Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression

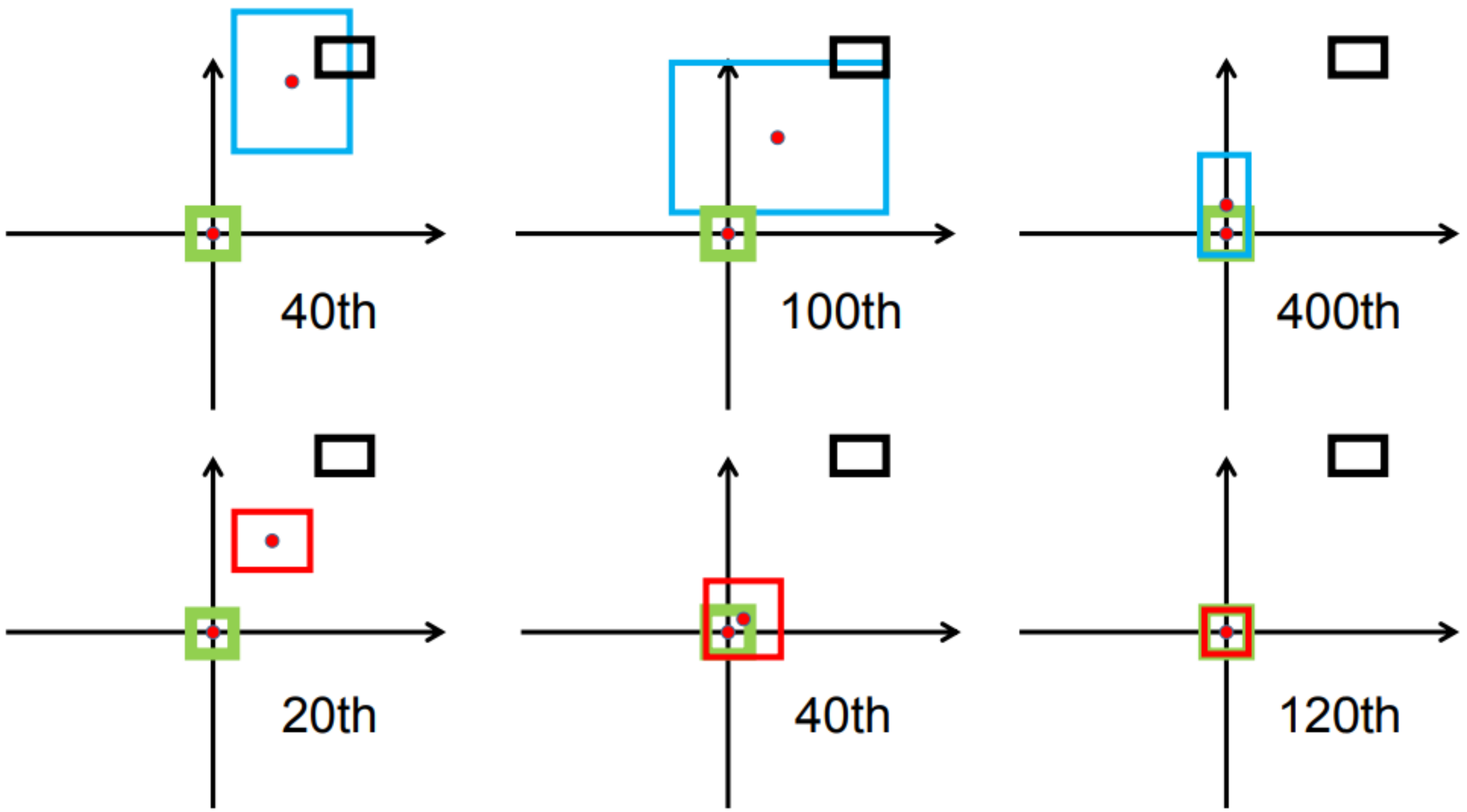
Paper Review

Sangho Kim

Contents

The limitation of GIoU

The limitations of GIoU



Contents

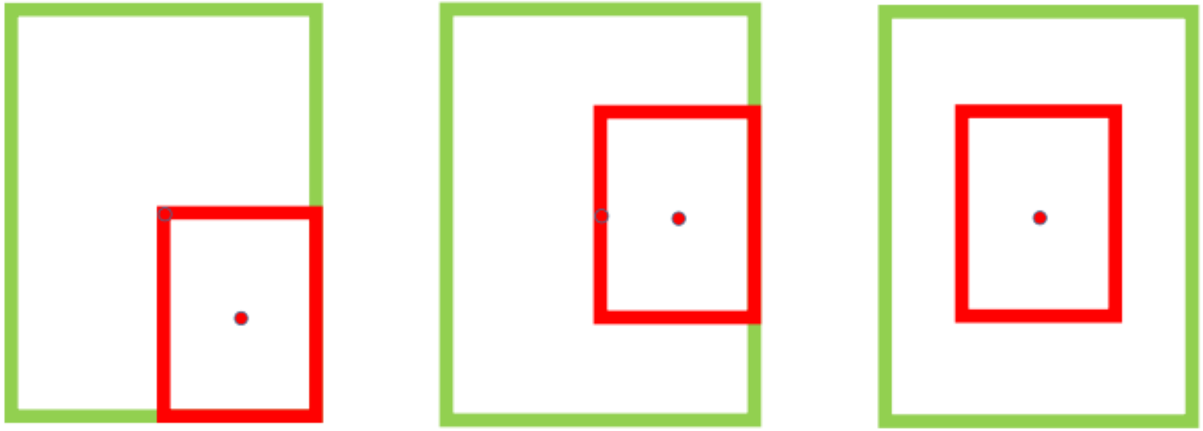
- The limitation of GIoU
- DIoU Loss
- CIoU Loss
- Experiments

Distance-IoU Loss

$$\mathcal{L}_{DIoU} = 1 - IoU + \frac{\rho^2(\mathbf{b}, \mathbf{b}^{gt})}{c^2}$$

- where $\boldsymbol{b}$ and $\boldsymbol{b}^{gt}$ denote the central points of B and B^{gt}, respectively,
- ρ is the Euclidean distance,
- and c is the diagonal length of the smallest enclosing box covering the two boxes

Distance-IoU Loss



$\mathcal{L}_{IoU} = 0.75$ $\mathcal{L}_{IoU} = 0.75$ $\mathcal{L}_{IoU} = 0.75$

$\mathcal{L}_{GIoU} = 0.75$ $\mathcal{L}_{GIoU} = 0.75$ $\mathcal{L}_{GIoU} = 0.75$

$\mathcal{L}_{DIoU} = 0.81$ $\mathcal{L}_{DIoU} = 0.77$ $\mathcal{L}_{DIoU} = 0.75$

Figure 2: GIoU loss degrades to IoU loss for these cases, while our DIoU loss is still distinguishable. Green and red denote target box and predicted box respectively.

Contents

- The limitation of GIoU
- DIoU Loss
- CIoU Loss
- Experiments

Complete IoU Loss

- A good loss for bbox regression should consider three important geometric factors, i.e., overlap area, central point distance and aspect ratio.
- Our proposed DIoU loss aims at considering the overlap area and central point distance of bboxes.
- However, the consistency of aspect ratios for bboxes is also an important factor.

Complete IoU Loss

- Therefore, based on DIoU loss, the CIoU loss is proposed by imposing the consistency of aspect ratio

$$\mathcal{L}_{CIoU} = 1 - IoU + \frac{\rho^2(\mathbf{b}, \mathbf{b}^{gt})}{c^2} + \alpha v.$$

$$v = \frac{4}{\pi^2}(arctan\frac{w^{gt}}{h^{gt}} - arctan\frac{w}{h})^2. \qquad \alpha = \frac{v}{(1 - IoU) + v}$$

- where α is positive trade-off parameter, and v measures the consistency of aspect ratio
- by which the overlap area factor is given higher priority for regression, especially for non-overlapping cases.

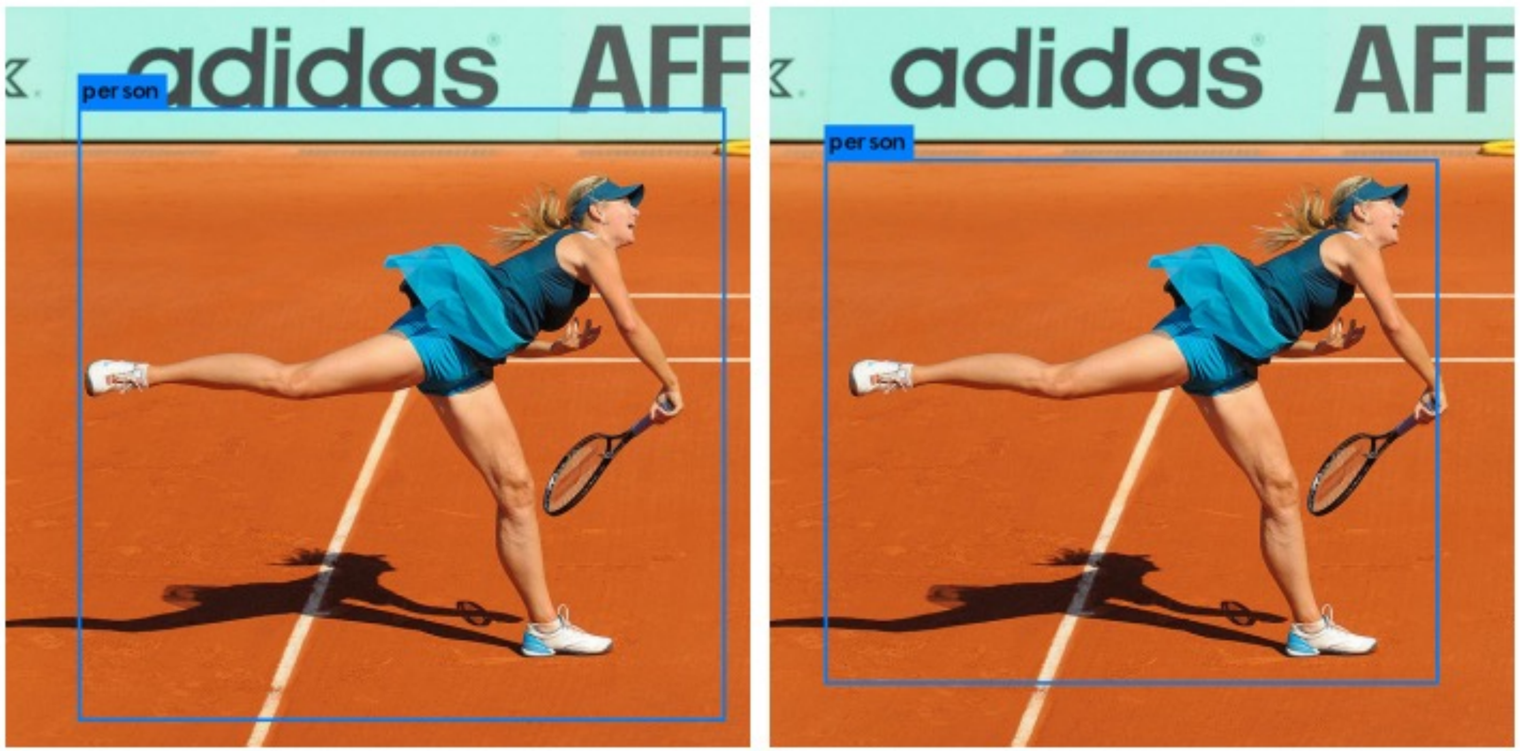
Contents

Compare with IoU losses

Table 1: Quantitative comparison of **YOLOv3** (Redmon and Farhadi 2018) trained using $\mathcal{L}_{IoU}$ (baseline), $\mathcal{L}_{GIoU}$, $\mathcal{L}_{DIoU}$ and $\mathcal{L}_{CIoU}$. (D) denotes using DIoU-NMS. The results are reported on the test set of PASCAL VOC 2007.

Loss / Evaluation	AP		AP75	
	IoU	**GIoU**	**IoU**	**GIoU**
$\mathcal{L}_{IoU}$	46.57	45.82	49.82	48.76
$\mathcal{L}_{GIoU}$	47.73	46.88	52.20	51.05
Relative improv. %	2.49%	2.31%	4.78%	4.70%
$\mathcal{L}_{DIoU}$	48.10	47.38	52.82	51.88
Relative improv. %	3.29%	3.40%	6.02%	6.40%
$\mathcal{L}_{CIoU}$	**49.21**	**48.42**	**54.28**	**52.87**
Relative improv. %	**5.67%**	**5.67%**	**8.95%**	**8.43%**
$\mathcal{L}_{CIoU}(D)$	**49.32**	**48.54**	**54.74**	**53.30**
Relative improv. %	**5.91%**	**5.94%**	**9.88%**	**9.31%**

Detection examples using YOLOv3

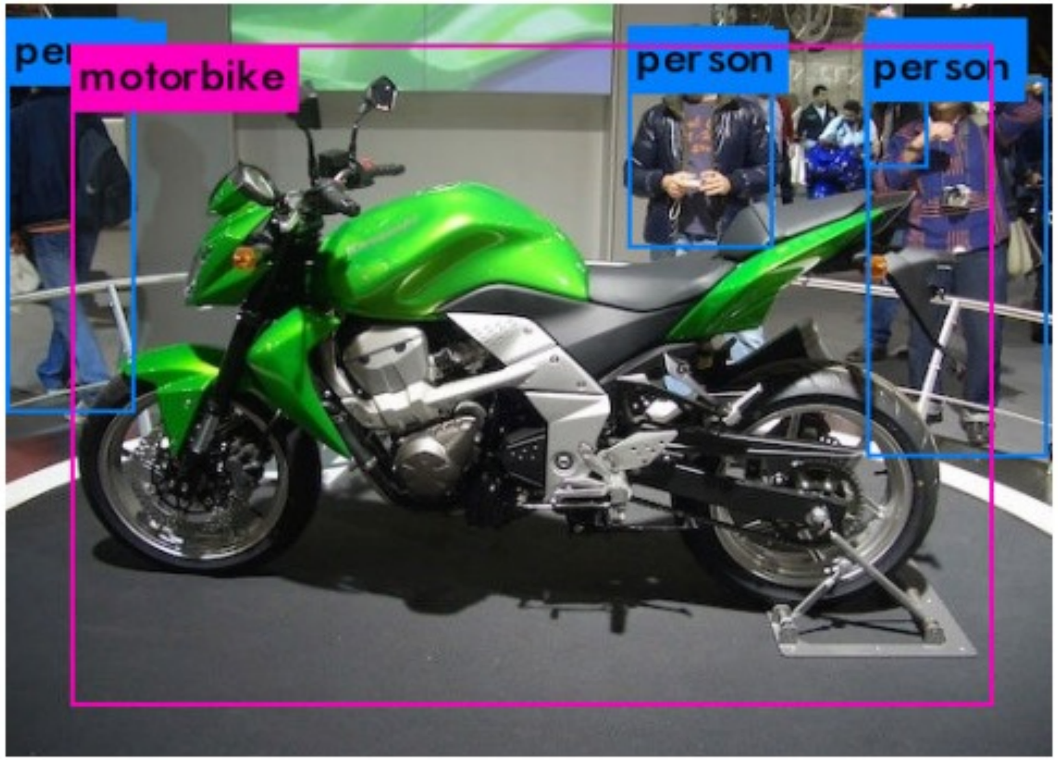


$\mathcal{L}_{GIoU}$ $\mathcal{L}_{CIoU}$

Detection examples using YOLOv3

$\mathcal{L}_{GIoU}$ $\mathcal{L}_{CIoU}$

Detection examples using YOLOv3



$\mathcal{L}_{GIoU}$

$\mathcal{L}_{CIoU}$