Aeden Jameson
9810912
EE 596 Autumn 2021
Homework 3

(1) Motivation – Why do the authors want to work on this problem

The authors want to establish a set of scientifically backed guidelines for adjustments made to training procedures, called tricks, for long tailed visual recognition problems (e.g. changes in the data distribution, re-weighting)

(2) Contributions – What are the accomplishments they achieved in this paper (others did not achieve)?

The authors major contributions are

- A novel CAM-based sampling approach tailored for two-stage training[1]
- They comprehensively explore long-tailed related tricks and identify optimal combinations that outperform state-of-the-art methods.

(3) Formulations – How do they solve the problems as mentioned/discussed in the introduction or related literature?

The authors systematically evaluated the known tricks on long tailed versions of CIRAR-10, CIRAR-100 and iNaturalist datasets. They reduced the number of training samples per class according to an exponential function $n = nt \times \mu t$ where t is the class index (0-indexed) and nt is the original number of training images with $\mu \in (0, 1)$[1] And the test set remained the same. Additionally the authors adopt ResNet-32 and ResNet-50 backbones for the CIFAR and iNaturalist datasets. Furthermore, the authors adopted state of the art data augmentation, initialization, and hyper parameters settings for training.

From there the authors classify tricks into four families: reweighting, re-sampling, mixup training, and two-stage training[1] and systemically find the optimal combination of tricks that produce the best results.

(4) Justification – Do the experiments/simulations support their claimed accomplishments?

From a scientific perspective it would appear their results are justified given they implement their change/experiments on established architectures. Thus they are controlling their experiments.

(5) Your Own Thoughts – What are you most impressed with in this paper?

The ability to know how to *systematically* go about improving the training for a given network on long tailed visual recognition problems is an impressive contribution. And as an aside after finishing problem three on domain adaptation I'm very curious about whether that technique won't prove useful in solving long tail distribution problems.

References

[1] Yongshun Zhang, Xiu-Shen Wei, Boyan Zhou, Jianxin Wu. Bag of Tricks for Long-Tailed Visual Recognition with Deep Convolutional Neural Networks