Project Proposal(CSI5139 2019 Fall)

Yuan Gao 300086076
Ao Peng 3000145650

October 6, 2019

1 Introduction

The topic of the paper we chose is Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression [1], which was released in CVPR 2019. This paper proposed a new loss in replace of bounding box regression MSE loss used in state-of-the art object detection algorithms [2] [3] [4]. By incorporating this generalized IoU (GIoU) as a loss into the state-of-the art object detection frameworks, the result shows a improvement on the performance.

2 Problem description

Intersection over Union (IoU) is the most popular evaluation metric used in the object detection benchmarks [5] [6]. However, there is a gap between optimizing the commonly used distance losses for regressing the parameters of a bounding box and maximizing this metric value. The optimal objective for a metric is the metric itself. However, IoU has a plateau making it infeasible to optimize in the case of non-overlapping bounding boxes. IoU as both a metric and a loss has a major issue: if two objects do not overlap, the IoU value will be zero and will not reflect how far the two shapes are from each other. In this case of non-overlapping objects, if IoU is used as a loss, its gradient will be zero and cannot be optimized.

This paper addressed the weaknesses of IoU by introducing a generalized version as both a new loss and a new metric. The authors addressed this weakness by extending the concept to non-overlapping cases. This generalization (a) follows the same definition as IoU, i.e. encoding the shape properties of the compared objects into the region property; (b) maintains the scale invariant property of IoU, and (c) ensures a strong correlation with IoU in the case of overlapping objects. This paper introduced this generalized version of IoU as a new metric for comparing any two arbitrary shapes. They also provided an analytical solution for calculating GIoU between two axis aligned rectangles, allowing it to be used as a loss in object detection tasks.

3 Project scope and objectives

- Pre-training a backbone(Darknet or Resnet series [7]) on Imagenet cls dataset [8].
- Implementing original YOLOv3 and training on VOC2007 [5] trainval dataset.
- Substituting bounding box regression loss with IOU and GIOU loss proposed by the paper and training on VOC2007 trainval dataset.
- Evaluating the resulst of 3 different losses.

References

[1] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized intersection over union: A metric and a loss for bounding box regression," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[2] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," in *arXiv preprint arXiv:1804.02767*, 2018.

[3] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds. Curran Associates, Inc., 2015, pp. 91–99. [Online]. Available: http://papers.nips.cc/paper/5638-faster-r-cnn-towards-real-time-object-detection-with-region-proposal-networks.pdf

[4] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask r-cnn," in *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[5] M. Everingham, L. Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *Int. J. Comput. Vision*, vol. 88, no. 2, pp. 303–338, Jun. 2010. [Online]. Available: http://dx.doi.org/10.1007/s11263-009-0275-4

[6] T.-Y. Lin, M. Maire, S. J. Belongie, L. D. Bourdev, R. B. Girshick, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context." in *ECCV*, vol. 8693, 2014.

[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[8] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *CVPR09*, 2009.