(CVPR 2020 AliProducts Challenge) Technical Report

Mingliang Zhang[† ‡], Baole Wei[* ‡], and Yirong Yang[† ‡]

[†]Institute of Automation, Chinese Academy of Sciences
[*]Institute of Information Engineering, Chinese Academy of Sciences
[‡]University of Chinese Academy of Sciences,
Email: {zhangmingliang2018, yangyirong2018}@ia.ac.cn, weibaole@iie.ac.cn

Abstract

In this competition, we have to deal with some challenges which often appears in e-commerce industry, that are classes imbalanced dataset, labels are not fully reliable and very large number of classes. In order to achieve better performance, we solve the imbalanced data problem by limiting the number of samples in the head, adding regularization terms, and using rebalanced data to train the backbone and the classifier respectively. Meanwhile, the regularization terms can relieve label noisy problem. Furthermore, training model with different r-balanced samplers in different stages gets better representation and calssifier. Our model wins the 6th place and the mean top-1 error is 0.0995. The corresponding code is public available now at `https://github.com/ZhangMingliangAI/AiProducts-Challenge`.

1. Main Challenges

- There are 50030 different classes. That will reasult in a very big fully connected layer with large memory consumption.
- The number of samples in different classes ranges from 5 to 8k. The imbalanced ratio gets 1/1600 between the minimum class and the maximum class. That is the core matter of long tailed data tasks.
- Some images have incorrect labels.
- The difference between the class level is small. There is no obvious relatedness in the same class level. The hierarchical structures of classes are hard to use.

Figure 1. Some instances in AiProducts which often appears in e-commerce industry. It is a class imbalanced dataset with very large number of classes, and labels are not fully reliable. More details please see `https://tianchi.aliyun.com/competition/entrance/231780/introduction`

2. Methods

2.1. Limiting sample number in the head.

Paper[1] has pointed out that very large classes always take up very little proportion compared with the whole classes in the long tailed tasks. Undersampling and oversampling are two main methods to solve it. The former is a better choice since it can reduce the training set scale and shorten training time. Paper[2] proposes the concept of valid samples and it holds the views that very large classes exist a lot of redundancies.
Under the premise of maintaining the expression power of large classes, we undersample these classes by randomly discarding some samples and keeping the maximum number of each class not exceed 250. This strategy can relieve

class imbalanced problem and speed up training process.

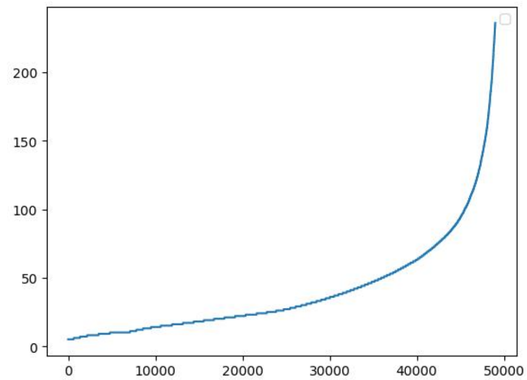

Figure 2. The class distribution after limiting sample in the head for narrowing the gap of the instance number between different categories.

2.2. Adding regularization terms.

Paper[4] has mensioned that we need to pay more attention to samples with little training loss since these data are more likely to be clean data. However, when dealing with imbalanced data, we should focus more on samples who has large training loss because these data are more likely to be in the tail classes. It is intractable to reweight different classes if the dataset is noisy and imbalanced.
Unfortunately, this competition offers this kind of data. We improve the generalization ability and reduce the noisy data's influences by adding L2 regularization terms.

2.3. Training the model with different re-balanced samplers

Paper[6] proposes that the instances rebalanced strategy can contribute to better representation learning while lead to worse classifier, while classes rebalance strategy has the opposite influences. Model[3] decouples the representation learning and the classifier to get better recognition performance. Specifically, they first train the whole model and after that froze the backbone and train the classifier or normalize the weight vectors of classifier.
We use the similar methods as papers[6][3], that is learning with the **instances re-balanced sampler** to get a better representation, and then finetuning using the **classes re-balanced sampler** with a lower learning rate to get better calssifier. Experiments demostrates that the improve performance margin is greater than the methods in paper [6][3].

3. Training Details

- Preprocess: Randomly throw some training samples in the head of classes to make sure the number of samples in each class is no more than 250.
- Backbone: We use Resnest-50[5]. It's a very efficient backbone which performs better than resnet-50 with less parameters and computation.
- Data augmentation: Some classical data augmentation methods like crop, resize, flip, normalize, etc.
- Regularization: L2 regularization with 1e-4.
- Optimizer: SGD with multi-step.
- Classifier(head): Fully connected layers.
- Loss function: Softmax + CrossEntropy.
- BatchSize: 512.
- Resolution rate: 112×112.
- Post-processing procedures: None.

4. Conclusion

This model uses some core strategies to relieve class imbalanced problem that (i) limiting sample number in the head, (ii) adding regularization terms, (iii) training the model with different rebalanced samplers in different stages. Deeper backbone, larger batchsize and higher resolution will further improve performance in this recognition task if hardware conditions permit. To the end, our model wins the 6th place and the mean top-1 error is 0.0995.

5. Acknowledgment

Thanks for CVPR organizing Committee and Aliyun TianChi platform to organize this competition, which provides the precious opportunity to train ourselves. More importantly, thanks for cooperation and hard work of everyone in our team.

6. References

References

[1] Mateusz Buda, Atsuto Maki, and Maciej A Mazurowski. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, 106:249–259, 2018.

[2] Yin Cui, Menglin Jia, Tsung-Yi Lin, Yang Song, and Serge Belongie. Class-balanced loss based on effective number of samples. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9268–9277, 2019.

[3] Bingyi Kang, Saining Xie, Marcus Rohrbach, Zhicheng Yan, Albert Gordo, Jiashi Feng, and Yannis Kalantidis. Decoupling representation and classifier for long-tailed recognition. *arXiv preprint arXiv:1910.09217*, 2019.

[4] Mengye Ren, Wenyuan Zeng, Bin Yang, and Raquel Urtasun. Learning to reweight examples for robust deep learning. *arXiv preprint arXiv:1803.09050*, 2018.

[5] Hang Zhang, Chongruo Wu, Zhongyue Zhang, Yi Zhu, Zhi Zhang, Haibin Lin, Yue Sun, Tong He, Jonas Mueller, R Manmatha, et al. Resnest: Split-attention networks. *arXiv preprint arXiv:2004.08955*, 2020.

[6] Boyan Zhou, Quan Cui, Xiu-Shen Wei, and Zhao-Min Chen. Bbn: Bilateral-branch network with cumulative learning for long-tailed visual recognition. *arXiv preprint arXiv:1912.02413*, 2019.