Learning Modulated Loss for Rotated Object Detection

Wen Qian1,2, Yue Yang1,2, Silong Peng1,2,∗, Junchi Yan1,4, Yue Guo1,2
1Institute of Automation, Chinese Academy of Sciences
2University of Chinese Academy of Sciences
3Department of Computer Science and Engineering, Shanghai Jiao Tong University
4MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University

Introduction

Rotation Object Detection
Task: Solving the discontinuity of loss which is caused by the contradiction between the definition of the rotated bounding box and the loss function.

Challenges
• Parameterization of rotated bounding box: two mainstream protocols for bounding box parameterization i.e. the five- and eight-parameter models.
• Discontinuity of Loss: The case exists both in the five-parameter and eight-parameter models caused by the contradiction between the definition of the rotated bounding box and the loss function.

Our main contributions
• Formulate the important while relatively ignored rotation sensitivity error (RSE) for region-based rotation detectors, which refers to the loss discontinuity.
• For the traditionally widely used five-parameter system and eight-parameter system, we devise a special treatment to ensure the loss continuity. The new loss is termed by \( \ell_{\text{mod}} \).
• Based on \( \ell_{\text{mod}} \), we respectively extend it to the one-stage and two-stage detection frameworks, which show state-of-the-art performance on DOTA and UCAS-AOD benchmarks.

Codes: https://github.com/Meganduo0/RSDLet-8P-4R

Proposed Approach

Overview
In this section, we firstly present two mainstream protocols for bounding box parameterization i.e. the five-parameter and eight-parameter models. Then we formally determine the loss discontinuity in the five-parameter and eight-parameter methods. We call such issues collectively as rotation sensitivity error (RSE) and propose a modulated rotation loss to achieve more smooth learning.

Parameterization of Rotated Bounding Box
Our five-parameter definition is in line with that in OpenCV.
The definition of eight-parameter is more simple: starting from the lower left corner, four clockwise vertices \((a, b, c, d)\) of the rotated bounding box are used to describe its location.

RSE in Five-parameter Methods
RSE is mainly caused by two reasons: i) The adoption of the angle param-eter and the exchange between width and height contribute to the sudden loss change (increase) in the boundary case. ii) Regression inconsistency of measure units exists in the five-parameter model.
The angle parameter causes the loss discontinuity. To obtain the predicted box that coincides with the ground truth box, the horizontal reference box is rotated counterclockwise, as shown in Fig.2a.

Different measurement units of five parameters make regression inconsistent. However, the impact of such artifacts is still unclear and has barely been studied in the literature. Relationships among all the parameters and IoU are empirically studied in Fig.3.

RSE in Eight-parameter Methods
The discontinuity of loss still exists in the eight-parameter regression model. Therefore, consider the situation of an eight-parameter regression in the boundary case, as shown in Fig.2b.

The Proposed Modulated Rotation Loss
\[
\ell_{\text{mod}} = \min \left( \ell_{\text{para}}(\text{modulated} - \text{para}), \right)
\]
\( \ell_{\text{para}} \) is the parameterized rotation loss. In this paper, we devise the following boundary constraints to modulate the loss as termed by modulated rotation loss \( \ell_{\text{mod}} \):

\[
\ell_{\text{mod}} = \min \left( \ell_{\text{para}}, \ell_{\text{modulated}} \right)
\]

\[
\ell_{\text{modulated}} = \min \left( \ell_{\text{modulated} - \text{para}} \right)
\]

Five-parameter Modulated Rotation Loss
Eight-parameter Modulated Rotation Loss

\[
\ell_{\text{RSE}} = \min \left( \ell_{\text{RSE}}(\text{modulated} - \text{para}) \right)
\]

\[
\ell_{\text{RSE}} = \min \left( \ell_{\text{RSE} - \text{para}} \right)
\]

Fig1: The five-parameter definition

Fig2: Boundary discontinuity analysis of five-parameter regression and eight-parameter regression. The red solid arrow indicates the actual regression process, and the red dotted arrow indicates the ideal regression process.

Fig3: Inconsistency in five-parameter regression model. between width and IoU.

Fig4: Comparison between two loss functions.

Experiments

Ablation experiments of \( \ell_{\text{mod}} \) on DOTA benchmark.

Ablation study using the proposed techniques on DOTA.

Ablation experiments of backbone data augmentation and balance.

Detection accuracy on DOTA.

Performance on UCAS-AOD:

Detection results on DOTA and ICDAR15.