Relational Matching for Weakly Semi-Supervised Oriented Object Detection

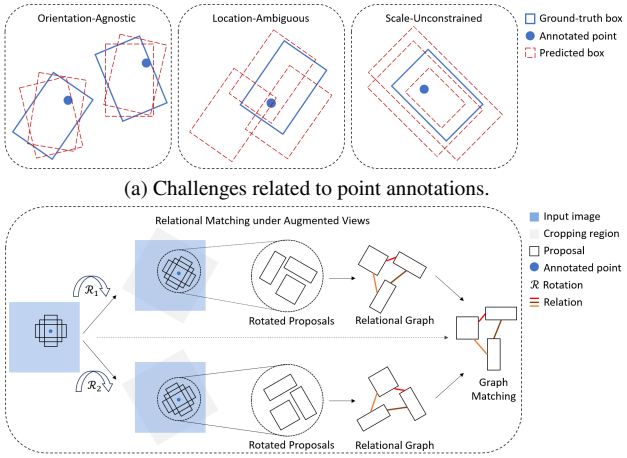
Wenhao Wu[1], Hau-San Wong[1]*, Si Wu[2], and Tianyou Zhang[2]
[1]Department of Computer Science, City University of Hong Kong
[2]School of Computer Science and Engineering, South China University of Technology
wenhaowu5-c@my.cityu.edu.hk, cshswong@cityu.edu.hk, cswusi@scut.edu.cn,
cszhangtianyou@mail.scut.edu.cn

Abstract

Oriented object detection has witnessed significant progress in recent years. However, the impressive performance of oriented object detectors is at the huge cost of labor-intensive annotations, and deteriorates once the annotated data becomes limited. Semi-supervised learning, in which sufficient unannotated data are utilized to enhance the base detector, is a promising method to address the annotation deficiency problem. Motivated by weakly supervised learning, we introduce annotation-efficient point annotations for unannotated images and propose a weakly semi-supervised method for oriented object detection to balance the detection performance and annotation cost. Specifically, we propose a Rotation-Modulated Relational Graph Matching method to match relations of proposals centered on annotated points between the teacher and student models to alleviate the ambiguity of point annotations in depicting the oriented object. In addition, we further propose a Relational Rank Distribution Matching method to align the rank distribution on classification and regression between different models. Finally, to handle the difficult annotated points that both models are confused about, we introduce weakly supervised learning to impose positive signals for difficult point-induced clusters to the base model, and focus the base model on the occupancy between the predictions and annotated points. We perform extensive experiments on challenging datasets to demonstrate the effectiveness of our proposed weakly semi-supervised method in leveraging point-annotated data for significant performance improvement.

1. Introduction

Oriented object detection, which aims at locating objects with the orientation property and identifying the corresponding categories, has achieved significant progress with the advance of deep learning and generic object detection.



(a) Challenges related to point annotations.

(b) Relational matching to align the relations of proposals centered around the annotated points between different models under augmented views.

Figure 1. Due to the ambiguity of the point annotations, the direct prediction consistency on annotated points is unsatisfactory. Our proposed relational matching method focuses the base model on the spatial and semantic information centered on the annotated points, thus alleviating the ambiguity of the point annotations in depicting the oriented objects.

However, the impressive performance of well-designed oriented object detectors depends on the availability of massive annotated data, which is costly in terms of both time and human resources. To alleviate the annotation burden, weakly supervised learning [5, 10, 11, 17], in which box annotations are replaced by image-level or point-level annotations, and semi-supervised learning [19, 23, 29], in which only limited images from the whole training dataset are annotated, have been separately studied in the field of generic object detection and oriented object detection. However, the problems of ambiguous supervision in weakly supervised learning and limited full supervision in semi-supervised learning result in unsatisfactory performance compared to their fully-supervised counterparts. To balance the annotation burden and detection performance, we resort to weakly semi-

*Corresponding author.

supervised learning on oriented object detection, in which limited annotated images and a large number of unannotated images with weak annotations are utilized to achieve comparable performance with their fully-annotated counterparts.

To alleviate the problem of missing annotation in the unannotated images, we introduce point annotations, which indicate the coarse locations of oriented objects, with categorical information as weak supervision for the unannotated data. As is common in semi-supervised object detection methods, we construct a teacher-student framework [36] for weakly semi-supervised oriented object detection, in which annotated images with manual annotations and unannotated images with pseudo annotations generated from the teacher model are utilized to enhance the student model. Orientation is a fundamental property in characterizing an oriented object, and the requirement of orientation estimation represents a more challenging problem compared to generic object detection, especially when the annotated data with accurate angular annotations is limited in the semi-supervised setting. Inspired by H2RBox [46], the consistency of predictions under different views facilitates the orientation estimation of oriented objects when weak annotations of circumscribed horizontal bounding boxes are provided. However, due to the ambiguity of point annotations in depicting the precise locations of oriented objects, the prediction consistency induced by only the point annotations is not enough to capture precise spatial and semantic information about oriented objects. As shown in Fig. 1a, each annotated point is orientation-agnostic, location-ambiguous, and scale-unconstrained, which leads to the difficulty in uniquely determining the only oriented box for each oriented object.

To address the ambiguity problem associated with the point annotations, we resort to matching the relations centered on the annotated points between the teacher and student models under different views. Specifically, we propose to set up the relational graphs over annotated points, with the features of the proposals centered on annotated points as vertices and the affinities between vertices as edges. The proposals centered on annotated points pinpoint the core regions that the base model mainly focuses on. Therefore, the relational graph with relational knowledge between proposals encodes context information around the annotated points. We match the relational graphs between different models and force the student model to follow the context recognition of the teacher model. We perform the relational graph matching weighted by the modulated orientation difference to focus the matching process on the relations between predictions with significant deviations on orientation estimation of different models. In addition, we propose a Relational Rank Distribution Matching method to align rank distributions on classification and regression between the teacher and student models. Finally, we introduce weakly supervised learning on the difficult points that both models are confused about. We perform alignment between aggregated categorical predictions of clusters and the categories associated with the related annotated points. We also force the regression outputs of the base model to enclose the related annotated points while excluding all neighboring annotated points, which enhances the discrimination of the base model on densely distributed objects in the aerial scenes.

We evaluate our proposed method on several challenging benchmarks, and achieve a significant performance gain over the baseline model. In particular, our proposed method outperforms the weakly semi-supervised method, Group R-CNN [50], under all semi-supervised settings.

We summarize our contributions as follows:

- We propose a weakly semi-supervised method for oriented object detection, in which sufficient images with point annotations are utilized to enhance the base model.
- To address the ambiguity problem from point annotations, we propose a Rotation-Modulated Relational Graph Matching method to align the contextual relations centered on the annotated points between different models.
- We propose a Relational Rank Distribution Matching method to further focus the base model on the relations between predictions from the annotated points by aligning rank distributions on both classification and regression.
- We introduce weakly supervised learning on difficult annotated points with inaccurate classification and regression outputs from both models.
- With readily available point annotations, our proposed weakly semi-supervised method contributes to significant performance gains on multiple challenging datasets.

2. Related Work

2.1. Oriented Object Detection

Oriented object detection, extended from generic object detection, aims to detect objects with the orientation property, and has achieved significant progress with the rapid development of deep learning based object detection. Ding *et al*. [8] extended the R-CNN [34] based framework with an additional head for angular prediction. Han *et al*. [15] alleviated the misalignment between axis-aligned features and arbitrary oriented objects through the proposed alignment convolution. Qian *et al*. [32] alleviated the boundary discontinuity problem based on the proposed modulated rotation loss. Yang and Yan [40] treated angular prediction as a classification task for precise rotation prediction of oriented objects. Yang *et al*. [42] further proposed a densely coded label representation for classification-based angular prediction. Pu *et al*. [31] rotated the convolution kernels to accommodate the orientation variations of oriented objects. SCRDet [41] and SCRDet++ [45] are devised to recognize rotated objects with small scales. Gaussian-based methods [43, 44, 47, 48] can alleviate the problems caused by differ-

ent definitions of oriented objects. Xu *et al*. [39] performed oriented object detection on the quadrilateral-based representation of oriented objects. In addition, some methods [14, 18, 25] represented oriented objects as point-sets, and constructed convex hulls for oriented object detection.

2.2. Semi-Supervised Object Detection

Semi-supervised learning, in which limited fully-annotated data and plenty of unannotated data are utilized for improved performance, is promising to solve the problem of annotation deficiency in object detection. Zhou *et al*. [52] proposed to generate pseudo annotations for unannotated data through a teacher model with continuously updated parameters. Tang *et al*. [35] further adopted soft labels as the training targets for the student model. Xu *et al*. [38] restricted the model's learning on backgrounds to alleviate the negative influence of false positives. Liu *et al*. [30] and Li *et al*. [23] proposed to alleviate the imbalance problem under the semi-supervised setting. Chen *et al*. [3] and Chen *et al*. [2] proposed to alleviate the confirmation bias issue [1] caused by noisy pseudo annotations. Scale variance is also one of the main challenges in semi-supervised object detection, and has been explored in recent works [13, 20, 29]. In addition to semi-supervised object detection with sparse predictions, recent works [21, 28, 51] have adopted the dense predictions of the teacher model as the training target for the student model. Point annotations, as weak supervision for unannotated data, have also been explored to facilitate semi-supervised object detection in recent works [4, 12, 49, 50]. Hua *et al*. [19] first achieved semi-supervised oriented object detection through the proposed rotation-aware loss.

3. Method

3.1. Problem Definition

The objective of weakly semi-supervised oriented object detection is to enhance the detection performance of the base detector through training on limited fully-annotated images $X_a = \{I_i^a, y_i^a\}_{i=1}^{N_a}$ and a large number of weakly-annotated images $X_p = \{I_i^p, y_i^p\}_{i=1}^{N_p}$, where y_i^a is the annotation for the fully-annotated image I_i^a with oriented bounding box and categorical information, and y_i^p is the annotation for the weakly-annotated image I_i^p with point location and categorical information. N_a and N_p denote the number of fully-annotated images and point-annotated images, respectively. With point-annotated images, we can achieve a reasonable balance of the detection performance on oriented object detection and annotation burden on a large number of images.

3.2. Overview

An overview of the proposed weakly semi-supervised framework is shown in Fig. 2. With the inspiration of Mean Teacher [36], we adopt the teacher-student framework, in which a teacher model and a student model share the same architecture but with different optimization. Specifically, the teacher model is responsible for generating reliable pseudo annotations for sufficient images with only point annotations. Along with limited fully-annotated images, these pseudo-annotated images will be utilized to enhance the student model by standard gradient descent, and the objective to be optimized is as follows:

$$\begin{aligned}\mathcal{L} = \mathcal{L}_a(X_a) + \lambda_p \mathcal{L}_p(X_p) + \lambda_g \mathcal{L}_g(X_p) + \\ \lambda_{rm} \mathcal{L}_{rm}(X_p) + \lambda_{ws} \mathcal{L}_{ws}(X_p),\end{aligned} \tag{1}$$

where $\mathcal{L}_a(X_a)$ and $\mathcal{L}_p(X_p)$ is the supervision loss on fully-annotated images and pseudo-supervision loss on unannotated images respectively, both of which consist of classification and regression losses. λ_p is the weighting factor for the pseudo-supervision loss. λ_g, λ_{rm} and λ_{ws} are utilized to control the contributions of the proposed Rotation-Modulated Relational Graph Matching, Relational Rank Distribution Matching, and weakly supervised learning methods on unannotated images, which will be detailed in Sec. 3.3, Sec. 3.4, and Sec. 3.5, respectively. We set λ_p, λ_g λ_{rm}, and λ_{ws} as 1.0, 0.1, 0.1, and 0.5 in all experiments.

After the optimization of the student model in each iteration, the teacher model is updated from the progressively updated student model by the exponential moving average (EMA) with a smoothing factor of 0.999.

3.3. Rotation-Modulated Relational Graph Matching

To enhance the perception of the base model on oriented objects, we propose a Rotation-Modulated Relational Graph Matching method, in which the relational graphs over annotated points should be aligned between different models under augmented views. Specifically, with the annotated points as the centers, we first generate N_k proposals on a given anchor setting with different scales and aspect ratios, and these proposals form a cluster for each annotated point. We perform random rotation to generate two different views, in which the unannotated images along with the proposals centered on the annotated points are rotated. The unannotated images under two augmented views are fed to the teacher model or the student model, respectively. As shown in Fig. 1b, from the perspective of the augmented view, the proposals are relatively static to the images and the related points, and the relations between proposals inside each cluster should be consistent regardless of any augmented views.

With rotated proposals, we extract the Rotated Region of Interests (RRoIs) from the feature maps through RRoI Align [8], and treat the RRoIs as vertices in the relational graph of each cluster. We calculate the cosine similarity between RRoIs as semantic relations to build the relational edges in

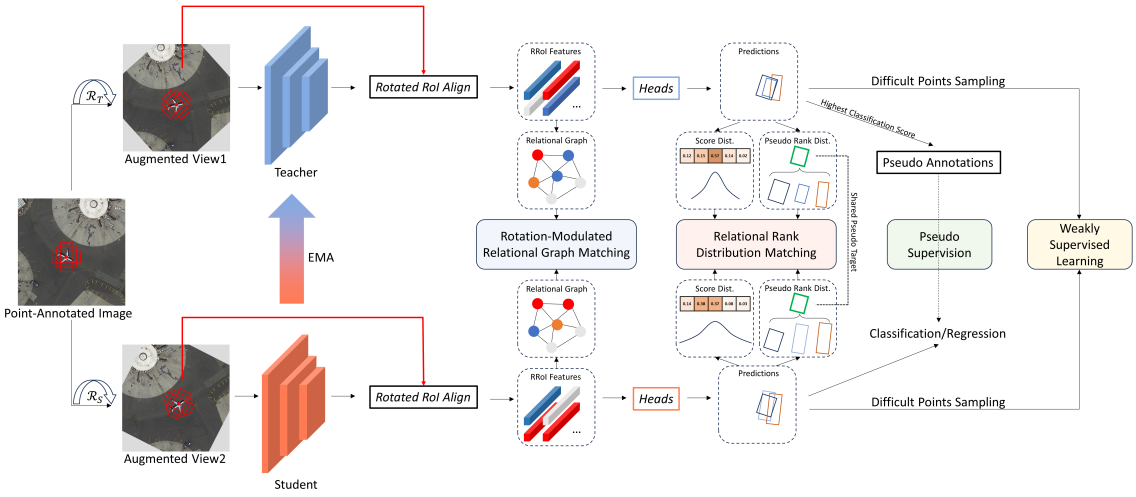


Figure 2. An overview of the proposed weakly semi-supervised framework for oriented object detection. In this framework, Rotation-Modulated Relational Graph Matching is proposed to match the relations of proposals centered on annotated points between the teacher and student models under augmented views. In addition, Relational Rank Distribution Matching is proposed to further align the relations of predictions over annotated points between different models by matching the rank distribution over classification and regression. We also introduce weak supervised learning to handle the difficult points with inaccurate classification and regression outputs.

the relational graph as follows:

$$\boldsymbol{E}^{\boldsymbol{R}} = \left\{ e_{ij}^{R} \middle| \frac{v_i^R v_j^R}{\|v_i^R\| \|v_j^R\|} \right\}, \tag{2}$$

where e_{ij}^{R} denotes the relational edge between vertices v_i^R and v_j^R in the same cluster under the augmented view $\boldsymbol{R}$, and $\boldsymbol{E}^{\boldsymbol{R}}$ is the edge matrix with e_{ij}^{R} as the element. We match the relational graphs between different models as follows:

$$\mathcal{L}_g = \frac{1}{N_g}(\lambda_g^v \mathcal{L}_g^v + \lambda_g^e \mathcal{L}_g^e), \tag{3}$$

where N_g denotes the number of the relational graphs corresponding to annotated points. $\mathcal{L}_g^v$ and $\mathcal{L}_g^e$ are the matching loss over vertices and edges, controlled by the weighting factors λ_g^v and λ_g^e, respectively. We set λ_g^v and λ_g^e to 1 by default.

The matching loss on vertices $\mathcal{L}_g^v$ facilitates the explicit distribution alignment between the teacher and student models over the proposals' features under different views. We propose a rotation-modulated mean squared error to focus the student model on the alignment of proposals with greater deviations to the teacher model over orientation estimation, which is as follows:

$$\mathcal{L}_g^v = \frac{1}{N_k} \frac{\sum_{i=1}^{N_k} d_{ii} \cdot \left(v_i^{R_t} - v_i^{R_s} \right)^2}{\sum_{i=1}^{N_k} d_{ii}} \tag{4}$$

where R_t and R_s denote the augmented views applied on inputs to the teacher model and the student model, respectively. d_{ii} is the modulated orientation difference as follows:

$$d_{ii} = min\left(\left| \frac{a_i^{R_t} - a_i^{R_s}}{\pi/2} \right|, \left| \frac{\pi - \left| a_i^{R_t} - a_i^{R_s} \right|}{\pi/2} \right| \right) \tag{5}$$

where a_i denotes the angular predictions from different models over features of the proposals under augmented views. The modulated orientation difference [33] can avoid the sudden changes of angular values in the boundary cases. It is noted that the angle predictions, though evaluated on different views, are relatively offset to the related rotated proposals under augmented views, and thus can be directly used for the modulated orientation difference between both models.

In addition, we further propose the rotation-modulated matching loss on edges $\mathcal{L}_g^e$ to facilitate the relational alignment between different models as follows:

$$\mathcal{L}_g^e = \frac{1}{N_k^2} \left(\left\| \left(\boldsymbol{E}^{R_t} - \boldsymbol{E}^{R_s} \right) \odot \boldsymbol{D_a} \odot \boldsymbol{D_s} \right\|_2^2 \right) \tag{6}$$

where $\boldsymbol{D_a}$ is a modulated matrix with modulated orientation difference d_{ij}, calculated as in Eq. 5, between the i-th and j-th vertices as elements, and $\boldsymbol{D_s}$ is a modulated matrix with entries s_{ij}, the product of classification scores between the i-th and j-th vertices from the teacher and student models. $\odot$ denotes element-wise multiplication. Since some outliers with low scores have a great orientation difference from any other predictions, we weight the matching by $\boldsymbol{D_s}$ to focus the base model on the relational alignment between high-confidence proposals and limit the contribution from the noisy predictions to the relational matching.

The relational graph builds an implicit connection between clusters of proposals and the corresponding oriented objects, and thus the matching of relational graphs between different models can distill the insight of the teacher model on inferring the oriented objects to the student model.

3.4. Relational Rank Distribution Matching

As a complement to Rotation-Modulated Relational Graph Matching, we further propose a Relational Rank Distribution Matching method, in which we align the rank distribution of predictions from proposals in each cluster to facilitate the transfer of rich relation information from the teacher model to the student model under different views.

For classification, we build the rank distributions over logits of proposals inside each cluster as follows:

$$\mathcal{D}(p_i, y^p) = \frac{\exp(p_i^{y^p}/T)}{\sum_{k=1}^{N_k} \exp(p_k^{y^p}/T)} \tag{7}$$

where p_i denotes the logit of the i-th proposals over the total categories, and y^p is the category from the point related to the i-th proposal. $p_i^{y^p}$ denotes the logit induced from the category y^p. T is the temperature factor, and set to 1 in all experiments. We then propose to match the rank distribution between the teacher model and the student model as follows:

$$\mathcal{L}_{rm}^{cls} = \frac{1}{N_k} \sum_{i}^{N_k} -\mathcal{D}(p_i^t, y^p) \log \frac{\mathcal{D}(p_i^s, y^p)}{\mathcal{D}(p_i^t, y^p)} \tag{8}$$

where p_i^t and p_i^s are the logits from the teacher model and the student model, respectively. Different from previous rank matching based methods [21, 22, 26], our proposed rank matching over classification focuses the base model on both the relations between the positives, and between the positives and hard negatives in the same cluster, thus enhancing the discrimination of the student model on hard negatives.

In addition, we further propose to match the rank distribution on regression outputs, which can induce the student model to achieve consistency of the predicted location, scale and angular distributions with the teacher model. However, a direct rank matching between different models on regression is not feasible, since there are no precise and shared targets to encode the regression outputs into the analogous rank distribution. To address this problem, we treat the regression outputs of the teacher model as the pseudo targets, and build pseudo rank distributions over the regression outputs of different models as follows:

$$\mathcal{D}(rb_i, rb_j^*, R^*) = RIoU(\mathrm{rot}(rb_i, R^*), rb_j^*) \tag{9}$$

where rb_i is the regression output of different models, and rb_j^* is the pseudo target from the teacher model. R^* is the rotational difference between augmented views applied on rb_i and rb_j^*, and rot is the rotation operation. $RIoU$ is the rotated intersection-over-union to evaluate the deviations between predicted oriented objects. Therefore, we propose a pseudo rank matching to align the spatial consistency between the teacher and student models as follows:

$$\mathcal{L}_{rm}^{reg} = \frac{1}{N_k^2} \sum_{j}^{N_k} \sum_{i}^{N_k} -\sigma(p_j^*)\, \mathcal{D}(rb_i^t, rb_j^*, R_t^*) \log \frac{\mathcal{D}(rb_i^s, rb_j^*, R_s^*)}{\mathcal{D}(rb_i^t, rb_j^*, R_t^*)} \tag{10}$$

where σ is the softmax operation to convert the pseudo targets' logits p_j^* to prediction scores, which are used to measure the quality of the corresponding pseudo targets. rb_i^t and rb_i^s are the regression outputs from the teacher model and the student model, respectively. According to the definition of R^*, R_t^* and R_s^* equal to 0 and $min\left(\left|\frac{R_t - R_s}{\pi/2}\right|, \left|\frac{\pi - |R_t - R_s|}{\pi/2}\right|\right)$, respectively.

Different pseudo targets from the teacher model lead to multiple pseudo rank distributions for both models. We treat all pseudo rank distributions on a cluster between different pseudo targets from the teacher model and different predictions from both models as a whole, and concatenate them as a matrix for further analysis. In this matrix, all elements in the principal diagonal from the teacher model are **1**, while those from the student model indicate the proposal-wise regression deviation to the teacher model. The off-diagonal elements in the matrix relate the spatial location between the pseudo targets and the regression outputs of both models, thus encoding the regression outputs of both models into comparable rank distributions. The pseudo rank matching, seen as the alignment of the concatenated matrix, between the teacher and student models can not only facilitate the proposal-wise regression output consistency, but the spatial relation consistency of proposals between both models.

With rank matching on classification and regression, we obtain the loss function as follows:

$$\mathcal{L}_{rm} = \frac{1}{N_g}(\lambda_{rm}^{cls}\mathcal{L}_{rm}^{cls} + \lambda_{rm}^{reg}\mathcal{L}_{rm}^{reg}) \tag{11}$$

where λ_{rm}^{cls} and λ_{rm}^{reg} are the weighting factors for the corresponding rank matching, which are both set to 1 by default.

3.5. Weakly Supervised Learning on Difficult Annotated Points

While the recognition of the base model on positives and relations between proposals in each cluster improves with the proposed relational matching methods, there still exist difficult annotated points that are confusing for both the teacher and student models: (1) The total number of proposals centered on these annotated points are underrated with low scores by both models. (2) It is not desirable for the confident regression outputs of both models to exclude the related annotated points or enclose the neighbor points, which

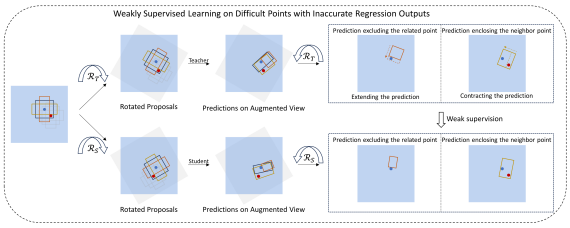


Figure 3. Weakly supervised learning on difficult points with regression predictions from both the teacher and student models inaccurately occupying annotated points.

are common cases in between the densely packed objects. The relation-based learning is not effective enough for these annotated points, since both the teacher and student models may not build a reasonable relational graph and rank distribution. Therefore, we propose to impose weakly supervised learning to handle the difficult annotated points.

To handle the points when all related proposals are underrated, we perform multi-instance learning (MIL) on the clusters of these points to impose positive signals at the cluster-level for the base model. Specifically, we treat the clusters, with all elements' classification scores lower than the threshold τ_m, as the low-quality bags, and average these scores to generate bag-level scores. We then apply the cross-entropy loss on the bag-level predictions and point-level categories. By applying MIL on these points, the base model is forced to carefully explore the positives in the related clusters.

We treat the predictions, with classification scores greater than the threshold τ_m and with regression outputs enclosing the neighbor points or excluding the related points from both models, as the second difficult case. As shown in Fig. 3, we calculate distances between annotated points and the four sides of the predicted boxes from the teacher model, and extend or contract the closest side to enclose the related points or exclude the neighbor points. We consider the refined predictions as the regression targets for the outputs of the student model, thus encouraging the student model to find the predicted boxes that approach the related points or depart from the neighbor points.

We obtain the loss function for weakly supervised learning as follows:

$$\mathcal{L}_{ws} = \lambda_{ws}^{cls}\mathcal{L}_{ws}^{cls} + \lambda_{ws}^{reg}\mathcal{L}_{ws}^{reg} \tag{12}$$

where $\mathcal{L}_{ws}^{cls}$ denotes the bag-level cross entropy loss for points with related proposals' scores less than τ_m, which is set to 0.1, and $\mathcal{L}_{ws}^{reg}$ denotes the smooth-L1 loss between predicted boxes, with predicted scores greater than τ_m but with inaccurate occupancy on annotated points, from the student model and the refined predicted boxes from the teacher model. It is noted that $\mathcal{L}_{ws}^{cls}$ is applied cluster-wise, while $\mathcal{L}_{ws}^{reg}$ is applied prediction-wise. λ_{ws}^{cls} and λ_{ws}^{reg} are the corresponding weighting factors, with both set to 1 by default.

4. Experiments

4.1. Datasets and Evaluation Metric

Datasets. We evaluate our proposed method on DOTA [9] and DIOR-R [7] datasets with the mean average precision (mAP) as the evaluation metric.

DOTA [9] is one of the largest datasets for oriented object detection. We implement the proposed method on DOTA-v1.0 with 15 categories and DOTA-v1.5 with an extra category. Both versions contain the same 2806 images, with 1/2, 1/6 and 1/3 of the images as the training set, the validation set, and the test set. We split the images into patches with a scale of 1024×1024 and with a pixel overlap of 200 between adjacent patches.

DIOR-R [7] is a challenging dataset with objects in the DIOR [24] dataset annotated with oriented bounding boxes. DIOR-R includes 11725 images as the trainval set and 11738 images as the test set with a uniform scale of 800×800 covering 20 categories.

We consider two semi-supervised settings, partially-annotated and fully-annotated, to evaluate the effectiveness of the proposed method in the scenes when different ratios of fully-annotated images are provided.

Partially-annotated. We follow [19] to sample different ratios of the training set from DOTA or the trainval set from DIOR-R as the annotated subset, with category distributions similar to the fully-annotated dataset. The remaining images are annotated with points randomly sampled inside the oriented bounding boxes. The evaluation is performed on the validation set of DOTA or the test set of DIOR-R.

Fully-annotated. We adopt the total fully-annotated images of the trainval set of DOTA as the annotated subset, and adopt the trainval set of DOTA-v2.0 with 555 images as the point-annotated subset. We evaluate the performance on the test set through the online evaluation server.

4.2. Implementation Details

We adopt Faster R-CNN [34] as the base model, with ResNet-50 [16] and FPN [27] as the backbone and neck respectively, to build the teacher model and the student model after warm-up on annotated images. With SGD as the optimizer with a momentum of 0.9 and weight decay of 0.0001, we train the model with an initial learning rate of 0.005 and a batch size of 2, half of which are from either the annotated subset or the point-annotated subset. We decay the initial learning rate at 8 and 11 epochs, and stop the training at 12 epochs. We follow [19] to utilize an asymmetric data augmentation strategy, with a strong augmentation for data fed into the student model and a weak augmentation for data fed into the teacher model. We perform all experiments on MMRotate [53] without using the multi-scale strategy. During inference, images without any annotations are fed into the teacher model for evaluation.

Methods	Type	DOTA-v1.0						DOTA-v1.5					
		5%	10%	20%	30%	50%	Fully	5%	10%	20%	30%	50%	Fully
FCOS$^\diamond$ [37]	FS	35.14	43.85	53.89	59.59	63.83	71.28	32.61	42.78	50.81	54.79	58.40	64.42
Faster R-CNN* [34]	FS	35.90	44.96	55.98	57.89	64.12	73.37	32.41	43.43	51.32	53.14	59.14	64.70
Dense Teacher$^\diamond$ [51]	SS	40.09	47.95	55.95	61.88	65.06	72.81	33.48	46.90	53.93	57.86	60.83	65.69
Soft Teacher* [38]	SS	44.11	50.29	58.43	60.24	64.33	74.00	37.02	48.46	54.89	57.83	61.25	65.10
SOOD$^\diamond$ [19]	SS	42.04	48.92	56.55	62.04	65.46	72.81	35.18	48.63	55.58	59.23	60.79	65.25
Group R-CNN$^\diamond$ [50]	WSS	51.24	57.58	62.78	65.92	66.07	73.10	33.99	51.99	58.20	59.01	60.76	65.58
Ours*	WSS	53.85	59.69	64.48	67.04	68.02	74.56	46.29	53.27	59.10	60.17	61.92	66.13

Table 1. Performance comparison of mAP on DOTA-v1.0 and DOTA-v1.5 under the partially-annotated and fully-annotated settings. FS, SS, and WSS denote training with annotated images, annotated images with unannotated images, and annotated images with point-annotated images, respectively. $\diamond$ and $*$ denote the base detectors used in different methods, including rotated Faster R-CNN and rotated FCOS.

Methods	Type	DIOR-R				
		5%	10%	20%	30%	50%
FCOS$^\diamond$ [37]	FS	36.99	44.10	51.50	54.00	58.69
Faster R-CNN* [34]	FS	39.91	44.65	52.24	54.82	59.07
Dense Teacher$^\diamond$ [51]	SS	46.35	50.41	55.91	57.20	60.51
Soft Teacher* [38]	SS	42.20	52.84	53.88	57.39	59.40
SOOD$^\diamond$ [19]	SS	40.60	44.46	51.97	54.34	58.91
Group R-CNN$^\diamond$ [50]	WSS	48.33	53.63	56.71	58.24	59.34
Ours*	WSS	50.76	54.69	57.53	58.80	61.72

Table 2. Performance comparison of mAP on DIOR-R under the partially-annotated setting.

4.3. Main Results

We compare our proposed method with the competing methods with and without point annotations. We re-implement these methods on oriented object detection under the same semi-supervised settings for a fair comparison.

Partially-annotated. We first conduct a performance comparison between our proposed method and competing semi-supervised methods on the DOTA and DIOR-R datasets under the partially-annotated setting. The results are shown in Tab. 1 and Tab. 2. When given different ratios of images annotated with box annotations, the proposed method consistently surpasses competing semi-supervised methods to a significant extent. In particular, our proposed method outperforms Group R-CNN [50] under the same weakly semi-supervised setting, demonstrating the effectiveness of our proposed method in utilizing readily available point-annotated images for improved oriented object detection. In addition, given an ideal average annotation time of 7s [9] or 0.8∼0.9s [6] to annotate a box or a point and average instances of 67.10 [9] per image on DOTA-v1.0, our proposed method trained on 20% of annotated data with about 107 annotation hours performs comparably to other settings trained on 50% of annotated data with about 183 annotation hours, confirming the effectiveness of our proposed method in balancing the detection performance and annotation cost.

Configs	DOTA-v1.0	DOTA-v1.5
Baseline	35.90	32.41
+ Pseudo Anno.	43.91	38.11
+ Graph Match.	48.78	41.89
+ Dist. Match.	50.22	44.11
+ Weakly Sup.	53.85	46.29

Table 3. Ablation study of different components in our proposed weakly semi-supervised learning method.

$\mathcal{L}_g^v$	$\mathcal{L}_g^e$	DOTA-v1.0	DOTA-v1.5
		48.53	40.98
✓		50.51	42.34
	✓	51.10	43.99
✓	✓	53.85	46.29

Table 4. Performance comparison of the base model trained with different combinations of vertex matching and edge matching.

Fully-annotated. We also demonstrate the success of our proposed method to further improve the detection performance when sufficient annotated images are available, as shown in Tab. 1. With additional point-annotated data, our proposed method achieves a significant performance improvement compared to the semi-supervised methods and the weakly semi-supervised Group R-CNN method, demonstrating the superiority of the proposed method in transferring ambiguous point information to reliable object-level semantic knowledge for enhanced detection performance.

4.4. Ablation Study

We further conduct extensive experiments to highlight the contributions of each component in our proposed method. Unless otherwise stated, we perform the ablation experiments when 5% of fully-annotated images are provided.

Pseudo Supervision. As shown in Tab. 3, the performance of the base model, trained on 5% of fully-annotated images, drops significantly on the validation set of both the DOTA-v1.0 and DOTA-v1.5 datasets. We incorporate the point-annotated images, with only the pseudo supervision loss on the pseudo annotations of highest scores on the annotated points, into the training of the base model, which contributes to limited performance improvements compared to the baseline model. With the proposed Rotation-Modulated Relational Graph Matching method, the base model is forced to focus more on the spatial and semantic knowledge centered on the annotated points, leading to a further improvement in detection performance. Collaborating with the relational graph matching method, the Relational Rank Distribution Matching method concentrates on the distillation of reliable rank distribution over classification and regression from the teacher model to the student model. To handle the difficult points that both models are confused about, we introduce weakly supervised learning to impose positive signals on the low-quality clusters, and alleviate the problem of inaccurate occupancy of the regression outputs on annotated points. As a result, all these components are required to achieve a significant performance gain over the baseline model.

Rotation-Modulated Relational Graph Matching. We next explore the effectiveness of matching vertices and edges in the relational graph between different models, as shown in Tab. 4. Matching on vertices of RRoI features allows the student model to closely follow the feature distribution from the teacher model, while matching on edges of relations between RRoI centered on annotated points encourages the student model to be aware of the precise relational estimation from the teacher model. In addition, modulated orientation differences as weighting factors focus the relation recognition on the predictions with a greater deviation of orientation estimation between different models. Therefore, the base model is enhanced to a significant extent with matching on both vertices and edges in the relational graph.

Relational Rank Distribution Matching. We also adopt different combinations of rank distribution matching on classification and regression, the results of which are shown in Tab. 5. Matching on both rank distributions facilitates the distillation of the relational information between clustered predictions from the teacher model to the student model, leading to a further performance gain over the base model.

Weakly Supervised Learning. We also analyze the cases when implementing different combinations of weakly supervised learning methods on classification or regression in Tab. 6. Weakly supervised learning can well alleviate the inaccurate classification outputs, in which classification scores of all proposals clustered around the difficult points are underrated by both models, and regression outputs, in which the predicted boxes have inappropriate occupancies with annotated points. Both weakly supervised methods are effective for further improving the detection performance.

$\mathcal{L}_{rm}^{cls}$	$\mathcal{L}_{rm}^{reg}$	DOTA-v1.0	DOTA-v1.5
		49.58	42.12
✓		51.84	44.18
	✓	51.73	45.10
✓	✓	53.85	46.29

Table 5. Performance comparison of the base model trained with different rank distribution matching settings.

$\mathcal{L}_{ws}^{cls}$	$\mathcal{L}_{ws}^{reg}$	DOTA-v1.0	DOTA-v1.5
		50.22	44.11
✓		51.97	45.51
	✓	51.62	45.17
✓	✓	53.85	46.29

Table 6. Performance comparison of the base model trained with different weakly supervised learning methods.

5. Conclusion

In this work, we propose a weakly semi-supervised learning framework for oriented object detection, in which we introduce point-annotated data to improve the base detector when limited annotated data are given. To alleviate the ambiguity from annotated points in depicting the oriented objects, we propose a Rotation-Modulated Relational Graph Matching method to align the relations of proposals centered on annotated points between the teacher and student models under augmented views. In addition, we further propose a Relational Rank Distribution Matching method to distill the rich relation information between predictions over annotated points from the teacher model to the student model by matching rank distributions on classification and regression. Finally, we introduce weakly supervised learning on difficult points with inaccurate classification and regression outputs through the weak supervision of the categorical and location information from these points. With our proposed weakly semi-supervised method, the resulting model achieves significant performance gains on multiple challenging datasets.

Acknowledgement

This work was supported in part by the Research Grants Council of the Hong Kong Special Administration Region (Project No. CityU 11206622), in part by the National Natural Science Foundation of China (Project No. 62072189), in part by the GuangDong Basic and Applied Basic Research Foundation (Project No. 2022A1515011160), and in part by TCL Science and Technology Innovation Fund (Project No. 20231752).

References

[1] Eric Arazo, Diego Ortego, Paul Albert, Noel E O'Connor, and Kevin McGuinness. Pseudo-labeling and confirmation bias in deep semi-supervised learning. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2020. 3

[2] Binbin Chen, Weijie Chen, Shicai Yang, Yunyi Xuan, Jie Song, Di Xie, Shiliang Pu, Mingli Song, and Yueting Zhuang. Label matching semi-supervised object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14381–14390, 2022. 3

[3] Changrui Chen, Kurt Debattista, and Jungong Han. Semi-supervised object detection via virtual category learning. In *European Conference on Computer Vision (ECCV)*, 2022. 3

[4] Liangyu Chen, Tong Yang, Xiangyu Zhang, Wei Zhang, and Jian Sun. Points as queries: Weakly semi-supervised object detection by points. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8823–8832, 2021. 3

[5] Pengfei Chen, Xuehui Yu, Xumeng Han, Najmul Hassan, Kai Wang, Jiachen Li, Jian Zhao, Humphrey Shi, Zhenjun Han, and Qixiang Ye. Point-to-box network for accurate object detection via single point supervision. In *European Conference on Computer Vision*, pages 51–67. Springer, 2022. 1

[6] Bowen Cheng, Omkar Parkhi, and Alexander Kirillov. Pointly-supervised instance segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2617–2626, 2022. 7

[7] Gong Cheng, Jiabao Wang, Ke Li, Xingxing Xie, Chunbo Lang, Yanqing Yao, and Junwei Han. Anchor-free oriented proposal generator for object detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–11, 2022. 6

[8] Jian Ding, Nan Xue, Yang Long, Gui-Song Xia, and Qikai Lu. Learning roi transformer for oriented object detection in aerial images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2849–2858, 2019. 2, 3

[9] Jian Ding, Nan Xue, Gui-Song Xia, Xiang Bai, Wen Yang, Michael Ying Yang, Serge Belongie, Jiebo Luo, Mihai Datcu, Marcello Pelillo, et al. Object detection in aerial images: A large-scale benchmark and challenges. *IEEE transactions on pattern analysis and machine intelligence*, 44(11):7778–7796, 2021. 6, 7

[10] Xiaoxu Feng, Xiwen Yao, Gong Cheng, and Junwei Han. Weakly supervised rotation-invariant aerial object detection network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14146–14155, 2022. 1

[11] Xiaoxu Feng, Xiwen Yao, Hui Shen, Gong Cheng, Bin Xiao, and Junwei Han. Learning an invariant and equivariant network for weakly supervised object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023. 1

[12] Yongtao Ge, Qiang Zhou, Xinlong Wang, Chunhua Shen, Zhibin Wang, and Hao Li. Point-teaching: Weakly semi-supervised object detection with point annotations. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(1): 667–675, 2023. 3

[13] Qiushan Guo, Yao Mu, Jianyu Chen, Tianqi Wang, Yizhou Yu, and Ping Luo. Scale-equivalent distillation for semi-supervised object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14522–14531, 2022. 3

[14] Zonghao Guo, Chang Liu, Xiaosong Zhang, Jianbin Jiao, Xiangyang Ji, and Qixiang Ye. 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*, pages 8792–8801, 2021. 3

[15] Jiaming Han, Jian Ding, Jie Li, and Gui-Song Xia. Align deep features for oriented object detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–11, 2021. 2

[16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 6

[17] Shitian He, Huanxin Zou, Yingqian Wang, Boyang Li, Xu Cao, and Ning Jing. Learning remote sensing object detection with single point supervision. *IEEE Transactions on Geoscience and Remote Sensing*, 2024. 1

[18] Liping Hou, Ke Lu, Jian Xue, and Yuqiu Li. Shape-adaptive selection and measurement for oriented object detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36 (1):923–932, 2022. 3

[19] Wei Hua, Dingkang Liang, Jingyu Li, Xiaolong Liu, Zhikang Zou, Xiaoqing Ye, and Xiang Bai. Sood: Towards semi-supervised oriented object detection. *arXiv preprint arXiv:2304.04515*, 2023. 1, 3, 6, 7

[20] Gang Li, Xiang Li, Yujie Wang, Yichao Wu, Ding Liang, and Shanshan Zhang. Pseco: Pseudo labeling and consistency training for semi-supervised object detection. In *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part IX*, pages 457–472. Springer, 2022. 3

[21] Gang Li, Xiang Li, Yujie Wang, Wu Yichao, Ding Liang, and Shanshan Zhang. Dtg-ssod: Dense teacher guidance for semi-supervised object detection. *Advances in Neural Information Processing Systems*, 35:8840–8852, 2022. 3, 5

[22] Gang Li, Xiang Li, Yujie Wang, Shanshan Zhang, Yichao Wu, and Ding Liang. Knowledge distillation for object detection via rank mimicking and prediction-guided feature imitation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 1306–1313, 2022. 5

[23] Jiaming Li, Xiangru Lin, Wei Zhang, Xiao Tan, Yingying Li, Junyu Han, Errui Ding, Jingdong Wang, and Guanbin Li. Gradient-based sampling for class imbalanced semi-supervised object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 16390–16400, 2023. 1, 3

[24] Ke Li, Gang Wan, Gong Cheng, Liqiu Meng, and Junwei Han. Object detection in optical remote sensing images: A survey and a new benchmark. *ISPRS journal of photogrammetry and remote sensing*, 159:296–307, 2020. 6

[25] Wentong Li, Yijie Chen, Kaixuan Hu, and Jianke Zhu. Oriented reppoints for aerial object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1829–1838, 2022. 3

[26] Xiang Li, Wenhai Wang, Lijun Wu, Shuo Chen, Xiaolin Hu, Jun Li, Jinhui Tang, and Jian Yang. Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection. *Advances in Neural Information Processing Systems*, 33:21002–21012, 2020. 5

[27] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117–2125, 2017. 6

[28] Chang Liu, Weiming Zhang, Xiangru Lin, Wei Zhang, Xiao Tan, Junyu Han, Xiaomao Li, Errui Ding, and Jingdong Wang. Ambiguity-resistant semi-supervised learning for dense object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 15579–15588, 2023. 3

[29] Liang Liu, Boshen Zhang, Jiangning Zhang, Wuhao Zhang, Zhenye Gan, Guanzhong Tian, Wenbing Zhu, Yabiao Wang, and Chengjie Wang. Mixteacher: Mining promising labels with mixed scale teacher for semi-supervised object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7370–7379, 2023. 1, 3

[30] Yen-Cheng Liu, Chih-Yao Ma, Zijian He, Chia-Wen Kuo, Kan Chen, Peizhao Zhang, Bichen Wu, Zsolt Kira, and Peter Vajda. Unbiased teacher for semi-supervised object detection. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2021. 3

[31] Yifan Pu, Yiru Wang, Zhuofan Xia, Yizeng Han, Yulin Wang, Weihao Gan, Zidong Wang, Shiji Song, and Gao Huang. Adaptive rotated convolution for rotated object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023. 2

[32] Wen Qian, Xue Yang, Silong Peng, Junchi Yan, and Yue Guo. Learning modulated loss for rotated object detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(3):2458–2466, 2021. 2

[33] Wen Qian, Xue Yang, Silong Peng, Junchi Yan, and Yue Guo. Learning modulated loss for rotated object detection. In *Proceedings of the AAAI conference on artificial intelligence*, pages 2458–2466, 2021. 4

[34] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28, 2015. 2, 6, 7

[35] Yihe Tang, Weifeng Chen, Yijun Luo, and Yuting Zhang. Humble teachers teach better students for semi-supervised object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3132–3141, 2021. 3

[36] Antti Tarvainen and Harri Valpola. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. *Advances in neural information processing systems*, 30, 2017. 2, 3

[37] Zhi Tian, Chunhua Shen, Hao Chen, and Tong He. Fcos: Fully convolutional one-stage object detection. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9627–9636, 2019. 7

[38] Mengde Xu, Zheng Zhang, Han Hu, Jianfeng Wang, Lijuan Wang, Fangyun Wei, Xiang Bai, and Zicheng Liu. End-to-end semi-supervised object detection with soft teacher. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3060–3069, 2021. 3, 7

[39] Yongchao Xu, Mingtao Fu, Qimeng Wang, Yukang Wang, Kai Chen, Gui-Song Xia, and Xiang Bai. Gliding vertex on the horizontal bounding box for multi-oriented object detection. *IEEE transactions on pattern analysis and machine intelligence*, 43(4):1452–1459, 2020. 3

[40] Xue Yang and Junchi Yan. Arbitrary-oriented object detection with circular smooth label. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VIII 16*, pages 677–694. Springer, 2020. 2

[41] Xue Yang, Jirui Yang, Junchi Yan, Yue Zhang, Tengfei Zhang, Zhi Guo, Xian Sun, and Kun Fu. Scrdet: Towards more robust detection for small, cluttered and rotated objects. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 8232–8241, 2019. 2

[42] Xue Yang, Liping Hou, Yue Zhou, Wentao Wang, and Junchi Yan. Dense label encoding for boundary discontinuity free rotation detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 15819–15829, 2021. 2

[43] Xue Yang, Junchi Yan, Qi Ming, Wentao Wang, Xiaopeng Zhang, and Qi Tian. Rethinking rotated object detection with gaussian wasserstein distance loss. In *International Conference on Machine Learning*, pages 11830–11841. PMLR, 2021. 2

[44] Xue Yang, Xiaojiang Yang, Jirui Yang, Qi Ming, Wentao Wang, Qi Tian, and Junchi Yan. Learning high-precision bounding box for rotated object detection via kullback-leibler divergence. *Advances in Neural Information Processing Systems*, 34:18381–18394, 2021. 2

[45] Xue Yang, Junchi Yan, Wenlong Liao, Xiaokang Yang, Jin Tang, and Tao He. Scrdet++: Detecting small, cluttered and rotated objects via instance-level feature denoising and rotation loss smoothing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2):2384–2399, 2022. 2

[46] Xue Yang, Gefan Zhang, Wentong Li, Yue Zhou, Xuehui Wang, and Junchi Yan. H2rbox: Horizontal box annotation is all you need for oriented object detection. In *The Eleventh International Conference on Learning Representations*, 2022. 2

[47] Xue Yang, Gefan Zhang, Xiaojiang Yang, Yue Zhou, Wentao Wang, Jin Tang, Tao He, and Junchi Yan. Detecting rotated objects as gaussian distributions and its 3-d generalization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. 2

[48] Xue Yang, Yue Zhou, Gefan Zhang, Jirui Yang, Wentao Wang, Junchi Yan, XIAOPENG ZHANG, and Qi Tian. The KFiou loss for rotated object detection. In *The Eleventh In-*

ternational Conference on Learning Representations, 2023. 2

[49] Dingyuan Zhang, Dingkang Liang, Zhikang Zou, Jingyu Li, Xiaoqing Ye, Zhe Liu, Xiao Tan, and Xiang Bai. A simple vision transformer for weakly semi-supervised 3d object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 8373–8383, 2023. 3

[50] Shilong Zhang, Zhuoran Yu, Liyang Liu, Xinjiang Wang, Aojun Zhou, and Kai Chen. Group r-cnn for weakly semi-supervised object detection with points. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9417–9426, 2022. 2, 3, 7

[51] Hongyu Zhou, Zheng Ge, Songtao Liu, Weixin Mao, Zeming Li, Haiyan Yu, and Jian Sun. Dense teacher: Dense pseudo-labels for semi-supervised object detection. In *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part IX*, pages 35–50. Springer, 2022. 3, 7

[52] Qiang Zhou, Chaohui Yu, Zhibin Wang, Qi Qian, and Hao Li. Instant-teaching: An end-to-end semi-supervised object detection framework. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4081–4090, 2021. 3

[53] Yue Zhou, Xue Yang, Gefan Zhang, Jiabao Wang, Yanyi Liu, Liping Hou, Xue Jiang, Xingzhao Liu, Junchi Yan, Chengqi Lyu, et al. Mmrotate: A rotated object detection benchmark using pytorch. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 7331–7334, 2022. 6