.06</td>
</tr>
</tbody>
</table>

Table 4: **Quantitative results on the ImageNet dataset.** We report the mean IS (↑) and FID (↓) scores of three training runs.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Full data IS ↑</th>
<th>FID ↓</th>
<th>50% data IS ↑</th>
<th>FID ↓</th>
<th>25% data IS ↑</th>
<th>FID ↓</th>
</tr>
</thead>
<tbody>
<tr>
<td>BigGAN [7]</td>
<td>90.48 ± 12.7</td>
<td>8.60 ± 1.08</td>
<td>80.26 ± 5.55</td>
<td>9.83 ± 0.94</td>
<td>61.05 ± 6.43</td>
<td>18.22 ± 2.59</td>
</tr>
<tr>
<td>BigGAN + $R_L$ (Ours)</td>
<td>93.00 ± 13.27</td>
<td>7.27 ± 0.14</td>
<td>89.94 ± 6.67</td>
<td>9.13 ± 0.84</td>
<td>65.66 ± 4.96</td>
<td>14.47 ± 1.73</td>
</tr>
</tbody>
</table>

Figure 5: **Zero-centered gradient penalty values.** We visualize the values of zero-centered gradient penalty (GP-0) [45] during the training stage. The proposed regularization also constrains the values without explicitly minimizing the GP-0 loss.

4.2. **Comparison to State-of-the-art on ImageNet**

ImageNet [54] is a challenging dataset since it contains more categories and images with higher resolution. Considering the variance of the model performance, we follow the evaluation protocol in the BigGAN paper [7]. Specifically, we run the training/evaluation pipeline three times using different random seeds, and then report the average performance. We present the quantitative results in Table 4. The proposed method improves the resistance of the BigGAN model against the scarce training data issue (e.g., ↓ 3.75 in FID under 25% data). It is noteworthy that the performance variance of our models is reduced in most cases (e.g., 2.59 → 1.73 in FID under 25% data), suggesting its capability in stabilizing the training process.

Table 5 demonstrates the quantitative results compared to the state-of-the-art model that uses the DA [71] method. Both the quantitative results and qualitative comparison presented in Figure 6 validate that the proposed method complements the data augmentation approach. We achieve state-of-the-art performance on the limited (e.g., 10%) ImageNet dataset by combining our regularization and the data augment approaches.

4.3. **Comparison with Data Augmentation**

We use the StyleGAN dataset to conduct the experiments. Experiment details are provided in the supplementary document. As presented in Table 3 and Table 6, the proposed method improves the performance of the StyleGAN2 model trained with(out) data augmentation [28] in all cases. We note that different from BigGAN, the StyleGAN2 model minimizes the non-saturated [13] GAN loss and uses the gradient penalty GP-0 [45] in the default setting. This shows that the proposed regularization scheme can be applied to other GAN loss functions along with existing regularization approaches.

We make a comparison in Table 7 to summarize the (dis)advantages of the data augmentation and our methods. First, the data augmentation approaches yield more...
significant gain than the proposed method when the training data is extremely limited. Nevertheless, our method can further improve the performance of data augmentation due to the complementary nature of the two methods. Second, the data augmentation approaches may degrade the performance when the training images are sufficiently diverse (e.g., the full dataset). This is consistent with the observation described in [28]. In comparison, our regularization method may not suffer the same problem.

4.4. Analysis and Ablation Studies

We use the BigGAN model and the CIFAR-10 dataset to conduct the analysis and ablation studies.

Regularization strength for $R_{LC}$. We conduct a sensitive study on the regularization weight $\lambda$. As shown in Figure 7(b), weights greater than 0.5 degrade the performance. This agrees with our analysis in Eq. (7) that larger weights $\lambda$ result in negative divergence values. Generally, the proposed method is effective when the weight $\lambda$ is in a reasonable range, e.g., $[0.1, 0.5]$ in Figure 7(b).

Regularizing real vs. generated image predictions. Our default method regularizes the predictions of both real images $D(x)$ and generated images $D(G(z))$. In this experiment, we investigate the effectiveness of separately regularizing the two terms $D(x)$ and $D(G(z))$. As shown in

Table 5: Comparison to the state-of-the-art on the limited ImageNet training data. We train and evaluate the BigGAN [7] model following the same evaluation protocol in [71]. † denotes the result is quoted from [71].

<table>
<thead>
<tr>
<th>Methods</th>
<th>Full data</th>
<th>50% data</th>
<th>25% data</th>
<th>10% data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IS ↑</td>
<td>FID ↓</td>
<td>IS ↑</td>
<td>FID ↓</td>
</tr>
<tr>
<td>DA [71] (Zhao et al.)</td>
<td>100.8 ± 0.2†</td>
<td>6.80 ± 0.02†</td>
<td>91.9 ± 0.5†</td>
<td>8.88 ± 0.06†</td>
</tr>
<tr>
<td>DA + $R_{LC}$ (Ours)</td>
<td>108.0 ± 0.6</td>
<td>6.54 ± 0.03</td>
<td>91.7 ± 0.6</td>
<td>8.59 ± 0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>64.7 ± 0.6†</td>
</tr>
</tbody>
</table>

Figure 6: Qualitative comparisons under limited training data. We show the generation results on the (top) 10% and (bottom) 25% ImageNet dataset. The baseline models trained with our approach synthesize more realistic images.
Table 6: Quantitative results of the StyleGAN2 [29] model. We report the average FID (\(\downarrow\)) scores of three evaluation runs.

<table>
<thead>
<tr>
<th>Method</th>
<th>70k images</th>
<th>30k images</th>
<th>10k images</th>
<th>5k images</th>
<th>1k images</th>
</tr>
</thead>
<tbody>
<tr>
<td>StyleGAN2 [29]</td>
<td>3.79 ± 0.02</td>
<td>6.19 ± 0.05</td>
<td>14.96 ± 0.05</td>
<td>25.88 ± 0.09</td>
<td>72.07 ± 0.04</td>
</tr>
<tr>
<td>StyleGAN2 + (R_{\text{LC}}) (Ours)</td>
<td>3.66 ± 0.02</td>
<td>5.78 ± 0.03</td>
<td>14.58 ± 0.04</td>
<td>23.83 ± 0.11</td>
<td>63.16 ± 0.11</td>
</tr>
</tbody>
</table>

Table 7: Comparisons with data augmentation methods. We report the FID (\(\downarrow\)) scores of the StyleGAN2 backbone.

<table>
<thead>
<tr>
<th>Method</th>
<th>Full data</th>
<th>1K data</th>
</tr>
</thead>
<tbody>
<tr>
<td>StyleGAN2 [29]</td>
<td>3.71 ± 0.05</td>
<td>72.07 ± 0.04</td>
</tr>
<tr>
<td>+ DA [71]</td>
<td>4.21 ± 0.03</td>
<td>25.17 ± 0.09</td>
</tr>
<tr>
<td>+ ADA [28]</td>
<td>3.81 ± 0.04</td>
<td>23.27 ± 0.14</td>
</tr>
<tr>
<td>+ (R_{\text{LC}})</td>
<td>3.66 ± 0.02</td>
<td>63.16 ± 0.11</td>
</tr>
<tr>
<td>+ ADA + (R_{\text{LC}})</td>
<td>3.49 ± 0.04</td>
<td>21.70 ± 0.06</td>
</tr>
</tbody>
</table>

Table 8: Ablation study on regularizing real vs. generated image predictions. We train and evaluate the BigGAN [7] model on the CIFAR-10 dataset, then report the average FID (\(\downarrow\)) scores.

<table>
<thead>
<tr>
<th>Real</th>
<th>Generated</th>
<th>Full data</th>
<th>20% data</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>✓</td>
<td>8.73 ± 0.04</td>
<td>20.47 ± 0.36</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
<td>8.79 ± 0.09</td>
<td>18.18 ± 0.08</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
<td>8.81 ± 0.03</td>
<td>15.27 ± 0.10</td>
</tr>
</tbody>
</table>

Figure 7: Different (a) model sizes and (b) regularization strengths. The scores are computed on the (a) 10% and (b) 20% CIFAR-10 datasets.

Figure 8: Discriminator predictions. We visualize the discriminator predictions from the BigGAN model on the CIFAR-10 dataset during the training stage. The proposed method prevents the predictions of real images \(D(x)\) and generated images \(D(G(z))\) from diverging under the limited (e.g., 20%) data setting.

In future, we plan the training data scarcity issue for 1) the conditional GAN tasks such as image extrapolation, image-to-image translation, etc, and 2) the robust GAN learning on large-scale noisy training data.

5. Conclusion and Future Work

In this work, we present a regularization method to train the GAN models under the limited data setting. The proposed method achieves a more robust training objective for the GAN models by imposing a regularization loss to the discriminator during the training stage. In the experiments, we conduct experiments on various image generation datasets with different GAN backbones to demonstrate the efficacy of the proposed scheme that 1) improves the performance of the GAN models, especially under the limited data setting and 2) can be applied along with the data augmentation methods to further enhance the performance. In future, we plan the training data scarcity issue for 1) the conditional GAN tasks such as image extrapolation, image-to-image translation, etc, and 2) the robust GAN learning on large-scale noisy training data.

Acknowledgements

This work is supported in part by the NSF CAREER Grant #1149783.
References


