A Appendix

A.1 More Ablations and Visualizations

Effect of Blocking Gradient of $f(s_j - s_i; \theta_4)$. As mentioned in Section 3.2, we compare the performance of different detectors with or without blocking the gradient of $f(s_j - s_i; \theta_4)$ on the COCO benchmark [5] in Table 1. These results indicate that blocking the gradient of $f(s_j - s_i; \theta_4)$ can greatly boost the performance. We attribute this to the unstable training caused by the gradient from the denominator, so they are blocked out by default in the experiments.

Visualization of Searched Parameterized Functions. Figure 1 visualizes the searched parameterized functions for different detectors on the COCO benchmark [5]. Each line corresponds to an independent parameterized function in Eq. (6). The dots on each line represent the control points for each parameterized function. It can be observed that loss functions for different detectors seem to differ from each other. Their intrinsic differences can lead to distinct loss functions.

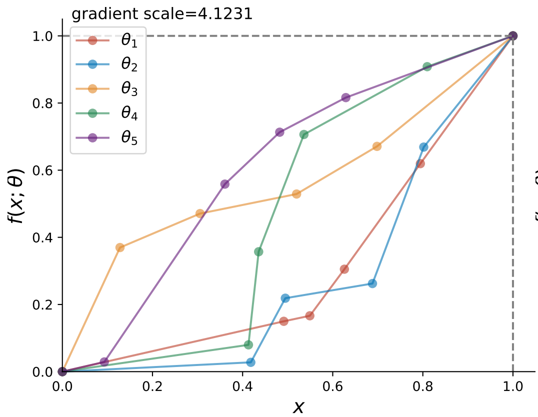
Parameterized AP Loss on PASCAL VOC Benchmark [3]. We also search Parameterized AP Loss on the PASCAL VOC beanchmark and compare its performance with the commonly used Focal Loss [7] and L1 combination. We adopt RetinaNet [7] with ImageNet [2] pre-trained ResNet50 [4] and FPN [6] as the backbone. The training setting strictly follows the default config in MMDetection codebase [1]. During search, we train the object detector for one epoch as the proxy task. Table 2 shows that the searched loss can perform well on the PASCAL VOC benchmark, bringing around 3.0 AP_{50} improvement.

Table 1: The effect of blocking gradient of $f(s_j - s_i; \theta_4)$ on the COCO benchmark.

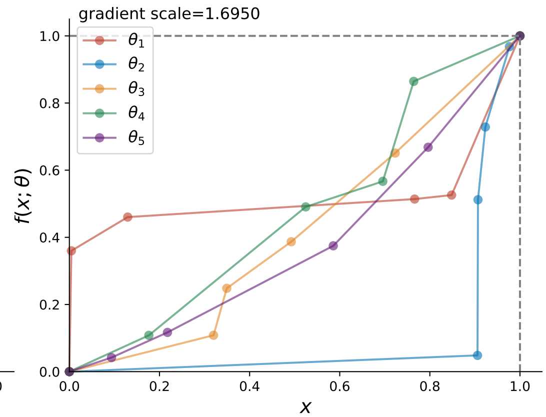
Model	Block Gradient	AP	AP_{50}	AP_{75}	AP_{S}	AP_{M}	AP_{L}
RetinaNet [7]		0.8	1.3	0.9	0.5	0.9	1.2
ResNet-50 [4] + FPN [6]	✓	40.5	59.0	43.4	23.9	44.9	56.1
Faster R-CNN [8]		29.5	42.4	31.4	14.4	32.3	42.6
ResNet-50 [4] + FPN [6]	✓	42.0	60.7	45.0	25.3	46.6	57.7
Deformable DETR [9]		27.8	54.9	25.8	16.8	31.5	35.6
ResNet-50 [4]	✓	45.3	63.1	49.6	27.9	49.3	60.2

Table 2: Comparison on the PASCAL VOC benchmark with RetinaNet.

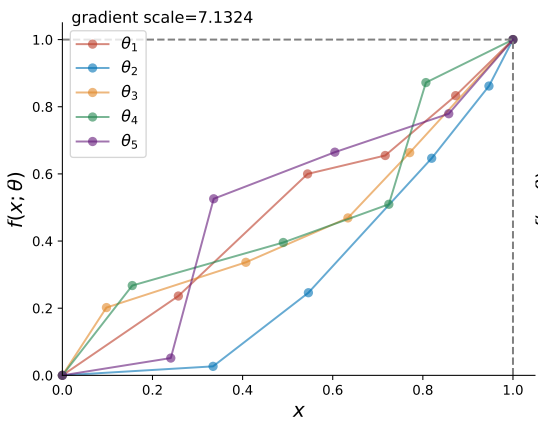
Loss	AP_{50}
Focal Loss [7] + L1	77.3
Parameterized AP Loss	80.2



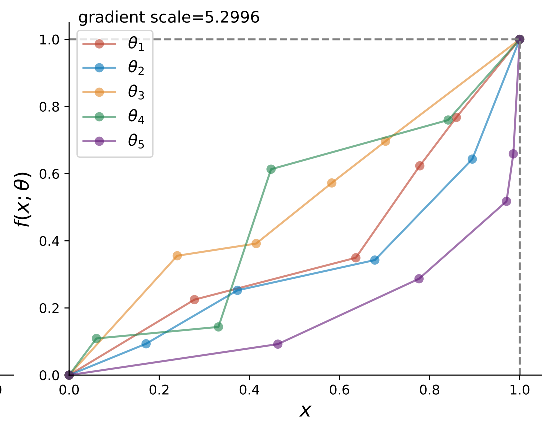
(a) RetinaNet



(b) RPN head in Faster R-CNN



(c) R-CNN head in Faster R-CNN



(d) Deformable DETR

Figure 1: Visualization of the searched parameterized functions on the COCO benchmark.

References

[1] K. Chen, J. Wang, J. Pang, Y. Cao, Y. Xiong, X. Li, S. Sun, W. Feng, Z. Liu, J. Xu, et al. Mmdetection: Open mmlab detection toolbox and benchmark. *arXiv preprint arXiv:1906.07155*, 2019.

[2] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *CVPR*, 2009.

[3] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman. The pascal visual object classes (voc) challenge. *IJCV*, 2010.

[4] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *CVPR*, 2016.

[5] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. Microsoft coco: Common objects in context. In *ECCV*, 2014.

[6] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie. Feature pyramid networks for object detection. In *CVPR*, 2017.

[7] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár. Focal loss for dense object detection. In *ICCV*, 2017.

[8] S. Ren, K. He, R. B. Girshick, and J. Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *TPAMI*, 2015.

[9] X. Zhu, W. Su, L. Lu, B. Li, X. Wang, and J. Dai. Deformable detr: Deformable transformers for end-to-end object detection. *arXiv preprint arXiv:2010.04159*, 2020.