Learning Modulated Loss for Rotated Object Detection

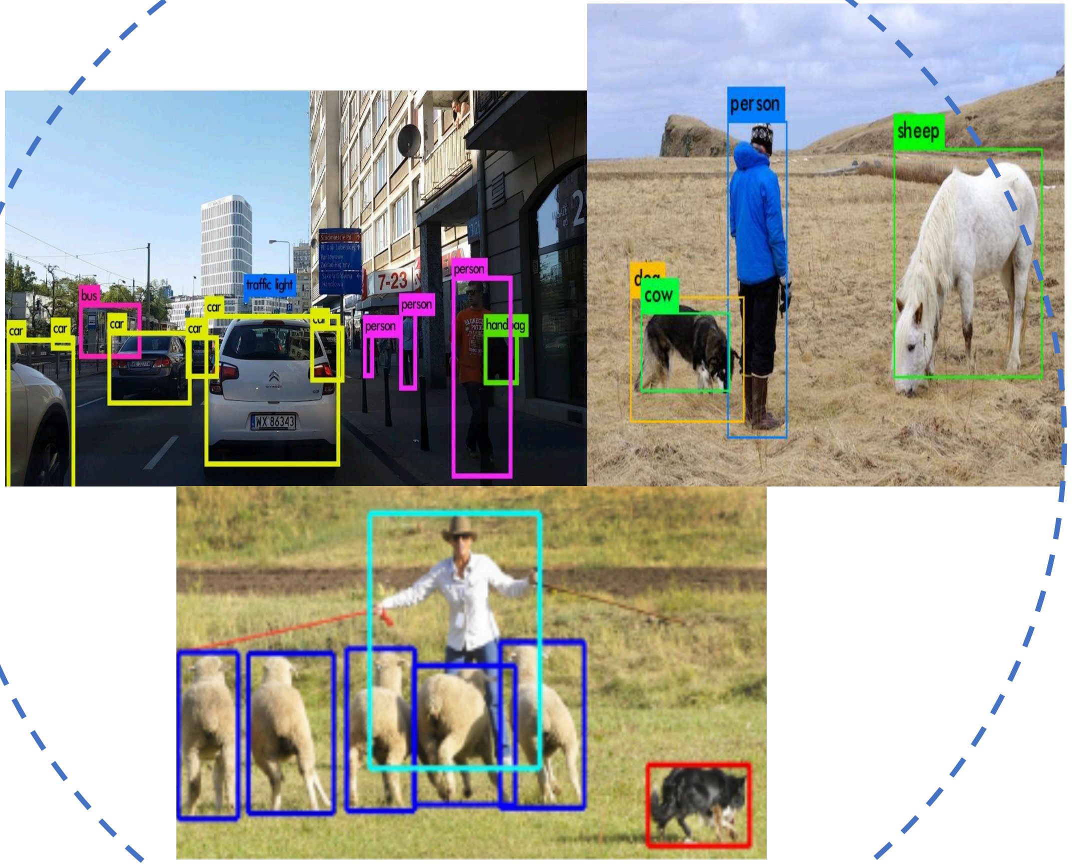
Wen Qian [1,2], Xue Yang[3,4], Silong Peng[1,2,*], Junchi Yan[3,4], Yue Guo[1,2]

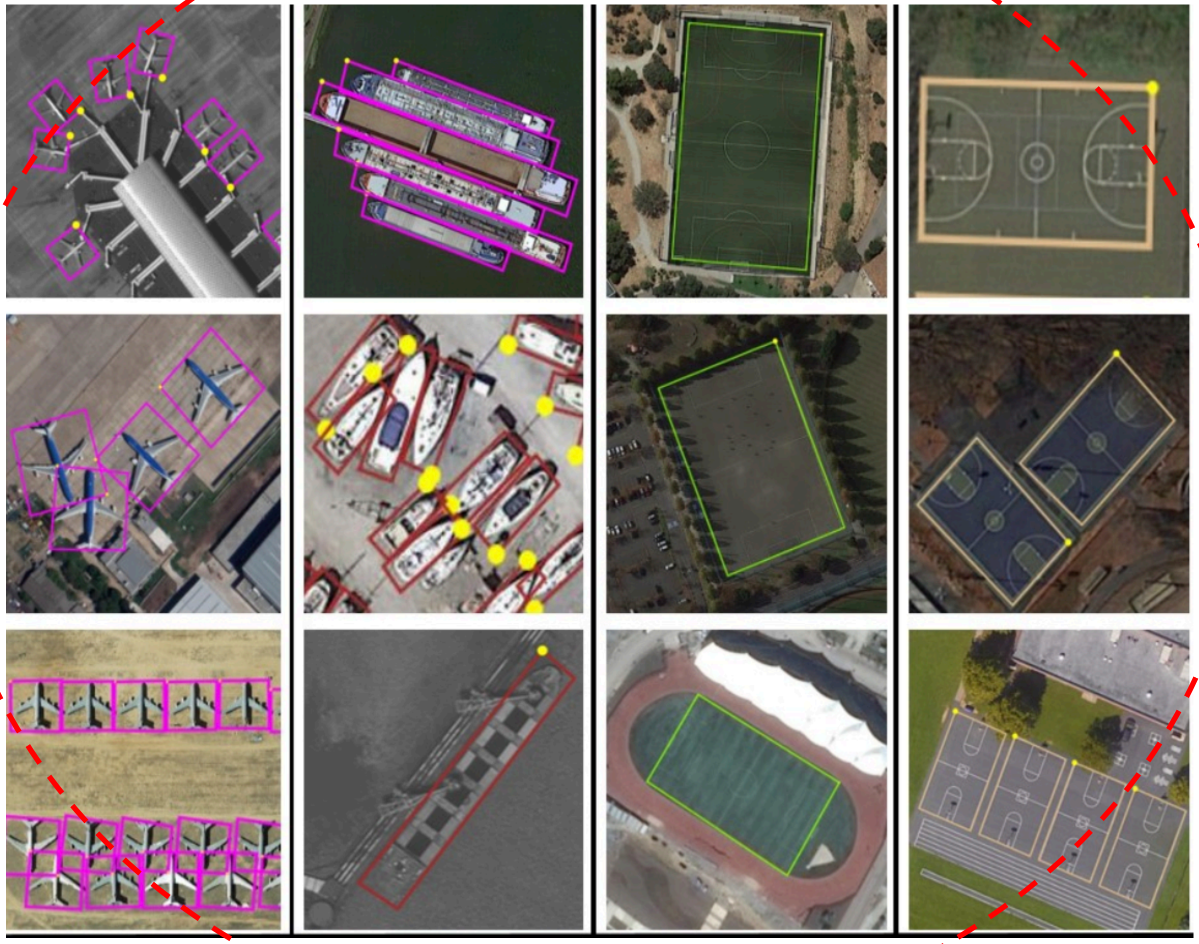
[1]Institute of Automation, Chinese Academy of Sciences
[2]University of Chinese Academy of Sciences
[3]Department of Computer Science and Engineering, Shanghai Jiao Tong University
[4]MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University
{qianwen2018, silong.peng, guoyue2013}@ia.ac.cn {yangxue-2019-sjtu, yanjunchi}@sjtu.edu.cn

35th AAAI Conference on Artificial Intelligence

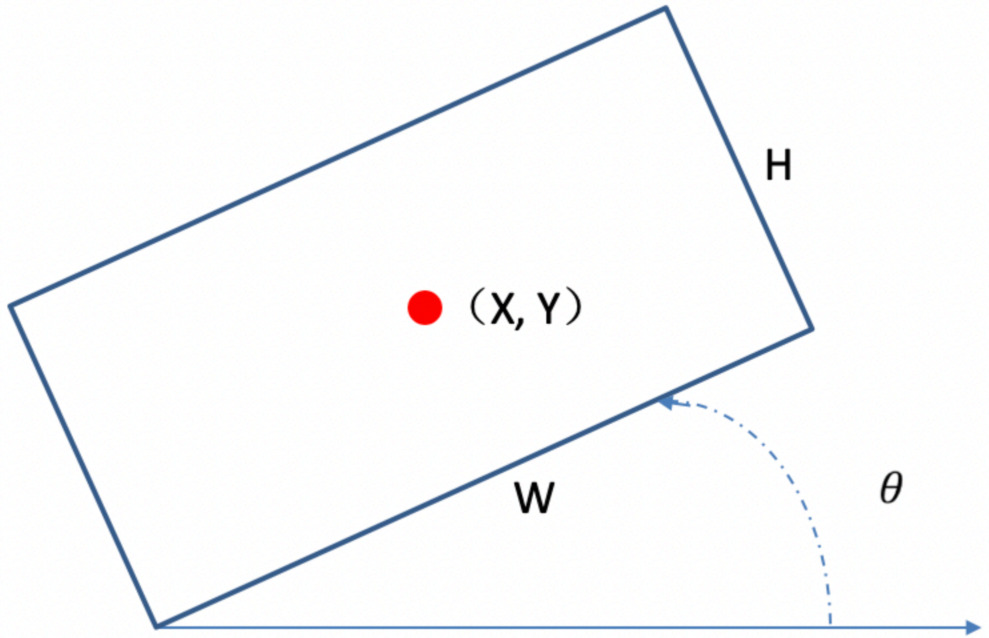
A Virtual Conference

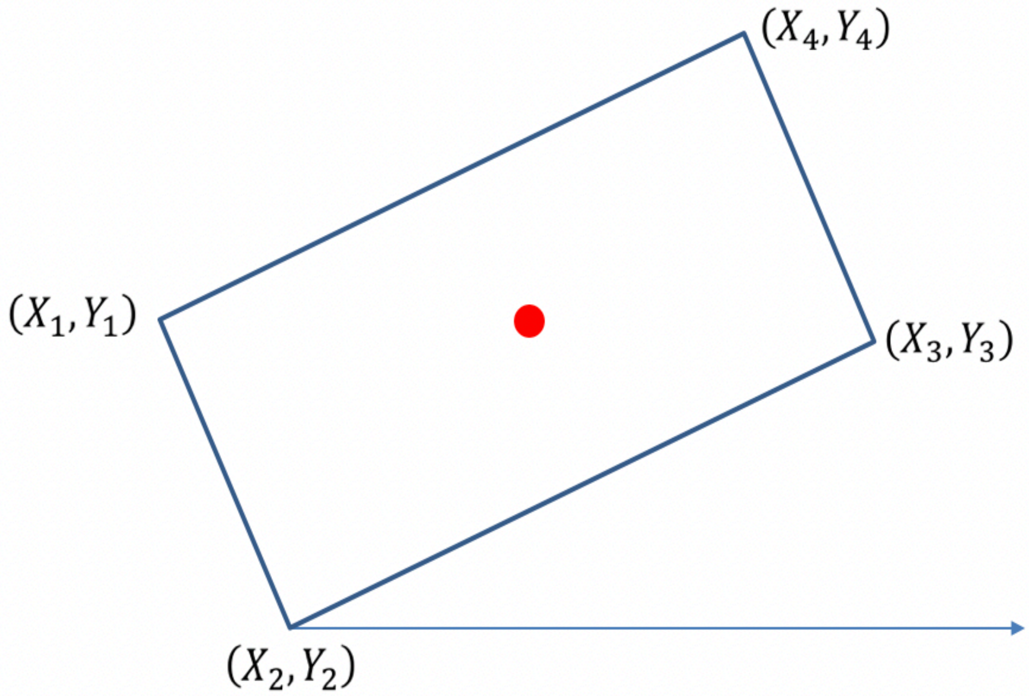
Horizontal object detection and Rotated object detection



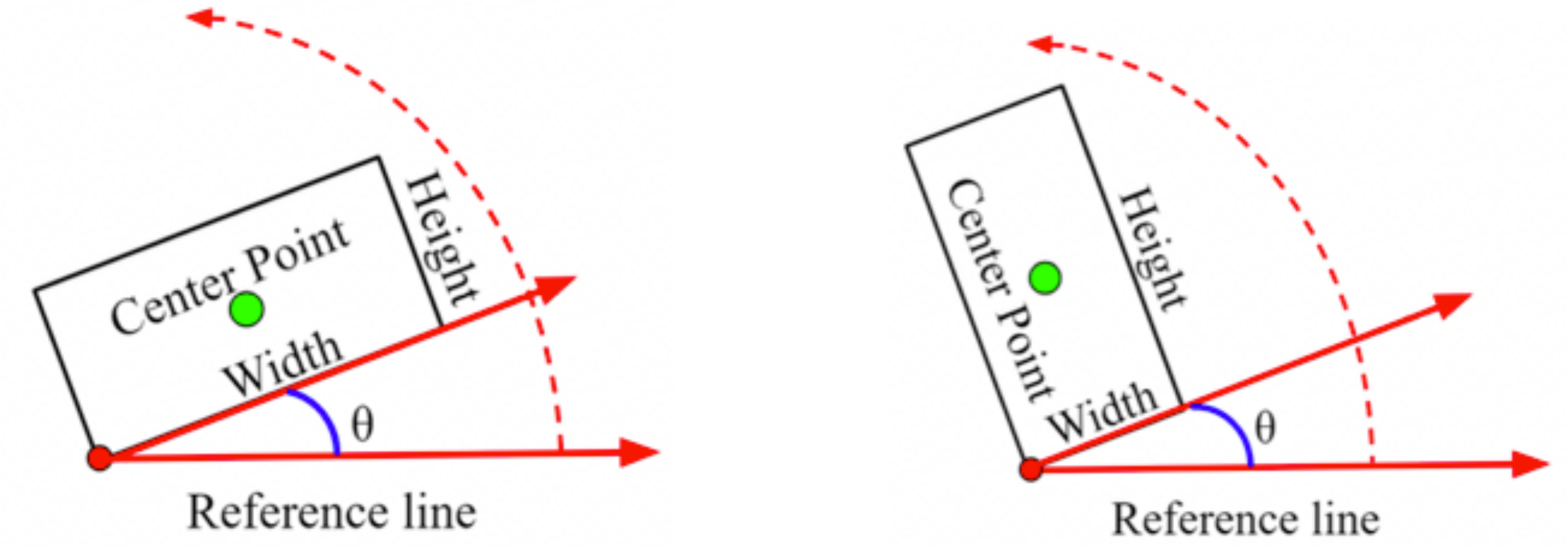


Traditional Definitions of rotated bounding box





The discontinuity of loss in five-parameter methods



(a) Width is longer than height. (b) Height is longer than width.

The discontinuity of loss in five-parameter methods

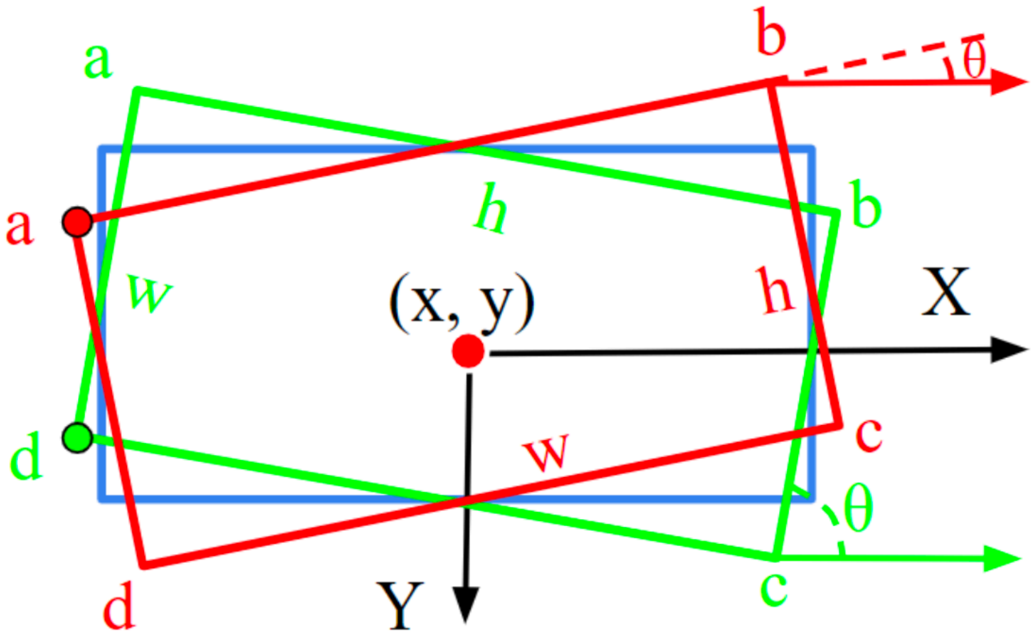
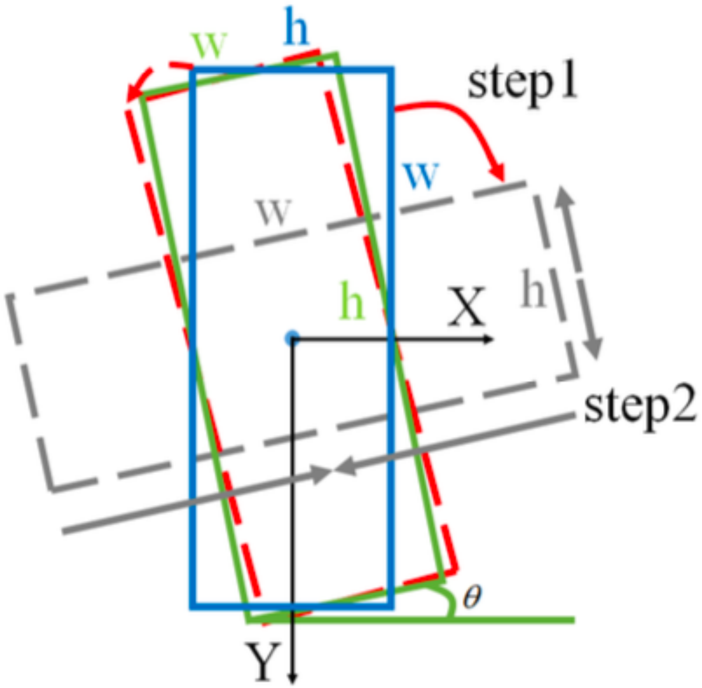
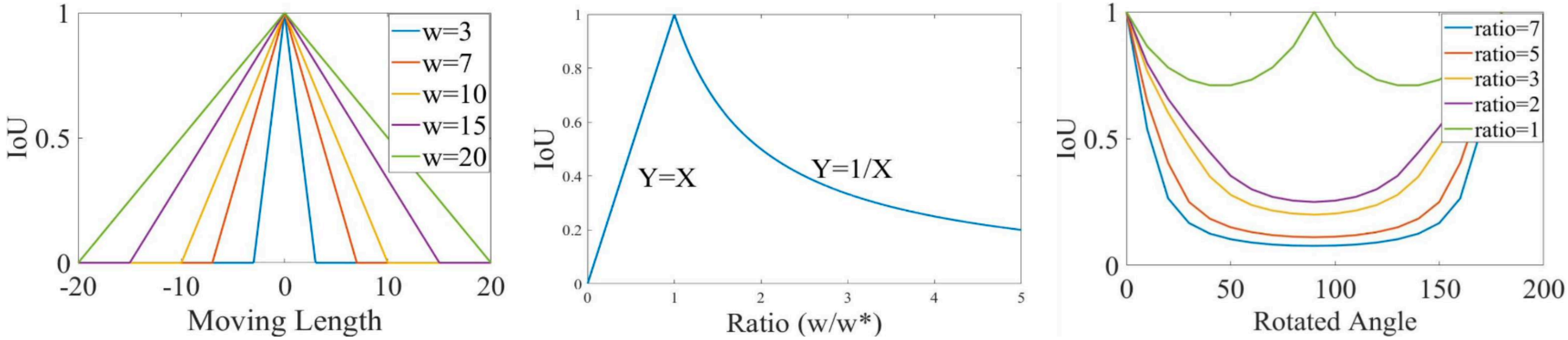


Figure 2: Loss discontinuity: rectangles in blue, red, and green respectively denote reference box, ground truth, and prediction. Here the reference box is rotated one degree clockwise to get the ground truth and is rotated similarly counterclockwise to obtain the prediction. Then the three boxes are described with five parameters: reference (0, 0, 10, 25, -90°), ground truth (0, 0, 25, 10, -1°), and prediction (0, 0, 10, 25, -89°). Here ℓ_1 loss is far more than 0.

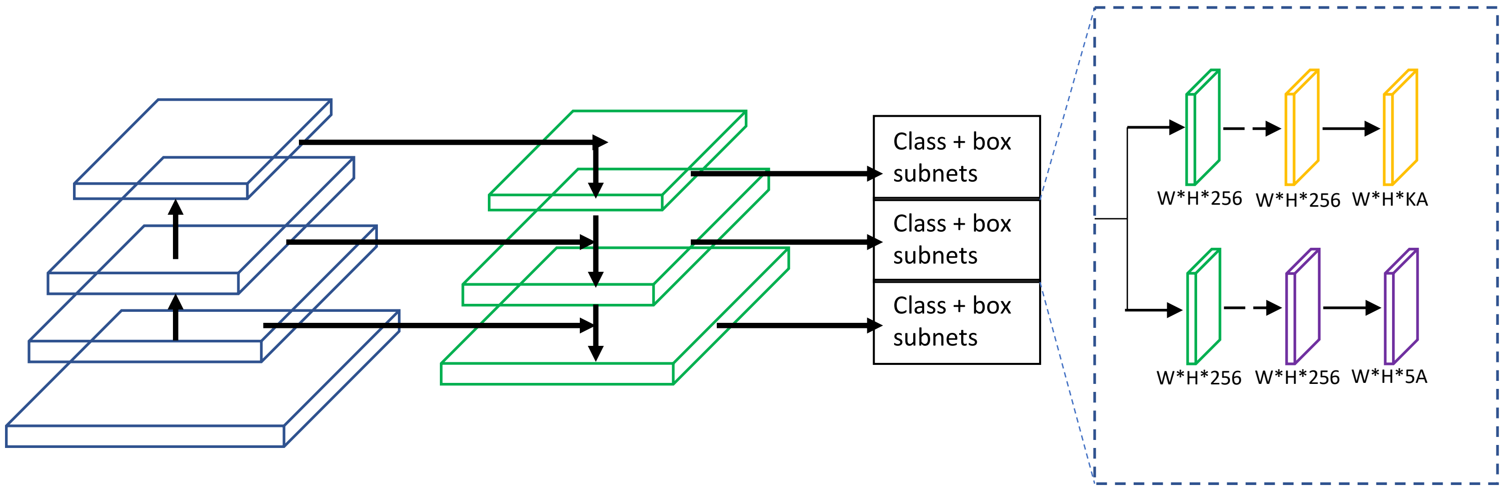
The discontinuity of loss in five-parameter methods



The discontinuity of loss in five-parameter methods



(a) Relation between center point and IoU (b) Relation between width and IoU (c) Relation between angle and IoU



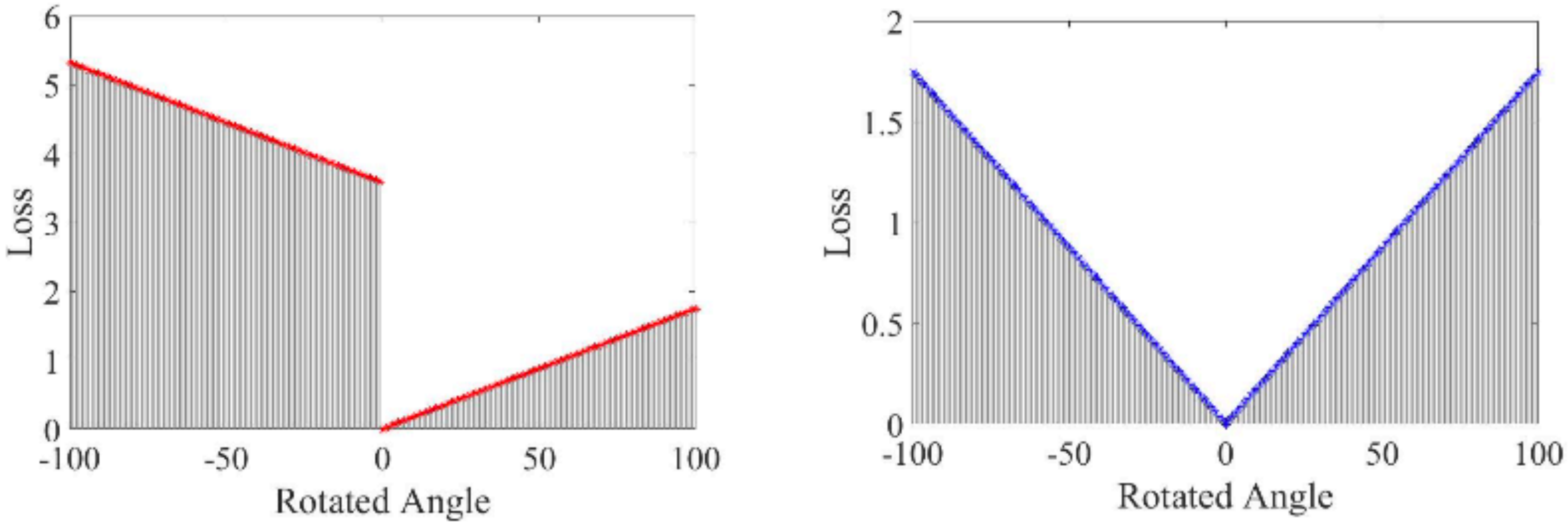
$$t_x = \frac{x - x_a}{w_a} \qquad t_y = \frac{y - y_a}{h_a} \qquad t_w = \log\left(\frac{w}{w_a}\right)$$

$$t_h = \log\left(\frac{h}{h_a}\right) \qquad t_\theta = \frac{\theta\,\pi}{180}$$

The discontinuity of loss in five-parameter methods

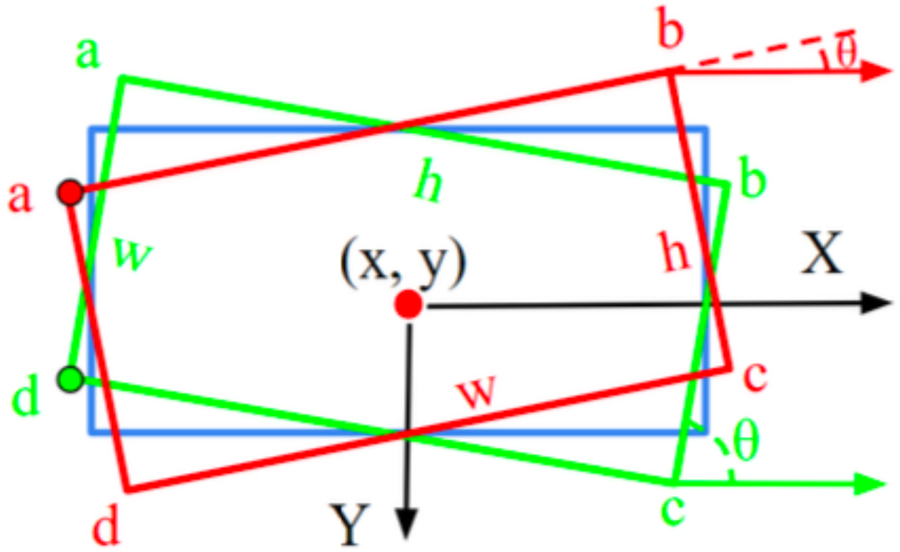
$$\ell_{cp} = |x_1 - x_2| + |y_1 - y_2|, \tag{2}$$

$$\ell_{mr}^{5p} = \min \begin{cases} \ell_{cp} + |w_1 - w_2| + |h_1 - h_2| + |\theta_1 - \theta_2| \\ \ell_{cp} + |w_1 - h_2| + |h_1 - w_2| \\ \quad + |90 - |\theta_1 - \theta_2||, \end{cases} \tag{3}$$



(a) Discontinuous ℓ_1-loss (b) Continuous ℓ_{mr}^{5p}

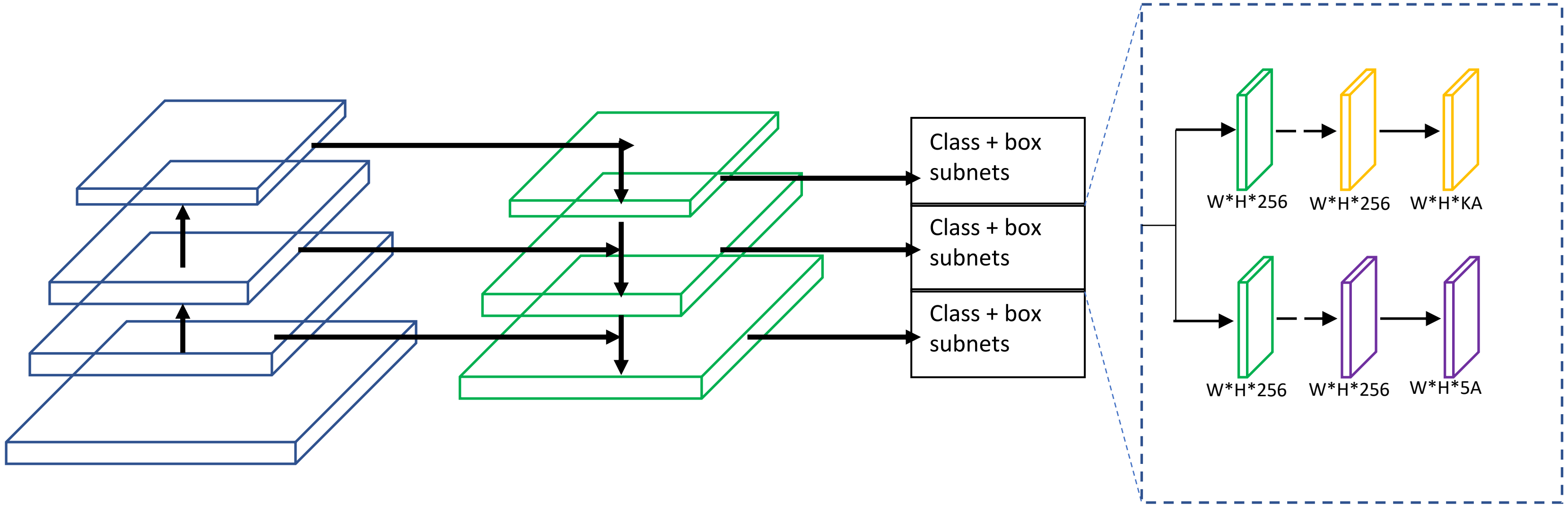
The discontinuity of loss in eight-parameter methods



The discontinuity of loss in eight-parameter methods

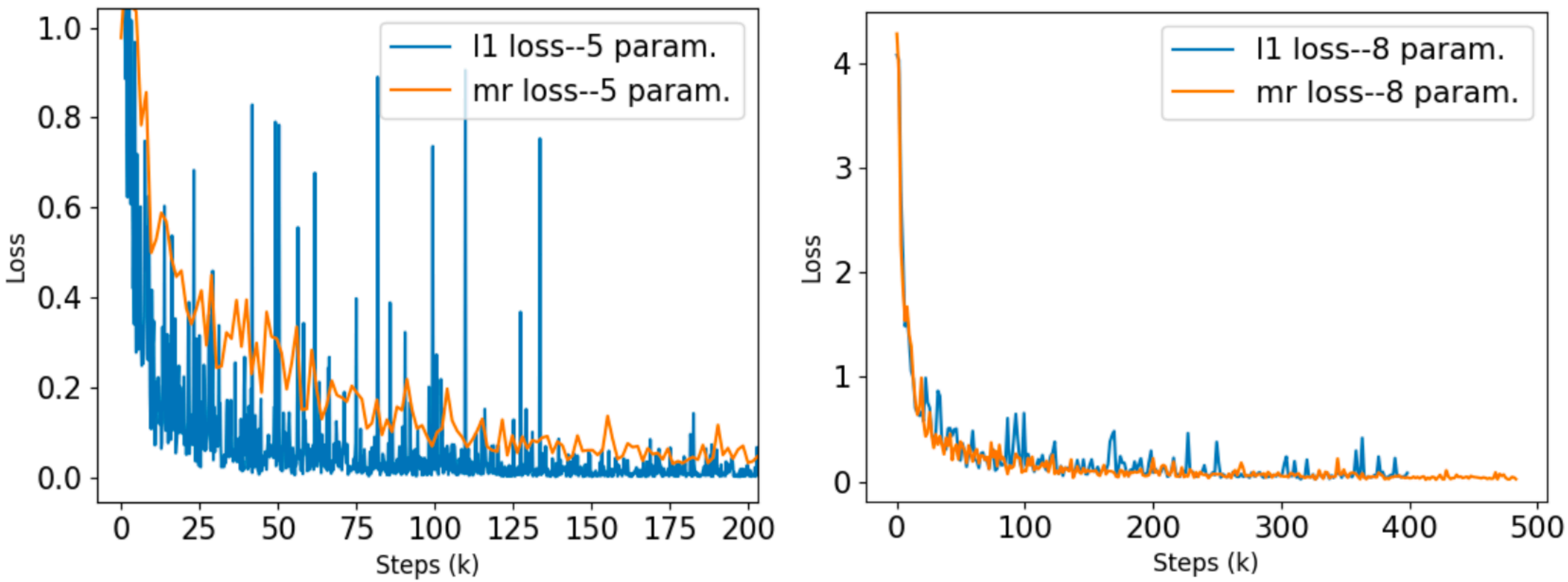
$$\ell_{mr}^{8p} = \min \begin{cases} \sum_{i=0}^{3} \left(\frac{|x_{(i+3)\%4} - x_i^*|}{w_a} + \frac{|y_{(i+3)\%4} - y_i^*|}{h_a} \right) \\ \sum_{i=0}^{3} \left(\frac{|x_i - x_i^*|}{w_a} + \frac{|y_i - y_i^*|}{h_a} \right) \\ \sum_{i=0}^{3} \left(\frac{|x_{(i+1)\%4} - x_i^*|}{w_a} + \frac{|y_{(i+1)\%4} - y_i^*|}{h_a} \right) \end{cases} \tag{6}$$

The discontinuity of loss in eight-parameter methods



$$t_x = \frac{x - x_a}{w_a} \qquad t_y = \frac{y - y_a}{h_a} \qquad t_w = \log\left(\frac{w}{w_a}\right)$$

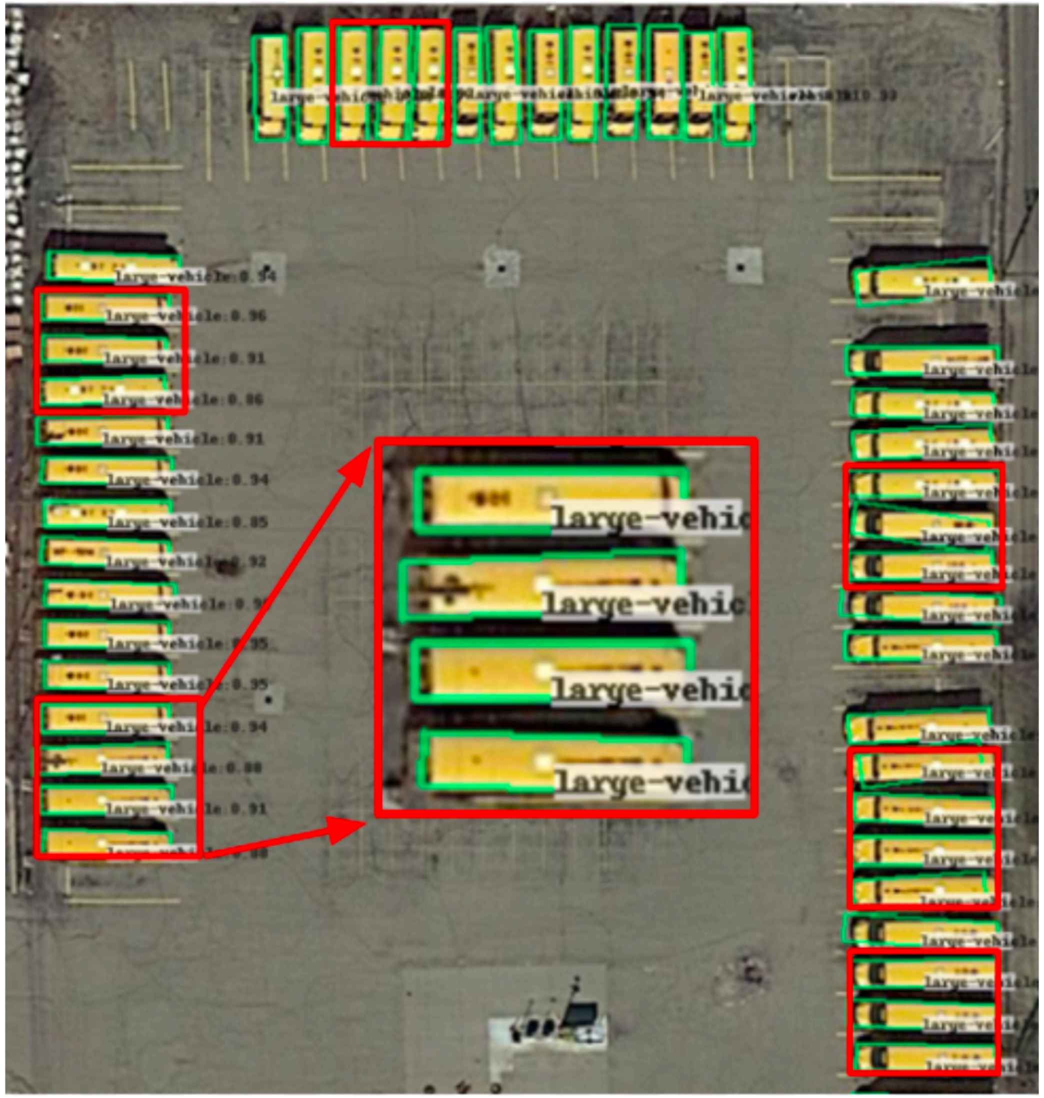
$$t_h = \log\left(\frac{h}{h_a}\right) \qquad t_\theta = \frac{\theta\,\pi}{180}$$



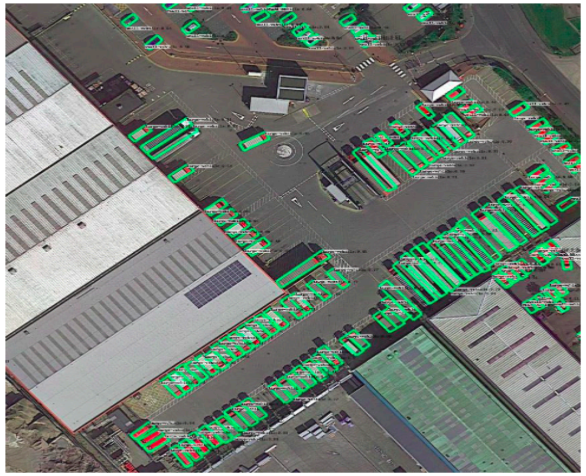
(a) Loss curves (five-param.) (b) Loss curves (eight-param.)

CASIA

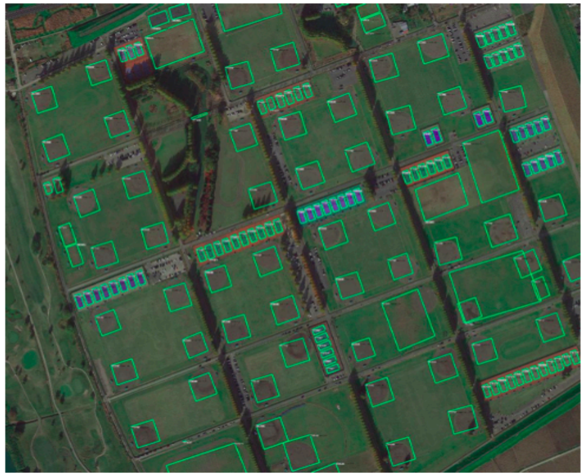

large-vehic
large-vehic
large-vehic
large-vehic


large-vehic
large-vehic
large-vehic
large-vehic

CASIA



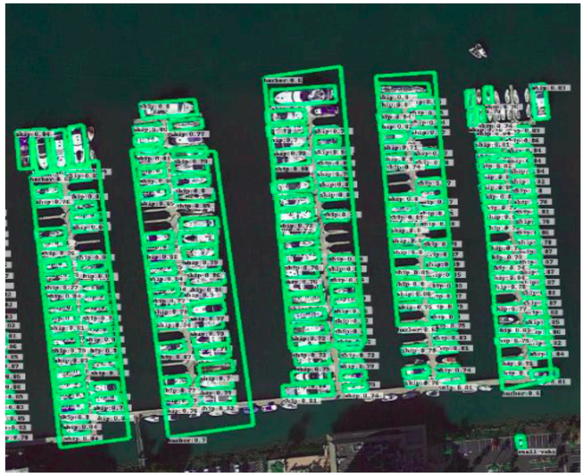
(a) Vehicles



(b) Tennis field

(c) Text in signs

(d) Storage tank



(e) Harbor and ship

(f) Text in elevator

THANKS