이미지 분류 문제에서의 반지도학습을 위한 경계성 데이터 생성 비교

Marcella Astrid[1,2], 이승익 [1,2]
[1] 과학기술연합대학원대학교, [2] 한국전자통신연구원

marcella.astrid@ust.ac.kr, the_silee@etri.re.kr

The Comparisons of Boundary Data Generation for Semi-Supervised Learning in Image Classification

Marcella Astrid[1,2], Seung-Ik Lee[1,2 (corresponding author)]
[1]University of Science and Technology
[2]Electronics and Telecommunications Research Institute

Abstract

In order to reduce the cost of labeling, semi-supervised learning technique uses only small amount of labeled data with unlabeled data to train a deep neural network. One way to exploit the available data, both labeled and unlabeled, is by training a generative model to provide more data to improve the main model, i.e., in this work, image classifier. Besides good quality generated data, bad quality generated data has also been shown effective to tighten the decision boundary of the trained classifier. However, generating all of bad quality data in the whole data space may be intractable. Therefore, generating boundary data, i.e., data just outside of the real data distribution, may be easier while still effective in tightening the classifier boundary. In this work, we compare two methods to define and generate boundary data: data between real and noise versus data between real and fake. Experiments on 100 labeled data MNIST show the data between real and noise is more superior.

Ⅰ. INTRODUCTION

As the efforts to reduce the need of data labeling, semi-supervised learning (SSL) techniques utilize only small amount of labeled data in addition to abundant unlabeled data. One kind of technique in SSL is by generating additional data to assist the training of the main model. In this paper, we focus on training image classifier model using the available data.

The target of SSL is to make the most of the unlabeled data so that the trained model achieves a better performance compared to if the model is trained only on the small amount of labeled data. Apart from generating good quality data similar to the typical generative models [1], generating bad data has also been known to be beneficial in helping the training of classifier in SSL setting [2, 3]. The bad data is used to tighten the boundary of the classifier such that the classifier has high confidence only on the real data while it has low confidence on the bad quality data.

However, covering all of the bad data space may be very difficult due to the immense size of the data space, while it may be not necessary to cover every bad data space. Therefore, to reduce the bad data space while keeping the benefit of bad data in assisting the training, we can focus on boundary data, i.e., data just outside of real/good data.

In this work, we compare two definitions of boundary data and approaches to generate it. First, boundary data as data between real and noise by Astrid *et al.* [2]. Second, inspired by Fence GAN [4] from anomaly detection field, we design boundary data as data between real and fake for our SSL setup. We conduct experiments on MNIST with 100 labeled data to compare both methods.

Ⅱ. METHODOLOGY

We use three players mechanism in [2] consisting of classifier, generator, and discriminator. Classifier takes image input and outputs its class prediction. It is the main model which we measure the performance after the training. It learns to predict high confident outputs for real data (labeled and unlabeled) while producing low confident in bad quality data. To be more specific, it is trained to minimize cross entropy loss, minimize entropy, and maximize entropy for labeled, unlabeled, and generated data respectively. Generator assists the classifier by providing fake image data. In this work, it learns with the aids from discriminator to generate either data between real and noise or between real and fake. To help the generator in generating data between real and noise, differently from typical generative models, Astrid *et al.*, [2] set the discriminator to learn discriminating between real image data and noise data. Meanwhile, for the case of generating data between real and fake, the discriminator is set similarly with classic generative models, i.e., to discriminate between real and fake data.

In this work, the generator, unlike the usual generative models, should not fool the discriminator completely. Rather than targeting to generate real data, i.e., make the discriminator outputs 1, it tries to make the discriminator produces a hyperparameter $\theta \in (0, 1)$ value.

Ⅲ. EXPERIMENTS

We train our baseline, our model with data between real and noise, and our model with data between real and fake using MNIST using 100 labeled and 59900 unlabeled data. The baseline classifier is trained fully supervised using only the 100 labeled data. Hyperparameter θ is set to 0.5 for generating data between real and noise while it is set to 0.9 for data between real and fake. Other training hyperparameters follow setup in [2]. The aforementioned three models (i.e., baseline, our model with data between real and noise, our model with data between real and fake) respectively achieve test error of 14.79%, 1.83%, and 2.90%. Both of the models using boundary data outperform the baseline, which demonstrate the importance of boundary data. Additionally, using data between real and noise outperforms using data between real and fake. The inferiority of the last model may be due to moving distribution of fake data compared to the fixed distribution of noise data. Moving target distribution may cause the training of the generator not optimal.

Ⅳ. CONCLUSION

We compare two methods in generating boundary data for semi-supervised image classification. First, we define boundary data as data between real and noise. The other method as data between real and fake. Through our experiments using MNIST dataset with 100 labeled data, we find data between real and noise is more superior as the other method may have instability as fake data distribution is moving as the model evolved.

ACKNOWLEDGMENT

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2022-0-00951, Development of Uncertainty-Aware Agents Learning by Asking Questions).

REFERENCES

[1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. “Generative adversarial networks,” Communications of the ACM, 63(11), pp.139-144, 2020.

[2] Astrid, M., Zaheer, M.Z. and Lee, S.I. “Semi Supervised Learning Using Data Between Real and Noise,” 한국통신학회 학술대회논문집, pp.545-546, 2022.

[3] Dai, Z., Yang, Z., Yang, F., Cohen, W.W. and Salakhutdinov, R.R. “Good semi-supervised learning that requires a bad gan.” Advances in neural information processing systems, 30, 2017.

[4] Ngo, P.C., Winarto, A.A., Kou, C.K.L., Park, S., Akram, F. and Lee, H.K. “Fence GAN: Towards better anomaly detection,” IEEE 31st International Conference on tools with artificial intelligence (ICTAI) (pp. 141-148), November 2019.