Rethinking Rotated Object Detection with Gaussian Wasserstein Distance Loss

Xue Yang1,2,4, Junchi Yan1,2, Qi Ming3, Wentao Wang1, Xiaopeng Zhang4, Qi Tian4
1Department of CSE, SJTU 2MoE Key Lab of Artificial Intelligence, AI Institute, SJTU 3School of Automation, Beijing Institute of Technology 4Huawei Inc.

Introduction:

Task: Design a novel multi-category rotation detector for small, cluttered and rotated objects.

Challenges:
- The inconsistency between metric and loss
- Boundary discontinuity
- Square-like problem

Our main contributions:
- We summarize three flaws in state-of-the-art rotation detectors, i.e., inconsistency between metric and loss, boundary discontinuity, and square-like problem, due to their regression based angle prediction nature.
- We propose to model the rotating bounding box distance by Gaussian Wasserstein Distance (GWD) which leads to an approximate and differentiable IoU induced loss. It resolves the loss inconsistency by aligning model learning with accuracy metric and thus naturally improves the model.
- Our GWD-based loss can elegantly resolve boundary discontinuity and square-like problem, regardless how the rotating bounding box is defined. In contrast, the design of most peer works are coupled with the parameterization of box.

Codes: https://github.com/yzhang0827/RotationDetection

Proposed Approach

Most of the IoU-based loss can be considered as a distance function. Inspired by this, we propose a new regression loss based on Wasserstein distance. First, we convert a rotating bounding box \( B(x,y,w,h,\theta) \) into a 2-D Gaussian distribution \( N(\mu,\Sigma) \).

\[
\Sigma^{1/2} = \text{RSR}^T = \begin{pmatrix}
\cos\theta & -\sin\theta & 0 & 0 \\
\sin\theta & \cos\theta & 0 & 0 \\
0 & 0 & \sin\theta & -\cos\theta \\
0 & 0 & \cos\theta & \sin\theta
\end{pmatrix}
\]

\[
\mu = \begin{pmatrix} x \\ y \\ \frac{w}{2} \sin\theta \\ \frac{h}{2} \cos\theta \end{pmatrix}
\]

Gaussian Wasserstein Distance Regression Loss:

\[
E = \frac{1}{N} \sum_{n=1}^{N} \frac{1}{\tau + f(d^2)^2} \sigma_{\text{gwd}}(\alpha_n)
\]

Experiments:

- Ablation study for GWD on three dataset.
- Ablation study for GWD on two scene text datasets.
- Peer method comparison.

AP on different objects and mAP on DOTA.