Improve Data Robustness With Multiple Data Processing

Hongyu Kan, Jun Shi, Hong An

[1]University of Science and Technology of China
{honeyk, shijun18, han}@mail.ustc.edu.cn

Abstract

With the wide application of machine learning technology, more and more application scenarios have higher requirements for the robustness of machine learning. Robust machine learning methods can deal with some unforeseen situations in the learning process, such as new application scenarios or external malicious attacks. In terms of robust learning, defense methods based on deep learning model have been proposed to reduce the potential threat against samples, but most methods pursue high-performance models under fixed constraints and data sets. Therefore, how to construct a general and effective data set to train robust models has not been widely explored. In this paper, we proposed a method to improve model robustness by processing training data. We have verified our method through CIFAR10, and our method significantly improves the robustness of the model.

Introduction

In the past decade, considerable progress has been made in many tasks in the field of machine learning, such as image classification, target detection, machine translation or question answering. The deep neural network can easily become the most advanced method (Canziani, Paszke, and Culurciello 2016; Liu et al. 2017). Verifying the effectiveness of a deep learning method is usually carried out on some well-known data sets such as MNIST, CIFAR and ImageNet (Deng 2012; Ho-Phuoc 2018; Krizhevsky, Sutskever, and Hinton 2012). In image classification, a large number of deep learning methods have emerged in recent years, such as Resnet, Densenet, Efficientnet and so on (He et al. 2016a; Huang et al. 2017; Tan and Le 2019). In some tasks, the classification accuracy of these networks even exceeds the human level.

Usually, the training set and test set are generated by the same underlying distribution, which makes the performance of the model under the change of distribution unknown. In view of the fact that machine learning technology is being used for sensitive tasks, such as self-driving cars and medical care, robustness should be considered as a key indicator when evaluating the performance of the model. Let's take the example of automatic driving. If the training data set only selects the street scenes of a certain city for a certain period of time when training the model of automatic driving, the trained model may only recognize these similar street scenes. If there is some unusual weather in the city, such as blizzard or fog, which does not appear in the scene of the training data set, those cars based on automatic driving are likely to have accidents when driving on the road, and the consequences may be unimaginable.

To improve the performance of the model and adapt to the robustness in different scenarios, many researchers have done a lot of research. For example, using pre-training parameters or self-supervised learning can improve the robustness of the model (Hendrycks, Lee, and Mazeika 2019; Hendrycks et al. 2019). In the process of network training, using randomness or adjusting better training parameters can also improve the performance of the network (Yang, Chen, and Gangopadhyay 2020; Jia, Fang, and Zhang 2021). In addition, in order to evaluate the robustness of the model, some researchers have proposed some feasible evaluation methods (Ilie, Popescu, and Stefanescu 2020).

Nevertheless, most of the current methods for robustness learning pursue high-performance models under fixed constraints and data sets. Therefore, the research on how to build general and effective data sets to train robust models is very limited. In the general training process, data augment is often an effective means to improve the generalization ability of the model. Common data augment, such as random rotation, random flip and random erase, have played a good role in some tasks (Zhong et al. 2020; Yu et al. 2021; Dabouei et al. 2021). However, in the presence of noise, the benefits of data enhancement may be greatly reduced (Yin et al. 2019)[Yin D et al., 2019]. Therefore, in order to further improve the robustness of the model, it is very important to take data noise into account.

In this paper, we proposed a method to improve model robustness by processing training data. In order to improve the robustness of the model from the perspective of data, we first carry out median filtering on the image to filter out the details of the image. Then we add a variety of noise to the image, and improve the anti-interference ability of the model in this way. Necessary data augment and appropriate training methods are also very important. We have obtained appropriate parameters through experiments to achieve the ideal effect.

Copyright © 2022, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

The main contributions of this paper are as follows: 1. We propose a method to improve model robustness by processing training data 2. Our method not only improves the anti-interference ability of the original scene, but also improves the robustness to different scenes. 3. Our method provides a solution for robust learning from the perspective of data.

Multiple Data Processing

Median Filter

For robustness learning, too many local details are often easy to make the model fall into overfitting, so as to ignore the robustness characteristics of the image. Median filtering can make the image blurred, thus masking the details of the image, making the processed image retain more robust features.

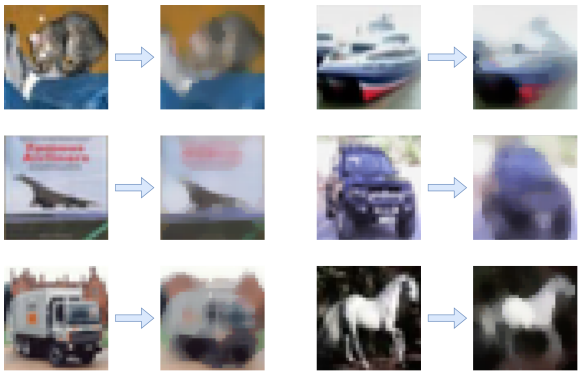

Figure 1: Examples of median filtering. We can see that the details of the picture are blurred.

.

Based on the above reasons, we first use a median filter to process the image. The processed image is shown in Figure 1. We can see that the objects in the image only retain the approximate contours, which will provide robust feature information for the model to be trained.

Adding Noise

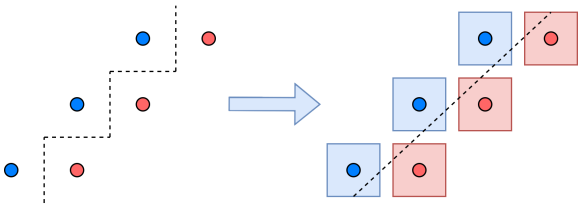

Figure 2: Change of classification boundary

.

In the related research of data robustness, increasing noise is often beneficial to improve robustness (Xie and Li 2019; Chan, Ho, and Nikolova 2005). Using appropriate noise can smooth the classification boundary, which will be beneficial to improve the robustness of the classifier. Because different noises will produce different effects, we use a variety of different noises such as gaussian noise, s&p noise and localvar noise.

The effect of noise is shown in Figure 3. Noise will not change the overall sense of the image but can change the local texture of the image, which can also prevent the model from falling into overfitting the local information.

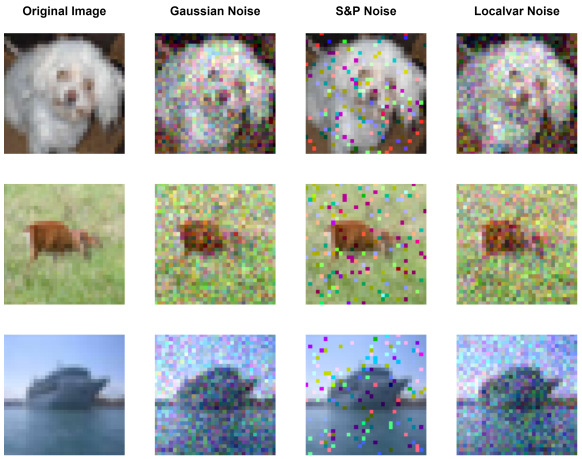

Figure 3: Some pictures with different noises

.

Data Augment

Data augment can improve the generalization ability of the classifier. Adjusting the brightness, hue and contrast of the image can change the style of the image. When the classifier learns more images of different styles, the classification accuracy of new images will also increase.

Figure 4: The image produced by adjusting brightness, hue and contrast

.

As shown in Figure 4, the images with data augment have different effects. They have different styles, but they won't let people misjudge them as other kinds of images. This effect will be beneficial to improve the generalization of the model.

We also use random cropping and random flipping in the process of training. Through data augment, the number of training samples will be greatly expanded, so as to improve the performance of the classifier.

Combination

In order to integrate the advantages of various data processing, we will process the training samples in combination with various parameters and random probability when preparing the data set. We use the median filter with a blur radius of 3 and 5, and add different noise to the image with different probabilities. For data augment, we select a part to process during training, which can reduce the number of samples in the training set. Some examples are shown in Figure 5.

Figure 5: Some sample pictures combining multiple data processing methods

.

Training Method And Experiment

Training Method

Appropriate training methods can make the network play a better performance. We use an SGD optimizer to optimize the network, compared with AdamW, SGD can better prevent overfitting. CosineAnnealingLR can make the learning rate decline more smoothly. At the same time, selecting a smaller batch size can make the network update more widely at each step, so that the network can converge to a relatively flat area more easily.

We trained 200 epochs for each model on a NVIDIA Titan V, and we took the weight saved in the last round of the model as the test weight. To verify the robustness of the data, we selected different network structures for testing in different stages of the experiment.

Experiment

Our experiment is based on CIFAR10 and is divided into two parts. Both parts are tested by an independent test set based on CIFAR10. In order to evaluate the quality of the data set, we use the average result of different networks as the evaluation of the corresponding data set in each part of the experiment.

In the first part of the experiment, we will test the anti-interference ability of the data. The training set and test set we selected are generated from the same data set. We added a variety of different noise and data processing methods to the test set and generated more than ten times the original data. In order to ensure the quality of the training set, we limit the number of samples in the training set to less than 5 times the number of original samples.

Metheds	Score
Baseline	72.34
Ordinary Augment	87.54
Gaussian Noise	96.77
Multiple Noise	97.75
Our Method	**98.46**

Table 1: Performance of anti-interference ability

We use Resnet50 and Densenet121 as the test models in the first stage (He et al. 2016a; Huang et al. 2017). We compare the effects of other conventional data augment or adding noise alone. The results are shown in Table 1, our method has achieved the best results. We find that compared with ordinary data augment, noise is more effective in increasing anti-interference ability. Based on combining different noises, our method further improves the performance of the classifier.

In the second stage of experiments, we verify the robustness of the model trained based on the data set in the face of new scenes. We used a public data set as the test set. These data have the same category as CIFAR10, but have not appeared in CIFAR10 data set. Similarly, these data have been added noise and processed.

Metheds	Score
One batch only	69.89
All train set	81.11
Gaussian Noise	81.99
Multiple Noise	81.99
Noise Plus Median Filter	82.64
Our Method	**84.04**

Table 2: Robustness in new scenarios

The training data for our second stage is the training set of cifar10. We also apply a variety of data processing methods to generate training sets and limit the number of training sets to 50000 images. At this stage, we use wideresnet and preactresnet18 as test models (Zagoruyko and Komodakis 2016; He et al. 2016b). The test results are shown in Table 2, our method shows the best robustness.

Compared with the experimental results of the first stage, we find that more data is more meaningful to improve the robustness of the model. Noise and median filtering can still increase the performance of the classifier, but to further improve the performance, the combination of multiple data processing methods will be more effective.

Conclusion

Constructing a general and effective data set to train robust models is of great significance to robust learning. In this paper, we proposed a method to make the data set more effective by processing data. We validate our method based on CIFAR10. As the experimental results show, median filter, adding noise and some data augment are all useful for improving data robustness. Our method improves the anti-interference ability and robustness of the classifier. For future work, our method can also be combined with better models to achieve better performance.

Acknowledgments

We thank the security AI challenger program launched by Alibaba Group and Tsinghua University.

References

Canziani, A.; Paszke, A.; and Culurciello, E. 2016. An analysis of deep neural network models for practical applications. *arXiv preprint arXiv:1605.07678*.

Chan, R. H.; Ho, C.-W.; and Nikolova, M. 2005. Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. *IEEE Transactions on image processing*, 14(10): 1479–1485.

Dabouei, A.; Soleymani, S.; Taherkhani, F.; and Nasrabadi, N. M. 2021. Supermix: Supervising the mixing data augmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 13794–13803.

Deng, L. 2012. The mnist database of handwritten digit images for machine learning research [best of the web]. *IEEE signal processing magazine*, 29(6): 141–142.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016a. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016b. Identity mappings in deep residual networks. In *European conference on computer vision*, 630–645. Springer.

Hendrycks, D.; Lee, K.; and Mazeika, M. 2019. Using pre-training can improve model robustness and uncertainty. In *International Conference on Machine Learning*, 2712–2721. PMLR.

Hendrycks, D.; Mazeika, M.; Kadavath, S.; and Song, D. 2019. Using self-supervised learning can improve model robustness and uncertainty. *Advances in Neural Information Processing Systems*, 32.

Ho-Phuoc, T. 2018. CIFAR10 to compare visual recognition performance between deep neural networks and humans. *arXiv preprint arXiv:1811.07270*.

Huang, G.; Liu, Z.; Van Der Maaten, L.; and Weinberger, K. Q. 2017. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4700–4708.

Ilie, C.-A.; Popescu, M.; and Stefanescu, A. 2020. Robustness as Inherent Property of Datapoints. In *AISafety@ IJCAI*.

Jia, Z.; Fang, H.; and Zhang, W. 2021. Mbrs: Enhancing robustness of dnn-based watermarking by mini-batch of real and simulated jpeg compression. In *Proceedings of the 29th ACM International Conference on Multimedia*, 41–49.

Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.

Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; and Alsaadi, F. E. 2017. A survey of deep neural network architectures and their applications. *Neurocomputing*, 234: 11–26.

Tan, M.; and Le, Q. 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, 6105–6114. PMLR.

Xie, T.; and Li, Y. 2019. Adding Gaussian Noise to DeepFool for Robustness based on Perturbation Directionality. *Aust. J. Intell. Inf. Process. Syst.*, 16(3): 44–54.

Yang, F.; Chen, Z.; and Gangopadhyay, A. 2020. Using randomness to improve robustness of tree-based models against evasion attacks. *IEEE Transactions on Knowledge and Data Engineering*.

Yin, D.; Gontijo Lopes, R.; Shlens, J.; Cubuk, E. D.; and Gilmer, J. 2019. A fourier perspective on model robustness in computer vision. *Advances in Neural Information Processing Systems*, 32.

Yu, D.; Zhang, H.; Chen, W.; Yin, J.; and Liu, T.-Y. 2021. How Does Data Augmentation Affect Privacy in Machine Learning? In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 10746–10753.

Zagoruyko, S.; and Komodakis, N. 2016. Wide residual networks. *arXiv preprint arXiv:1605.07146*.

Zhong, Z.; Zheng, L.; Kang, G.; Li, S.; and Yang, Y. 2020. Random erasing data augmentation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, 13001–13008.