Machine Learning Paper Review (ECE57000 Fall 2022)

Anonymous Authors[1]

Abstract

This paper will be summarizing 3 machine learning papers related to modifying conventional training and testing data for image classification and object detection machine learning models to yield better results and higher model accuracy. For every paper review there will be a summary of the key points of the paper, then following by a critical review.

1. Between class learning for image classification

1.1. Summary

This paper was accepted to the 2018 CVPR conference and discusses their novel machine learning method called Between-Class learning. The main focus of their method was on image classification, using image-classification models to test the accuracy and performance of their method.

With this, the researcher tried two methods of mixing, one slightly more complicated than the other. The researchers found that their Between-Class machine learning method achieved 19.4% and 2.26% top-1 errors on ImageNet-1K and CIFAR-10 respectively (Tokozume, 2018). This is important as it shows that we can further generate more meaningful test data to better train our models.

The main motivation of their paper was the idea of how humans can identify different sounds of different pitches despite them being superimposed on one another. The researchers thought that if we can abstract information from what seems like nonsense if we fed the sound to a machine, they thought that mixing two images together, which is meaningless data to us, may yield results in a machine as they may be able to abstract the two images from one another.

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

Similar to the concept of sound, the researchers treat the input image as a sound wave and mix two images together, and instead of having the model predict the image, the researchers want to model to predict the mix ratio of the two images.

1.2. Critical review

1.2.1. Strengths

The strengths of the paper were that it cited and used a lot of related work. They based their between-class model on a paper (Dai, 2017) that figured out the mathematics of converting waveforms to an input vector by incorporating the ”energy” of the waveform. This gives the paper a lot of credibilities as they worked on a previously accepted mathematical operation to transform their image inputs.

Furthermore, the paper shows a lot of statistics with community-accepted image classification datasets and models. The researchers used large datasets such as ImageNet-1K, CIFAR-100, and CIFAR-10. The researchers used large image classification models such as ResNeXt-29, ResNet-29, DenseNet, and Shake-Shake Regularization to make sure that their Between-Class learning method worked.

Lastly, I believe that it was interesting that the researchers included comparisons between the BC and BC+ model, showing us how a slight optimization made the error rate decrease.

1.2.2. Weakness

The weakness associated with this paper is that the models have not gone through a large number of epochs or trials to be determined consistent. The highest epoch that the researchers conducted was on the ImageNet-1K, having 150 epochs. Furthermore, I thought that their way of mixing the labels and finding the ratio was unconvincing as it appeared to only be limited to classifying two classes at a time due to the ratio, which is not applicable to large datasets with 1000+ classes.

Lastly, the researchers claimed to have imposed a constraint on the feature distribution by simply giving us a 3D Data Visualization of the data points after PCA dimensionality reduction without giving us information about the variance, mean etc...

2. Human uncertainty makes classification more robust

2.1. Summary

This paper was accepted to the CVPR 2019 conference. This paper explores the effectiveness and accuracy of adding human uncertainty into an image classification dataset. The researchers modified a popular community-accepted dataset CIFAR10 and added human uncertainty soft labels to the dataset, calling it CIFAR10H. The CIFAR10H dataset itself consists of 10,000 images with 511,400 human categorization decisions (Peterson, 2019).

The researchers discovered that modifying their dataset to be soft-labeled with human uncertainty instead of one hot encoding, it has shown the following improvements to model training:

- The dataset improved the generalizability of the model as compared to if the model was trained with hard labels.
- lower softmax errors whenever the model makes an incorrect prediction of an image class.
- Model was significantly more resistant to adversarial attacks.

2.2. Critical review

2.2.1. Strength

The strengths of the paper were that the researchers were very comprehensive. they utilized 5 image classification datasets (CIFAR10, CIFAR10 .1v6, v4, CINIC10, ImageNet-Far) other than their novel CIFAR10H. This makes their research a lot more reliable and factual as it compared their results with widely accepted datasets.

Moreover, the researchers provided a lot of informative graphs, tables, and statistics to back up their claims on the training accuracy and error of the training results of their paper. I also really appreciated how the researchers delved into the topic of whether or not the model was learning the human uncertainty when given the modified dataset or if the model still utilizes the information that is present in the image.

Lastly, a final strength of the paper was that they detailed every step of the making of their dataset, breaking down the costs, manpower, and time taken to create the dataset. Moreover, the dataset itself is rich in information, having more than half a million human categorization decisions.

I believe that this paper does discover a novel idea of dataset image labeling as their results do show promise. By training ImageNet-Far with CIFAR10H, its cross-entropy error rate was reduced by 38% on average as compared to training the model with the original CIFAR10 (Peterson, 2019).

2.2.2. Weakness

A weakness of this paper is that the researchers did not fully back up their claim on the robustness of adversarial attacks. The researchers did provide a graph showing that the cross-entropy loss of soft labels is a lot less than hard labels during an adversarial attack, but that is the only evidence, that I felt was a bit lackluster.

Another point that I would like to point out is that the researchers brought in a lot of calculations and assumptions without elaboration on how or where these methods came to be or were derived from.

3. Between class learning for image classification

3.1. Summary

This paper was accepted to the 2019 CVPR conference and discusses their solution to solve loss based on class imbalances in image classification datasets. The issue with long-tailed datasets is that there is an imbalance of training data for different classes, resulting in poor predictions for classes with a smaller sample of training data. Moreover, there also is a diminishing return to classes with a large sample of training data as it leads to waste in computation and issues with overfitting.

Numerous studies (Bengio, 2015; Ouyang, 2016; Huang, 2016) have tried to solve this issue through cost-sensitive re-weighting, which is to set a hyper-parameter on the loss function based on the ratio between the number of training data for a certain class and the total size of the training set. However, the paper argues that this is not generalizable and in some cases hurts the model by overfitting.

The researchers designed an effective re-weighting scheme that uses the effective samples for each class to re-balance the loss to calculate a class-balanced loss. (Yin Ciu, 2019) They trained their model on artificial and naturally long-tailed datasets on CNN models and found that the network is able to achieve significant performance gains as compared to without class re-balancing.

3.2. Critical Review

3.2.1. Strength

What I believed was a big strength of this paper was the number of references it had to other papers that tackled a similar problem. The researchers found similar results and referenced those papers. The paper also explained their motivation and methods very clearly, every step was very concrete and had rich references as to why they chose to set up their equations the way they are without bombarding us with unappetizing mathematics.

In terms of their experiment, the researchers used a wide range of widely popular image classification datasets such as iNaturalist 2017, 2018, and ImageNet data (ILSVRC 2012). The researchers also modified CIFAR10 and CIFAR100 to skew the image distributions such that it was tailed to mimic real-world data. The wide range of datasets used in the experiment debiases the experiment and prevents the conclusion that the results are only applicable to a specific dataset.

3.2.2. Weakness

The weaknesses of this paper would be the conclusion of their experiment. I believe that the paper set up a great theory of their class-based solution and came up with great solutions such as the modification of focal loss, sigmoid loss, and cross-entropy loss. But the results of the experiment showed little to no data backing up that their experiment worked. The researchers showed a few graphs with the error rate of each model with a specific class-based loss function. However, the researchers failed to elaborate what the conditions were in these experiments, lacking information such as epochs, optimizers, folds etc...

Not a weakness but rather an observation would be that the researchers conclude by declaiming their original claim that ratio-based loss functions are inefficient, contradicting it. This is because they found that if they set their hyperparameter to 0.99, that is when the class-based re-balancer works the best. However, 0.99 is close, if not the same as the ratio-based rebalancing method.

References

Yuji Tokozume. Between-class learning for image classification, 2018. URL https://openaccess.thecvf.com/content_cvpr_2018/papers/Tokozume_Between-Class_Learning_for_CVPR_2018_paper.pdf.

W. Dai. Very deep convolutional neural networks for raw waveforms., 2017. URL https://arxiv.org/pdf/1610.00087.pdf.

Joshua C. Peterson. Human uncertainty makes classification more robust, 2019. URL https://openaccess.thecvf.com/content_ICCV_2019/papers/Peterson_Human_Uncertainty_Makes_Classification_More_Robust_ICCV_2019_paper.pdf.

Samy Bengio. Sharing representations for long tail computer vision problems, 2015. URL https://arxiv.org/pdf/2103.16370.pdf.

Wanli Ouyang. Factors in finetuning deep model for object detection with long-tail distribution, 2016. URL https://openaccess.thecvf.com/content_cvpr_2016/papers/Ouyang_Factors_in_Finetuning_CVPR_2016_paper.pdf.

Chen Huang. Learning deep representation for imbalanced classification., 2016. URL https://openaccess.thecvf.com/content_cvpr_2016/papers/Huang_Learning_Deep_Representation_CVPR_2016_paper.pdf.