A. Details of compared methods.

We provide the training procedure for compared methods. All methods build upon the DiffAug CR-GAN [7] baseline except the experiments of Tab. 6. DiffAug CR-GAN only use labeled samples for training. Other baseline methods have label distribute function $C$ for unlabeled samples. $S^3$GAN [2], OSSGAN [1], and CurriculumGAN have a classifier trained simultaneously. RandomGAN has $C(x) = e^y$ where $y$ are sampled from $\mathcal{Y}$ uniformly. Here, $e^y$ is the $y$-th standard basis vector of $\mathbb{R}^K$. SingleGAN has $C(x) = [1/K, 1/K, \ldots, 1/K]^T$. Semi-supervised baselines assign temporal labels to unlabeled samples using $C$. Namely, the labels the discriminator takes are defined as

$$y' = \begin{cases} y & x \text{ is labeled} \\ C(x) & \text{otherwise} \end{cases}, \quad (1)$$

where $x$ is a sample, and $y$ is a corresponding label. For extended baselines with re-labeling strategy and CurriculumGAN, the labels for the discriminator are defined as

$$y' = \begin{cases} \frac{(y + C(x))}{2} & x \text{ is labeled} \\ C(x) & \text{otherwise} \end{cases}. \quad (2)$$

CurriculumGAN assigns the confidence $c$ to real unlabeled samples in the curriculum learning manner of open-set semi-supervised learning [6] with Otsu-threshold [3]. To achieve the best performance, CurriculumGAN corrects labels of labeled data and applies the curriculum learning to only unlabeled data. In other words, it assigns the high confidence $c = 1$ to all labeled data.

B. Additional results

We run additional experiments on TinyImageNet [5] to see the behavior of our method on different numbers of unlabeled samples. We conduct experiments on five configurations: 10%, 30%, 50%, 70%, 100% usage ratios. Other parameters are the same: 150 closed-set classes, 30% label noise, and 20% labeled samples. Our method achieves a consistent performance gain on different numbers of unlabeled samples, as shown in Fig. A. Figure B provide qualitative comparison in the experiments on ImageNet [4] with 200 closed-set classes and 4% labeled data. We can observe the performance gains of our method in terms of both image quality and diversity.

References

Figure B. Visual comparisons of class-conditional image synthesis results on ImageNet. Our method produces diverse images while respecting the given condition.


