Learning Oriented Remote Sensing Object Detection via Naive Geometric Computing

Yanjie Wang¹,* , Xu Zou¹,∗†, Zhijun Zhang¹, Wenhui Xu¹, Liqun Chen¹, Sheng Zhong¹, Luxin Yan¹, Guodong Wang²,

¹Huazhong University of Science and Technology
²Qingdao University

Abstract

Detecting oriented objects along with estimating their rotation information is one crucial step for analyzing remote sensing images. Despite that many methods proposed recently have achieved remarkable performance, most of them directly learn to predict object directions under the supervision of only one (e.g. the rotation angle) or a few (e.g. several coordinates) groundtruth values individually. Oriented object detection would be more accurate and robust if extra constraints, with respect to proposal and rotation information regression, are adopted for joint supervision during training. To this end, we innovatively propose a mechanism that simultaneously learns the regression of horizontal proposals, oriented proposals, and rotation angles of objects in a consistent manner, via naive geometric computing, as one additional steady constraint (see Figure 1). An oriented center prior guided label assignment strategy is proposed for further enhancing the quality of proposals, yielding better performance. Extensive experiments demonstrate the model equipped with our idea significantly outperforms the baseline by a large margin to achieve a new state-of-the-art result without any extra computational burden during inference. Our proposed idea is simple and intuitive that can be readily implemented. Source codes and trained models are involved in supplementary files.

Introduction

Different from general object detection (Redmon et al. 2016; Liu et al. 2016a; Ren et al. 2017; Law and Deng 2018; Duan et al. 2019; Yang et al. 2019b; Tian et al. 2019), oriented object detection, aiming at exhibiting objects using Oriented Bounding Boxes (OBBs) instead of Horizontal Bounding Boxes (HBBs) by estimating the rotation information of objects simultaneously (Han et al. 2021b; Guo et al. 2021; Yang et al. 2021a; Xu et al. 2021), has gained much attention especially in the community of remote sensing image analysis (Inglada and Michel 2009; Li et al. 2020; Zhao, Shi, and Zou 2021). For decades, many remarkable methods have been proposed and advanced the progress of oriented remote sensing detection while deep learning based models have achieved state-of-the-art performance recently. These artworks could be roughly categorized into 3 fashions: (1) Regressing rotation angle directly (Figure 2(a)). Methods of this fashion (Zhang et al. 2018; Ding et al. 2019; Yang et al. 2019a, 2021b) always regress a 5-parameter tuple \(\{x, y, w, h, \theta\}\) with an additional rotation angle \(\theta\) for OBB detection compared to the traditional 4-parameter tuple of HBB detection. (2) Classifying rotation angle directly (Figure 2(b)). These methods (Yang and Yan 2020; Yang et al. 2021a) apply the classification operation instead of the regression to obtain the rotation angle \(\theta\). (3) Representing the rotation information via regressing coordinates (Figure 2(c)). Unlike above methods, this kind of methods (Xia et al. 2018; Qian et al. 2021; Fu et al. 2021; Guo et al. 2021) regress the coordinates of a set of points (e.g. corners of the OBB, \(\{x_1, y_1, ..., x_d, y_d\}\)) to indirectly represent the rotation information.

However, the common character of these methods is that the object rotation information is learned by taking only one (e.g. the rotation angle, methods of fashion 1 and 2) or a few (e.g. coordinates, methods of fashion 3) groundtruth values as the supervision individually. Such a straightforward learning fashion may considerably limit the performance of...
a deep learning-based model thus yield discontented results since there inevitably exists unsatisfied regression errors on occasion. Oriented object detection would be more robust if these occasional regression errors could be alleviated by adopting extra joint constraints among regressed object proposals and rotation information for supervision during training. In addition, it has been studied that regressing horizontal proposals is also important for oriented object detection, as they could help reduce the model complexity and increase the number of high-quality proposals (Ding et al. 2019). Thus, one question remains: can we robustly learn both horizontal and oriented proposals simultaneously with an additional simple joint constraint with respect to the estimated rotation angle in a consistent manner?

In this paper, we try to tackle the issue of constructing a simple and intuitive connection among “horizontal object proposal”, “oriented object proposal”, and “object rotation angle” as an additional joint constraint for training the model. Specifically, for a same oriented object, there is an objective rule that the oriented proposal of this object could be derived from its horizontal proposal via a transformation operation, with respect to the rotation angle of the object, which can be easily obtained by just naive geometric computing (see Figure 1). Apparently, this rule could be adopted as an additional joint consistent constraint for supervising to alleviate occasional regression errors. Besides, the widely used strategy of positive and negative sample assignment is no longer suitable for oriented objects (see Figure 2(d)), which would heavily affect both horizontal and oriented object proposal regression. To this end, we extra introduce a novel oriented center prior guided label assignment strategy for further promoting the quality of proposals. To summarize, our contributions are 3-fold:

1. We are the first, to the best of our knowledge, that try to construct a connection among “horizontal object proposal”, “oriented object proposal”, and “object rotation angle” via an objective geometric rule.

2. We innovatively use this objective rule as a joint consistent constraint for supervision during training the detection model. A novel label assignment strategy is extra introduced for further promoting the performance.

3. The proposed idea can be readily implemented via just naive geometric computing. Extensive experiments demonstrate the model equipped with our idea significantly outperforms the baseline to achieve a new state-of-the-art result without any computational burden during inference.

**Related Work**

Over the past few years, we have witnessed the success of convolution neural network (CNN) for visual object detection, such as R-CNN series (Girshick et al. 2014; Girshick 2015) [Ren et al. 2017], YOLO series (Redmon et al. 2016; Redmon and Farhadi 2017), SSD (Liu et al. 2016a), and RetinaNet (Lin et al. 2017b). These classical detectors always exhibit natural objects (dog, car, etc.) with their Horizontal Bounding Boxes, which can not accurately describe oriented objects in remote sensing images.

In the field of remote sensing, there exist challenges for oriented objects with dense distribution, arbitrarily oriented, and large aspect ratios. To solve these problems, some well-designed methods [Liu et al. 2016b; Li et al. 2019; Pan et al. 2020; Hou et al. 2020; Han et al. 2021a; Ming et al. 2021] are proposed. Different from general detectors that use Horizontal Bounding Boxes, remote sensing object detectors can locate objects’ positions more accurately by leveraging Oriented Bounding Boxes with their rotation information. Regressing a 5-parameter tuple with an additional object rotation angle \( \theta \) (Ma et al. 2018; Zhang et al. 2018; Ding et al. 2019; Yang et al. 2019a, 2021b) is the most intuitive idea for detecting oriented objects. For solving the discontinuous boundaries problem, Circular Smooth Label (CSL) (Yang and Yan 2020) and Densely Coded Labels (DEL) (Yang et al. 2021a) argue to apply the classification operation instead of the regression to obtain the object rotation angle. Recently, improved 8-parameter tuple regression methods (e.g. Gliding vertex (Xu et al. 2021) and RS-Det (Qian et al. 2021)) have been proposed to enhance the robustness by directly predicting coordinates or offsets of a set of predefined points of oriented objects. For achieving more accurate object detection results, some of the aforementioned methods adopt the rotating anchor. But densely arranged anchors will increase the computational burden. RoI Transformer (Ding et al. 2019) has been proposed to convert Horizontal RoI (HRoI) to Rotated RoI (RRoI) for avoiding a large number of anchors.

The common character of these methods is that the object rotation information is individually learned from only one or a few groundtruth values. Different from them, we introduce an additional joint consistent constraint among the “horizontal object proposal”, “oriented object proposal”, and “ob-
Figure 3: The overall framework. Based on the Faster RCNN as backbones, we present two additional modules to fit the rotated property of objects. (a) The consistent geometric constraint (CGC) loss is employed on stable object proposal learning, minimizing the difference between “horizontal proposal” and the “transformed oriented proposal” (O2MER) from RoI Transformer by means of strict geometric relationships. (b) For enhancing the discriminative ability of RPN module, we distinguish the importance of different samples from diverse spatial configurations even with the same IoU and assign new positive/negative/ignored samples guided by oriented center prior to better capture the oriented object representation. The O2MER means the minimum enclosing rectangle of the oriented proposals.

Methodology

Overview

The overall proposed framework is shown in Figure 3. In this framework, the horizontal and oriented proposals are firstly regressed by the RPN (Ren et al. 2017) and RoI Transformer (Ding et al. 2019) respectively. The oriented angle of the object is also directly predicted. In the training procedure, the horizontal proposal and the oriented proposal would be steadily constrained by the Consistent Geometric Constraint (CGC) via the predicted oriented angle for promoting the detection result.

Thus, the loss $L$ of the entire model is defined as:

$$ L = L_{RPN} + L_{CGC} + L_{RT} + L_{RCNN} $$

where $L_{RPN}$ represents the loss of RPN, which generates horizontal proposals, $L_{RT}$ and $L_{RCNN}$ indicate losses of regressing oriented proposals and the final detection result adopted by RoI Transformer. $L_{CGC}$ is used for constraining the consistency between horizontal proposals and oriented proposals via the geometric transformation, which will be detailed in the next subsection. Moreover, a novel label assignment strategy is employed on the RPN module to better assign positive and negative samples that more suitable for oriented objects, resulting in further improvement.

Consistent Geometric Constraint Loss

Typically, the output of a traditional object detector for natural images is formulated as:

$$ (x, y, w_h, h_h, c) $$

where $x, y$ is the centre point, width, and height of the object’s horizontal bounding box, respectively, with category $c$. Relatively, the output of an oriented object detector can be represented with an extra oriented angle $\theta$:

$$ (x, y, w_o, h_o, \theta, c) $$

where $w_o, h_o$ represents the width and height of the oriented object’s bounding box. $\theta \in [-90^\circ, 90^\circ]$ is the object’s oriented angle between the long side of the oriented object and the x-axis. We explore the connection among “horizontal proposal”, “oriented proposal”, and “oriented angle $\theta'$” for promoting the performance of oriented object detection.
The proposed CGC loss relies on an objective geometric rule to convert the Oriented proposal to its Minimum Enclosing Rectangle (O2MER) and a rule-based constraint to evaluate the prediction error between proposals. The objective geometric rule (shown in Figure 1) is defined as:

\[
\begin{align*}
    w_h &= w_o \times \cos \theta + h_o \times |\sin \theta| \\
    h_h &= w_o \times |\sin \theta| + h_o \times \cos \theta
\end{align*}
\]  

(4)

Because the horizontal and its oriented groundtruth have a natural geometric relationship, the proposals of the same object in the detection result should also accord.

To alleviate occasional regression errors of the output horizontal proposals and oriented proposals, we introduce the Consistent Geometric Constraint (CGC) loss. Specifically, we first convert oriented proposals obtained from RoI transformer into O2MER via Eqn 4. Then, driven by IoU loss (Yu et al. 2016), we introduce an effective rule-based constraint to minimizing the difference between horizontal proposals and O2MER in our implementation. Similarly, converting horizontal proposals into oriented proposals is also working. The CGC loss can be defined as:

\[
L_{CGC} = 1 - \frac{B_h \cap B_{o2mer}}{B_h \cup B_{o2mer}}
\]

(5)

where \(B_h\) is predicted horizontal proposals generated by RPN, and \(B_{o2mer}\) means O2MER converted from predicted oriented proposals by Eqn 4.

We adopt CGC loss via naive geometric computing as a joint additional constraint for supervision during the training. The CGC loss, associating the OBB’s width/height/oriented angle with the HBB’s width/height, can effectively relieve the occasional error of the horizontal proposal or the oriented proposal regression and better ensure the consistency of them.

Oriented Center Prior Guided Label Assignment

The RPN module, which generates the horizontal proposal, is the initial step of our framework. Thus it has high requirements for the quality of the horizontal proposal. The key operation to obtain the horizontal proposal from RPN is to assign whether the sample is object or background. By introducing the anchor, the label assignment operation, dividing anchors into object or background, is the main method to set up the corresponding relationship between bounding box and classification label, which has been successfully applied to many methods (Liu et al. 2016a; Redmon et al. 2016; Lin et al. 2017a; He et al. 2017). There are detailed improvements for matching strategies of label assignment in FreeAnchor (Zhang et al. 2019), ATSS (Zhang et al. 2020), AutoAssign (Zhu et al. 2020). AutoAssign indicates all of the existing general detectors obey the center prior, that is, the sampling location near the center of bounding boxes is effective.

However, the existing horizontal anchor label assignment can not effectively describe the oriented objects in remote sensing, which has an obvious negative impact on detection performance. As shown in Figure 2(d), horizontal boxes X and Y have the same IoU and the same L1 distance with the horizontal GT, but it is obvious that box X has a higher coverage of the oriented GT. Therefore, the above evaluation method obeying the center prior can not effectively assign the quality of samples on oriented object detection. This means that the sampling locations towards oriented object detection should be close to the center of OBB rather than the center of HBB, following oriented center prior.

To solve this problem, we propose a horizontal anchor label assignment strategy for the oriented object, named Oriented Center Prior Guided (OCP-Guided) Label Assignment, which uses the prior information of horizontal space and the prior information of sample feature alignment to evaluate the sample jointly. Algorithm 1 describes how to label assign for an input image. For the object \(g\) in the input image, there are corresponding horizontal groundtruth \(G_h\) and oriented groundtruth \(G_o\). Firstly, we screen the candidate samples from all anchors \(A\) and select whose IoU with \(G_h\) is greater than 0.3 as the candidate positive samples and whose IoU is less than 0.3 as the negative samples. Then we formulate the degree of spatial overlap \(d_{gh}\) and the degree of sample feature alignment overlap \(d_{go}\) as follows:

\[
d_{gh} = \frac{c_g \cap g_h}{c_g \cup g_h}
\]

(6)

\[
d_{go} = \frac{c_g \cap g_o}{g_o}
\]

(7)

After that, we calculate the average of \(d_{gh}\) and \(d_{go}\) as the quality score \(t_g\) of the candidate \(c_g\). When \(t_g\) is greater than the Threshold, it is regarded as a positive sample. It is
worth noting that we take the area of $G_o$ as the denominator rather than the union of $g_o$ and $c$, while calculating $d_{go}$. The main reason is that the area of $G_o$ as the denominator can reflect the overlap proportion of $c$ and the oriented object features, and the value range of $d_{go}$ is 0-1.

This method can significantly improve the quality of the proposal, especially the object with a large aspect ratio, only by distinctively selecting the samples with high object feature proportion at the label assignment level.

Discussion

The proposed method does not need any extra manual annotations. Only horizontal and oriented groundtruth are adopted for supervision in our method, as same as ones adopted by other models. No extra annotation is required.

The proposed method will not cause extra computational burden during inference. The proposed idea is only involved during the training to learn a better detection model. Thus the structure of the detector and the number of parameters of our method would be totally as same as the baseline.

| Method | Backbone | PL | BD | BR | GTF | SV | LV | SH | TC | BC | ST | SBF | RA | HA | SP | HC | mAP |
|--------|----------|----|----|----|-----|----|----|----|----|----|----|-----|----|----|----|----|-----|
| R101  | 80.42    | 77.17 | 17.70 | 64.05 | 35.30 | 38.02 | 37.15 | 89.44 | 69.64 | 59.28 | 50.30 | 52.91 | 47.89 | 47.40 | 46.30 | 54.13 |
| R101-FPN | 81.36    | 74.30 | 47.70 | 70.32 | 64.89 | 67.82 | 69.88 | 90.76 | 79.06 | 78.20 | 53.64 | 62.90 | 67.02 | 64.17 | 50.23 | 68.16 |
| CAD* ‖ Deng, Liu, and Zhou 2019 | 87.80 | 82.40 | 49.40 | 73.50 | 71.10 | 63.50 | 76.60 | 90.90 | 79.20 | 73.30 | 48.40 | 60.90 | 62.00 | 67.00 | 62.20 | 69.90 |
| DRN* ‖ Pan et al. 2021 | 88.91 | 80.22 | 43.52 | 63.35 | 73.48 | 70.69 | 84.94 | 90.14 | 83.85 | 84.11 | 50.12 | 58.51 | 64.71 | 68.26 | 52.50 | 70.70 |
| Ours* | 88.88 | 81.24 | 53.15 | 60.65 | 78.10 | 88.83 | 87.80 | 83.61 | 49.36 | 66.19 | 72.10 | 72.36 | 58.70 | 71.74 |
| SCRNet* ‖ Dong et al. 2019 | 89.98 | 80.65 | 52.09 | 68.36 | 68.36 | 60.32 | 72.44 | 90.85 | 87.94 | 86.66 | 65.02 | 66.68 | 64.25 | 62.21 | 57.61 |
| R2Det* ‖ Pan et al. 2021 | 89.49 | 81.17 | 50.53 | 66.10 | 70.92 | 78.66 | 78.21 | 90.81 | 85.26 | 84.23 | 61.81 | 63.77 | 68.16 | 69.83 | 67.17 | 73.74 |
| S3-A-Net* ‖ Han et al. 2021 | 89.15 | 82.84 | 48.37 | 71.11 | 78.11 | 78.39 | 87.25 | 90.83 | 89.40 | 85.64 | 60.36 | 62.80 | 65.26 | 69.13 | 57.94 | 74.12 |
| ReDet* ‖ Han et al. 2021 | 89.79 | 82.64 | 53.97 | 75.00 | 78.13 | 84.06 | 88.84 | 89.80 | 87.28 | 85.25 | 61.76 | 60.99 | 75.96 | 68.07 | 63.59 | 76.25 |
| Ours | 88.92 | 84.56 | 53.20 | 78.20 | 77.85 | 83.24 | 88.04 | 90.87 | 87.75 | 84.98 | 65.40 | 63.29 | 76.03 | 70.90 | 66.62 | 77.22 |

Table 1: Comparisons with state-of-the-art detectors on DOTA dataset. * means multi-scale training and testing. The BEST and the SECOND BEST performances are highlighted in bold and underlined respectively. Our proposed method achieve new state-of-the-art results for both single-scale and multi-scale training and testing.
Table 2: Comparisons with state-of-the-art detectors on HRSC2016. mAP(07) indicates that the result is evaluated under VOC2007 metrics, and other methods are evaluated under VOC2012 metrics. The BEST performances are highlighted in **bold**. Our method also outperforms others to achieve a new state-of-the-art result.

| Method          | Backbone | mAP(07) | mAP(12) |
|-----------------|----------|---------|---------|
| R$^2$CNN (Ma et al. 2018) | R101     | 73.07   | 79.73   |
| R$^2$PN (Zhang et al. 2018) | VGG16   | 79.6    | –       |
| RoI Trans (Ding et al. 2019) | R101-FPN | 86.20   | –       |
| Gliding Vertex (Xu et al. 2021) | R101-FPN | 88.20   | –       |
| DRN (Pan et al. 2020) | H104 | –       | 92.70   |
| CenterMap (Wang et al. 2021) | R50-FPN | –       | 92.8    |
| R$^3$Det (Yang et al. 2021b) | R101-FPN | 89.26   | 96.01   |
| DCL (Yang et al. 2021a) | R152-FPN | 89.46   | 96.41   |
| CSL (Yang and Yan 2020) | R152-FPN | 89.62   | –       |
| S$^2$A-Net (Han et al. 2021a) | R50-FPN | 90.17   | 95.01   |
| ReDet (Han et al. 2021b) | ReR50-ReFPN | 90.46 | 97.63   |
| Ours            | R50-FPN | **90.57** | **97.86** |

Results on DOTA. The comparison between our proposed method and the existing state-of-the-art methods on DOTA is shown in Table 1. The accuracy evaluation of test data is obtained by submitting our predictions to the DOTA Evaluation server in a fixed format. Compared with the existing methods in the single-scale training and testing, it achieves the best performance, which is more accurate for multi-class object detection and more suitable for oriented object detection in remote sensing images. From the perspective of overall performance, 80.43% of our proposed method is obtained in the multi-scale training and testing and exceeds most existing methods. The ReDet (Han et al. 2021b) has the best performance in bridge (BR), ground track field (GTF), large vehicle (LV), basketball court (BC), storage tank (ST), and helicopter (HC). The Proposed method has distinct performance in small vehicle (SV), ship (SH), soccer ball field (SBF), harbor (HA), and swimming pool (SP). Our proposed method achieves new state-of-the-art results for both single-scale and multi-scale training and testing. Figure 4 shows some qualitative results of the proposed method on DOTA.

Figure 4: Qualitative detection results on the DOTA dataset. The confidence threshold is set to 0.3 when visualizing results.

Figure 5: The number of (a) horizontal proposal and (b) oriented proposals in different IoU of our method and the baseline. The number of both horizontal and oriented proposals of our method is higher than the baseline’s in high IoUs. That is, our method can improve the quality of both horizontal and oriented proposals, which intuitively indicates the effectiveness of our proposed idea.
Table 3: Ablation studies on the effectiveness of each component in our method on DOTA. After training with the proposed CGC loss, the model equipped with our idea outperforms the baseline by a large margin. The performance could be further promoted once the proposed OCP-Guided Label Assignment strategy is adopted. So the effectiveness of the geometric consistency is brought into full play under this setting.

| Baseline | CGC loss | OCP-Guided |
|----------|----------|------------|
| ✓        | 88.94    | 81.82      |
| ✓        | 88.68    | 82.41      |
| ✓        | 88.92    | 84.56      |

Effectiveness of Consistent Geometric Constraint Loss. In order to validate the effectiveness of the proposed CGC loss, we add it to the baseline and compare it with the baseline results while keeping other conditions unchanged. The experimental results show that CGC loss improves the overall performance by 0.6%. The main categories to be improved are swimming pool (SP) and helicopter (HC), which are difficult to learn with complex boundaries and angles. After training with the proposed CGC loss, the model equipped with our idea outperforms the baseline by a large margin. This also validates that our proposed loss makes the oriented object learning more robust by jointly modeling angle, width, and height, which provides a more flexible idea for the angle modeling of oriented object detectors.

Effectiveness of Oriented Center Prior Guided Label Assignment. To analyze the proposed OCP-Guided Label Assignment effectiveness, we replace the original RPN’s label assignment with the proposed method and compare the effect with the above method. As shown in Table 3, it can be seen that OCP-Guided label assignment has brought relatively significant results, which shows that it is essential to cover more object features in label assignment. In order to better understand the proposed method, the positive sample assignment diagrams of different label assignment methods are visualized for the oriented object, as shown in Figure 6. It can be observed that the center of the candidate box of the classical anchor assignment strategy is distributed near the center of horizontal groundtruth. However, anchors selected by our proposed strategy tend to be closer to the center of oriented groundtruth. The performance could be further promoted once the proposed OCP-Guided Label Assignment strategy is adopted due to the higher quality of proposals.

Figure 6: Comparison of the proposed anchor assignment strategy (right) with the classical anchor assignment strategy (left) for the oriented object. Dots of different colors represent the centers of the relevant anchors. Anchors selected by our proposed strategy are more tended to be close to the center of the oriented object.

Results on HRSC2016. HRSC2016 dataset is mainly ship objects with arbitrary orientation and a large aspect ratio, which challenges the angle regression and aspect regression of oriented object detector. According to Table 2, we can see our results exceed the existing methods on HRSC2016, which shows that our method has a better advantage in large aspect ratio. The comparison result verifies the effectiveness of our method on oriented object detection in aerial images. Figure 5(a) and (b) respectively shows the number of the horizontal and oriented proposals in different IoU of our method and the baseline on the HRSC2016 test set. In high IoUs, the number of both horizontal and oriented proposals of our proposed method is above baseline. That is, our method can improve the quality of both horizontal and oriented proposals, which intuitively indicates the effectiveness of our proposed idea.

D. Ablation Study

To verify the method’s effectiveness, we conducted two ablation experiments on the DOTA dataset, namely, Effectiveness of Consistent Geometric Constraint loss, Effectiveness of Oriented Center Prior Guided Label Assignment. Table 3 summarizes the results of our model under different settings, and the following is a detailed comparison. We use Faster RCNN (Ren et al. 2017) with FPN (Lin et al. 2017a), Roi Transformer (Ding et al. 2019) and RRoI align (Ding et al. 2019) as our baseline. The label assignment in the original RPN is used to assign each instance. The angle supervision method is only constrained by the smooth L1 loss between the predicted and the ground truth angles.

In this paper, we innovatively propose a mechanism that simultaneously learns the regression of horizontal proposals, oriented proposals, and rotation angles of objects in a consistent manner via naive geometric computing as a joint additional constraint for supervision during the training. For enhancing the quality of proposals and further promoting performance, a novel Oriented Center Prior Guided Label Assignment strategy is extra introduced. The proposed idea is simple and can be readily implemented. Extensive experiments demonstrate the model equipped with our method significantly outperforms the baseline by a large margin to achieve a new state-of-the-art result without any extra computational burden during inference.

Conclusion

In this paper, we innovatively propose a mechanism that simultaneously learns the regression of horizontal proposals, oriented proposals, and rotation angles of objects in a consistent manner via naive geometric computing as a joint additional constraint for supervision during the training. For enhancing the quality of proposals and further promoting performance, a novel Oriented Center Prior Guided Label Assignment strategy is extra introduced. The proposed idea is simple and can be readily implemented. Extensive experiments demonstrate the model equipped with our method significantly outperforms the baseline by a large margin to achieve a new state-of-the-art result without any extra computational burden during inference.
References

Azimi, S. M.; Vig, E.; Bahmanyar, R.; Körner, M.; and Reinartz, P. 2018. Towards multi-class object detection in unconstrained remote sensing imagery. In Proceedings of Asian Conference on Computer Vision, 150–165.

Chen, k.; Wang, J. Q.; Pang, J. M.; Cao, Y. H.; Xiong, Y.; Li, X. X.; Sun, S. Y.; Feng, W. S.; Liu, Z. W.; Xu, J. R.; et al. 2019. MMDEtection: Open mmlab detection toolbox and benchmark. arXiv preprint arXiv:1906.07155.

Ding, J.; Xue, N.; Long, y.; and Xia, G. S. 2019. Learning roi transformer for oriented object detection in aerial images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Duan, K. W.; Bai, S.; Xie, L. X.; Qi, H. G.; Huang, Q. M.; and Tian, Q. 2019. Centernet: Keypoint triplets for object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).

Fu, K.; Chang, Z.; Zhang, Y.; and Sun, X. 2021. Point-Based Estimator for Arbitrary-Oriented Object Detection in Aerial Images. IEEE Transactions on Geoscience and Remote Sensing, 59(5): 4370–4387.

Girshick, R. 2015. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 1440–1448.

Girshick, R.; Donahue, J.; Darrell, T.; and Malik, J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Guo, Z. H.; Liu, C.; Zhang, X. S.; Jiao, J. B.; Ji, X. Y.; and X., Y.Q. 2021. Beyond Bounding-Box: Convex-hull Feature Adaptation for Oriented and Densely Packed Object Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Han, J.; Ding, J.; Li, J.; and Xia, G. S. 2021a. Align deep features for oriented object detection. IEEE Transactions on Geoscience and Remote Sensing, 0(0): 1–11.

Han, J. M.; Ding, J.; Xue, N.; and Xia, G. S. 2021b. ReDet: A Rotation-equivariant Detector for Aerial Object Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

He, K. M.; Gkioxari, G.; P., D.; and Girshick, R. 2017. Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).

Hou, L. p.; Lu, K.; Xue, J.; and Hao, L. 2020. Cascade detector with feature fusion for arbitrary-oriented objects in remote sensing images. In Proceedings of the IEEE International Conference on Multimedia and Expo.

Ingладa, J.; and Michel, J. 2009. Qualitative Spatial Reasoning for High-Resolution Remote Sensing Image Analysis. IEEE Transactions on Geoscience & Remote Sensing, 47(2): 599–612.

Law, H.; and Deng, J. 2018. CornerNet: Detecting objects as paired keypoints. In Proceedings of the European Conference on Computer Vision (ECCV).

Li, C.; Xu, C.; Cui, Z.; Wang, D.; Zhang, T.; and J., Y. 2019. Feature-attentioned object detection in remote sensing imagery. In 2019 IEEE International Conference on Image Processing (ICIP).

Li, K.; Wan, G.; Cheng, G.; Meng, L.; and Han, J. 2020. Object detection in optical remote sensing images: A survey and a new benchmark. ISPRS Journal of Photogrammetry and Remote Sensing, 159: 296–307.

Lin, T.; Dollar, P.; Girshick, R.; He, K.; and Hariharan, S., B. 2017a. Feature Pyramid Networks for Object Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Lin, T.; Goyal, P.; Girshick, R.; He, K.; and Dollar, P. 2017b. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).

Liu, Z. K.; Wang, H. Z.; Weng, L. B.; and Yang, Y. P. 2016b. Ship rotated bounding box space for ship extraction from high-resolution optical satellite images with complex backgrounds. IEEE Geoscience and Remote Sensing Letters, 13(8): 1074–1078.

Ma, J.; Shao, W.; Ye, H.; Wang, L.; Wang, H.; Zheng, Y.; and Xue, X. 2018. Arbitrary-oriented scene text detection via rotation proposals. IEEE Transactions on Multimedia, 20(11): 3111–3122.

Ming, Q.; Zhou, Z. Q.; Miao, L. J.; Zhang, H. W.; and Li, L. H. 2021. Dynamic anchor learning for arbitrary-oriented object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Pan, X. J.; Ren, Y. Q.; Sheng, K. K.; Dong, W. M.; Yuan, H. L.; Guo, X. W.; Ma, C. Y.; and Xu, C. S. 2020. Dynamic refinement network for oriented and densely packed object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Qian, W.; Yang, X.; Peng, S.; Guo, Y.; and Yan, C. 2021. Learning modulated loss for rotated object detection. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI).

Redmon, J.; Divvala, S.; Girshick, R.; and Farhadi, A. 2016. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Ren, S.; He, K.; Girshick, R.; and Sun, J. 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis & Machine Intelligence, 39(6): 1137–1149.

Tian, Z.; Shen, C. H.; Chen, H.; and He, T. 2019. Fcos: Fully convolutional one-stage object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).
Wang, J. W.; Yang, W.; Li, H. C.; Zhang, H. J.; and Xia, G. S. 2021. Learning Center Probability Map for Detecting Objects in Aerial Images. IEEE Transactions on Geoscience and Remote Sensing, 59(5): 4307–4323.

Wei, H. R.; Zhang, Y.; Chang, Z. H.; Li, H.; Wang, H. Q.; and Sun, X. 2020. Oriented objects as pairs of middle lines. ISPRS Journal of Photogrammetry and Remote Sensing, 169: 268–279.

Xia, G.; Bai, X.; Ding, J.; Zhu, Z.; Belongie, S.; Luo, J.; Datcu, M.; Pelillo, M.; and Zhang, L. 2018. DOTA: A large-scale dataset for object detection in aerial images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Xu, Y.; Fu, M.; Wang, Q.; Wang, Y.; Chen, K.; Xia, G.-S.; and Bai, X. 2021. Gliding Vertex on the Horizontal Bounding Box for Multi-Oriented Object Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(4): 1452–1459.

Yang, X.; Hou, L. P.; Zhou, Y.; Wang, W. T.; and Yan, J. C. 2021a. Dense Label Encoding for Boundary Discontinuity Free Rotation Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Yang, X.; Liu, Q.; Yan, J. C.; Li, A.; Zhang, Z.; and Yu, G. 2021b. R3Det: Refined Single-Stage Detector with Feature Refinement for Rotating Object. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI).

Yang, X.; and Yan, J. C. 2020. Arbitrary-oriented object detection with circular smooth label. In Proceedings of the European Conference on Computer Vision (ECCV).

Yang, X.; Yang, J.; Yan, J. C.; Zhang, Y.; Zhang, T.; Guo, Z.; Sun, X.; and Fu, K. 2019a. SCRDet: Towards more robust detection for small, cluttered and rotated objects. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).

Yang, Z.; Liu, S. H.; Hu, H.; Wang, L. W.; and Lin, S. 2019b. Reppoints: Point set representation for object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).

Yu, J. H.; Jiang, Y. J.; Y., W. Z.; Cao, Z. M.; and Huang, T. S. 2016. Unitbox: An advanced object detection network. In Proceedings of ACM International Conference on Multimedia.

Zhang, F. C.; Chi, F.; Q., Y. Y.; Lei, Z.; and Li, S. Z. 2020. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Zhang, G.; Lu, S.; and Zhang, W. 2019. CAD-Net: A Context-Aware Detection Network for Objects in Remote Sensing Imagery. IEEE Transactions on Geoscience and Remote Sensing, PP: 1–10.

Zhang, X. S.; Wan, F.; Liu, C.; Ji, R. R.; and Ye, Q. X. 2019. Freeanchor: Learning to match anchors for visual object detection. In Advances in Neural Information Processing Systems (NeurIPS).