Work Report 26/6/2021

Anshul Singh

Today I studies about optimization technique in yolo v4 specillay to increase the mean average precision of the model. These optimization on mAP require data augumenation like stuffs that’s why I am not retraining.

Some techniques that can be done for increasing the mAP of the yolo model are:

1. Data Augumentation technique:
2. **Visually Coherent Image Mix-up for Object Detection (+3.55% mAP Boost)**

The methodology has already been proven to be successful in lessening adversarial fears in network classification after testing it on COCO 2017 and PASCAL datasets with YOLOv3 models.

https://miro.medium.com/max/60/0*x9zbSrXVRab0dA0L?q=20



The difference here is that researchers introduce occlusions and spatial signal perturbations that are common in natural image presentation, in particular, the use of geometry preserved placements for image mix-up to avoid distorted images during initial training iterations. They also go for a beta distribution with more visually coherent ratios including a >= 1 and b >= 1 instead of following conventional image classification which achieves actual model performance improvements.



1. **Classification Head Label Smoothening (+2.16% mAP Boost)**

Existing models apply Softmax technique to compute a probability distribution for classes. But there’s a risk of the model becoming too confident in its predictions which can result to over-fitting. One possible solution to this is to relax our confidence on the labels. For instance, we can slightly lower the loss target values from 1 to, say, 0.9. And naturally, we increase the target value of 0 for the others slightly as such. This idea is called label smoothing. Consult this for more information.

<https://github.com/Kyubyong/label_smoothing/blob/master/Noisy%20Labels%20and%20Lable%20Smoothing.ipynb>

1. **Data Pre-processing (Mixed Results)**

For object detection preprocessing it is critical to take extra caution guards as detection networks are sensitive to geometrical transformations. Some proven data augmentation methods include:

Random geometry transformation for random cropping (with constraints), random expansion, random horizontal flip and random resize (with random interpolation).

Random color jittering for brightness, hue, saturation, and contrast

1. **Training Scheduler Revamping (+1.44% mAP Boost)**

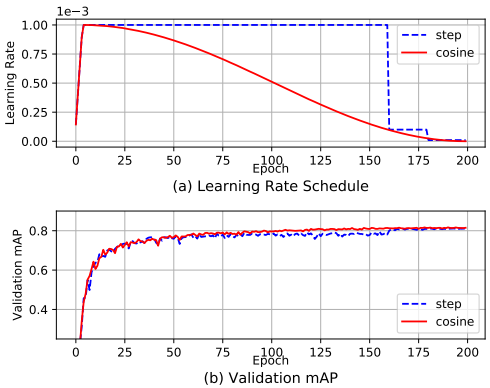
In model training, step scheduler is the most widely used learning rate schedule. It involves multiplying the learning rate by a constant number below 1 after a number of model iterations. For a Faster-RCNN, for example, the default step learning rate schedule aims at reducing the learning rate by 0.1 ratio at 60k iterations. Likewise, YOLOv3 uses the same ratio to lessen learning rate at 40k and 45k iterations.

The downside of the step scheduler in state-of-the-art object detectors is a sharp learning rate transition which may result in optimizer re-stabilization for subsequent iterations.

A better approach would involve training with:

**Cosine scheduler** — scales the learning rate according to the value of cosine function on 0 to pi. It starts by slowly reducing the learning rate, and quickly reducing it halfway to finally achieve a tiny slope which further reduces the learning rate to 0.

**Warm up scheduler** — aims to avoid gradient explosion during the initial model training iterations.



Cosine scheduler suffers less from plateau phenomenon and has been proven to outperform step scheduler. Applying both of these schedulers can help you achieve much better validation accuracy.

1. **Synchronized Batch Normalization (+0.56% mAP Boost)**

It’s true that batch normalization implementation on multiple devices (GPUs) is fast and doesn’t increase communication overheads. But, it reduces the batch size and alters statistics during computation which hurts model performance. The solution for this lies in synchronized batch normalization. The evaluation of synchronized batch normalization with YOLOv3 has been done to show the impacts of comparatively smaller batch-size on GPUs.

1. **Random Shapes Training for Single-stage ODN (+0.98% mAP Boost)**

To curb memory limitations and enable simpler batching, many single-stage object detection are trained with fixed shapes. Natural images come in a variety of unfixed shapes. To deal with the problem of overfitting that can be caused by training with fixed images and to improve the generalization of network predictions, the best approach is to implement a random shapes model training.

Various other techniques are mentioned in the previous report to improve mAP.