Proceedings of Machine Learning Research 222, 2023 ACML 2023

Supplementary Materials of "Generative Semi-supervised Learning with Meta-Optimized Synthetic Samples"

Shin'ya Yamaguchi SHINYA.YAMAGUCHI@NTT.COM
NTT, Tokyo

Editors: Berrin Yanıkoğlu and Wray Buntine

A. Extended Related Work

Dai et al. (2017) and Dhar et al. (2021) have shown semi-supervised learning methods based on generative models. A major difference between these and our work is the requirement to train a generative model on target datasets. It is well known that training generative models can be costly, unstable, and low-quality. In this sense, the existing methods have tackled the problems by improving the training generative models on target datasets. In contrast, our method skips this training by using generative foundation models and produces useful synthetic samples by latent meta optimization.

B. Application to Medical Imaging

To demonstrate the applicability, we further evaluated our method MP-SSL on the Chaoyang dataset, which is for a medical imaging task classifying cancers. Table I shows that our method performs well in medical imaging.

Table I: Performance comparison of ResNet-18 classifiers.

Method / Dataset	Chaoyang	Food-101
Base Model	$81.88^{\pm.11}$	$77.43^{\pm.03}$
Oracle SSL ($\mathcal{D} + \mathcal{D}_\mathrm{u}$)		
FreeMatch	$81.24^{\pm.29}$	$77.12^{\pm.81}$
Generative SSL ($\mathcal{D} + G_\mathrm{F}$)		
Naïve gSSL (FreeMatch)	$81.60^{\pm.76}$	$77.59^{\pm.20}$
P-SSL	$81.28^{\pm.76}$	$77.59^{\pm.20}$
MP-SSL (Ours)	$\mathbf{82.53^{\pm.32}}$	$\mathbf{78.59^{\pm.17}}$

C. Scalability on larger datasets and models.

In the real world, it is difficult to construct datasets with more than millions of samples, and thus, target datasets are basically small. Our experiments concentrated on the evaluation

Table II: Top-1 Acc. (%) of ResNet-18.

Method / Dataset	10%-Aircraft	10%-Birds	10%-Cars	10%-DTD	10%-Flower	10%-Pets
Base Model	$12.63^{\pm.61}$	$26.22^{\pm.65}$	$19.74^{\pm.15}$	$47.93^{\pm.19}$	$50.44^{\pm.57}$	$74.79^{\pm.56}$
Oracle SSL ($\mathcal{D} + \mathcal{D}_{\mathrm{u}}$)						
FreeMatch	$12.10^{\pm.34}$	$25.70^{\pm.47}$	$18.07^{\pm.83}$	$46.08^{\pm.52}$	$49.72^{\pm.69}$	$75.35^{\pm1.3}$
Generative SSL ($\mathcal{D} + G_{\mathrm{F}}$)						
P-SSL	$12.98^{\pm.28}$	$26.99^{\pm.28}$	$21.78^{\pm.31}$	$45.51^{\pm.28}$	$50.07^{\pm.11}$	$75.75^{\pm1.2}$
MP-SSL (Ours)	$\mathbf{15.48^{\pm.33}}$	$\mathbf{27.66^{\pm.13}}$	$\mathbf{24.62^{\pm.21}}$	$\mathbf{49.27^{\pm.58}}$	$\mathbf{54.78^{\pm.65}}$	$\mathbf{76.65^{\pm.50}}$

in such a realistic setting. Nevertheless, the meta-optimization of our method is done for each batch, so it works regardless of the dataset size; Table I shows that our method is effective on a larger Food-101 ($\approx$ 100,000 samples). Also, Table IV shows the scalability of our method for the larger architectures. However, since the proposed method requires backpropagation from the classifier during meta-training, the computational cost increases as the classifier size increases.

D. Additional Results on 10% Datasets

In the main paper, we only showed the results on smaller datasets, where the proposed method is more effective. Here, we show the full evaluation results on 10% datasets in Table II. We see that our method outperforms the baselines for all datasets.

E. Comparison to using real ImageNet

Table III shows that our method outperforms the SSL with real ImageNet (Transfer SSL), indicating the synthetic samples from our meta-learning-based method are superior to real samples. This also justifies the use of a generative model for SSL.

F. Ablation Study of $\mathcal{L}_{\mathrm{gap}}$

As the alternative metric, we also tried maximum mean discrepancy (MMD), which is a measure of the distribution gap, but it does not improve mean squared error (MSE). Since MSE is the simplest to implement, we chose the form of Eq. (8) as our proposed method.

References

Zihang Dai, Zhilin Yang, Fan Yang, William W Cohen, and Russ R Salakhutdinov. Good semi-supervised learning that requires a bad gan. *Advances in neural information processing systems*, 30, 2017.

Sauptik Dhar, Javad Heydari, Samarth Tripathi, Unmesh Kurup, and Mohak Shah. Universum gans: Improving gans through contradictions. *arXiv preprint arXiv:2106.09946*, 2021.

Table III: Top-1 Accuracy (%) of ResNet-18.

Method / Dataset	Cars
Base Model	$71.62^{\pm.30}$
Oracle SSL $(\mathcal{D}+\mathcal{D}_\mathrm{u})$	
FreeMatch	$\underline{82.73}^{\pm.41}$
SCR	$75.11^{\pm.14}$
Transfer SSL $(\mathcal{D}+\mathcal{D}_\mathrm{s})$	
FreeMatch	$72.76^{\pm2.4}$
SCR	$73.53^{\pm.21}$
Generative SSL $(\mathcal{D}+G_\mathrm{F})$	
Naïve gSSL (FreeMatch)	$73.67^{\pm.67}$
P-SSL	$72.45^{\pm.30}$
MP-SSL (Ours)	$\mathbf{76.33}^{\pm\mathbf{.31}}$

Table IV: Top-1 Accuracy (%) on Cars.

Method / Architecture	ResNet-50	ResNet-101
Base Model	$77.83^{\pm.30}$	$78.24^{\pm.28}$
Oracle SSL $(\mathcal{D}+\mathcal{D}_\mathrm{u})$		
FreeMatch	$84.87^{\pm.21}$	$85.68^{\pm.09}$
Generative SSL $(\mathcal{D}+G_\mathrm{F})$		
Naïve gSSL (FreeMatch)	$79.65^{\pm.23}$	$80.48^{\pm.37}$
P-SSL	$78.68^{\pm.16}$	$78.10^{\pm.20}$
MP-SSL (Ours)	$\mathbf{81.59}^{\pm\mathbf{.55}}$	$\mathbf{83.65}^{\pm\mathbf{.26}}$

Table V: Comparison of $\mathcal{L}_\mathrm{gap}$ for MP-SSL

Loss form	Cars Test Acc.(%)
MSE (Eq. (8))	$76.33^{\pm.31}$
MMD	$76.12^{\pm.77}$