Distilling Object Detectors via Decoupled Features (Supplementary Material)

Jianyuan Guo^{1,2}, Kai Han¹, Yunhe Wang¹, Han Wu², Xinghao Chen¹, Chunjing Xu¹, Chang Xu^{2*}

¹ Noah's Ark Lab, Huawei Technologies.

² School of Computer Science, Faculty of Engineering, University of Sydney.

{jianyuan.guo, kai.han, yunhe.wang}@huawei.com; c.xu@sydney.edu.au



Figure 1: Per-channel distance between teacher and student detector. (a) Student is trained without KD technique; (b) Student is distilled via object only regions in neck features; (c) Student is distilled via background only regions in neck features; (d) Student is distilled via all regions in neck features; (e) Student is distilled via decoupled regions (DeFeat) in neck features.

In this supplementary material, we show more cases about (i) the L_2 -Norm of the gradient in neck feature during back propagation, where the student is distilled via treating all regions uniformly, and (ii) the PR curve and error analysis of different distillation methods on COCO minival.

1. Distance between teacher and student.

We calculate the per-channel distance between teacher and student to better understand why decoupled regions can attain higher performance compared to its counterparts. We randomly choose 20 images from COCO minival and calculate the absolute value of the discrepancy between the same region in neck features from teacher and student. As shown in Figure 1, the teacher and the student are ResNet152-FPN and ResNet50-FPN, respectively. X axis stands for each channel (256 in total) from the first layer of FPN. Y axis stands for the average distance of pixels belonging to object/background regions per channel. Distilling student detector via object regions in neck features can substantially narrow the distances in object regions, but fail to promote the background regions. Distilling student object via treating all regions equally can decrease the distance by a fair margin. Our DeFeat could further minimize the distance to the largest extent for both object and background regions.

2. Visualization of the gradient

We randomly select images from COCO training set, and the student detector has been distilled by 2 epochs via treating all regions equally. We use the hook function in Pytorch [1] to get gradients during back prorogation and show them in Figure 2 and Figure 3. The magnitude of gradients from object regions are consistently larger than that from background regions, which reminds us of assigning different importance to object regions and background regions during the distillation.

3. PR curve and error analysis

Figure 4-9 present the error analyses on six categories of COCO benchmark. For example, the first row of Figure 3 denotes the result of baseline detector trained without distillation technique; the second row denotes the result of student detector distilled via object only regions; the third row denotes the result of student detector distilled via back-ground only regions. In addition, the student distilled via background regions achieve the same performance, *i.e.*, 39.9% mAP on COCO minival. We can find that in most cases, distillation via background only regions can effectively decrease the background false positive rate compared to distillation via object only regions.

References

[1] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. In Advances in neural information processing systems, pages 8026– 8037, 2019. 1



Figure 2. L_2 -Norm of the gradient in neck feature during back propagation, the darkest blue indicates the largest norm value.



Figure 3. L_2 -Norm of the gradient in neck feature during back propagation, the darkest blue indicates the largest norm value.



Figure 4. Precision-recall curves and error analyses of category "apple" on COCO minival.



Figure 5. Precision-recall curves and error analyses of category "backpack" on COCO minival.



Figure 6. Precision-recall curves and error analyses of category "chair" on COCO minival.



Figure 7. Precision-recall curves and error analyses of category "refrigerator" on COCO minival.



Figure 8. Precision-recall curves and error analyses of category "spoon" on COCO minival.



Figure 9. Precision-recall curves and error analyses of category "skis" on COCO minival.