

Training-Based Model Refinement and Representation Disagreement for Semi-Supervised Object Detection

Seyed Mojtaba Marvasti-Zadeh
University of Alberta
seyedmoj@ualberta.ca

Nilanjan Ray
University of Alberta
nray1@ualberta.ca

Nadir Erbilgin
University of Alberta
erbilgin@ualberta.ca

Abstract

Semi-supervised object detection (SSOD) aims to improve the performance and generalization of existing object detectors by utilizing limited labeled data and extensive unlabeled data. Despite many advances, recent SSOD methods are still challenged by inadequate model refinement using the classical exponential moving average (EMA) strategy, the consensus of Teacher-Student models in the latter stages of training (i.e., losing their distinctiveness), and noisy/misleading pseudo-labels. This paper proposes a novel training-based model refinement (TMR) stage and a simple yet effective representation disagreement (RD) strategy to address the limitations of classical EMA and the consensus problem. The TMR stage of Teacher-Student models optimizes the lightweight scaling operation to refine the model's weights and prevent overfitting or forgetting learned patterns from unlabeled data. Meanwhile, the RD strategy helps keep these models diverged to encourage the student model to explore additional patterns in unlabeled data. Our approach can be integrated into established SSOD methods and is empirically validated using two baseline methods, with and without cascade regression, to generate more reliable pseudo-labels. Extensive experiments demonstrate the superior performance of our approach over state-of-the-art SSOD methods. Specifically, the proposed approach outperforms the baseline Unbiased-Teacher-v2 (& Unbiased-Teacher-v1) method by an average mAP margin of 2.23, 2.1, and 3.36 (& 2.07, 1.9 and 3.27) on COCO-standard, COCO-additional, and Pascal VOC datasets, respectively.

1. Introduction

Object detection has experienced significant progress owing to the availability of large-scale benchmark datasets containing pairs of class labels and bounding boxes for various objects within images. However, collecting and accurately annotating object detection datasets is extremely

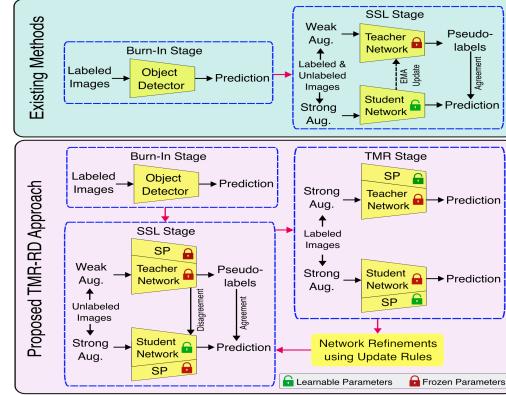


Figure 1. An illustration of the difference between existing SSOD methods and the proposed approach. Existing methods use Teacher-Student models (initialized in the Burn-In stage) where the student model is trained with pseudo-labels generated by the teacher model, and teacher model weights are updated using the classical EMA strategy. In the proposed approach, during the SSL stage, the teacher model weights and scaling operation parameters (denoted as SPs) remain frozen during the training of the student model. After completing N iterations, the weights of the models are frozen, and the TMR stage optimizes lightweight SPs for N' iterations, followed by adaptive refinement of model weights using updating rules. The SSL and TMR stages are iteratively applied until the completion of the training process. Best viewed in color.

expensive, time-consuming, and labor-intensive due to the lack of domain experts, limited resources, and the complicated nature of the problem. Meanwhile, acquiring a large amount of unlabeled data is relatively easy and provides valuable insights into the data distribution from which robust representations (under various transformations) can be learned [19]. Hence, leveraging unlabeled data for training object detectors has become increasingly popular due to its potential to significantly reduce annotation costs & efforts while improving model performance and generalization. *Semi-supervised object detection* (SSOD) aims to harness unsupervised information in scenarios where labeled data is limited. Most recent SSOD methods rely on self-training techniques, in which the pseudo-labels generated from the teacher model(s) are utilized to train the student

Table 1. Comparative review of SSOD methods. Best viewed with zoom-in.

Method	Motivation(s)	Backbone(s)	Detector(s)	Teacher Update	Weak Augmentation(s)	Strong Augmentation(s)
CSD [7]	Efficient training process (than self-training)	ResNet-101	SSD, RCFN	×	Horizontal flip	×
STAC [22]	Combining self-training & consistency regularization	ResNet-50-FPN	Faster-RCNN	×	Color/geometric transformations, cutout	
ISM7 [30]	Detecting training iteration discrepancies	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip	Color jitter
Soft-Teacher [29]	More reliable pseudo-labels	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip	Scale/solarize/brightness/contrast/sharpness jitters, translation, rotate, shift, cutout
Humble-Teacher [24]	More reliable pseudo-labels, Teacher ensemble	ResNet-50-FPN, ResNet-152-FPN	Faster-RCNN, Cascade-RCNN	Classical EMA	Resize, flip	color/sharpness/contrast jitters, Gaussian noise, cutout
Instant-Teaching [35]	More reliable pseudo-labels, Confirmation bias	ResNet-50-FPN	Faster-RCNN	Co-rectify scheme	×	Color/geometric transformations, cutout, mixup, mosaic
Unbiased-Teacher-v1 [15]	Class imbalance, Pseudo-label bias	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip	Color jitter, grayscale, Gaussian blur, cutout
Active-Teacher [18]	More reliable pseudo-labels, Data initialization	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip, color jitter, grayscale, Gaussian blur, cutout	
AGCo [13]	Class imbalance, Bias-corrected pseudo-labels	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip, crop	Color jitter, Gaussian blur, cutout
CAPI [1]	Class imbalance, Confirmation bias	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip	Color jitter, grayscale, Gaussian blur, cutout
SED [6]	Class imbalance, Large object size variance	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip, multi-scale	Color jitter, grayscale, Gaussian blur, cutout
Label-Match [3]	Label mismatch, Confirmation bias	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip	Color/geometric transformations/contrast/sharpness jitters, translation, rotate, shift, cutout
MA-GCP [8]	Relation between labeled & unlabeled data	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip	Scale/solarize/brightness/contrast jitters, Gaussian blur, cutout
MUM [8]	More complex data augmentation	ResNet-50-FPN, SwinTransformer	Faster-RCNN	Classical EMA	Mixing image tiles, color jitter, grayscale, Gaussian blur, cutout	
Unbiased-Teacher-v2 [16]	Anchor-free detectors, Misleading pseudo-labels	ResNet-50-FPN	Faster-RCNN, FCOS	Classical EMA	Horizontal flip	Color/scale jitters, grayscale, Gaussian blur, cutout
Diverse-Learner [12]	Maintain networks distinctiveness	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip, random size	Random erasing, rotation, color jitters, etc.
PseCo [10]	Noisy pseudo-labels, Scale-invariant learning	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip	Scale/solarize/brightness/contrast/sharpness jitters, translation, rotate, shift, cutout
VC-Learner [4]	Confirmation bias, Confusing samples	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip	Color/scale jitters, grayscale, Gaussian blur, cutout
De-biased Teacher [26]	IoU matching bias	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip	Color jitter, grayscale, Gaussian blur, cutout
Pseudo-Polish [33]	Limited teacher model generalization	ResNet-50-FPN	Faster-RCNN	Classical EMA	Horizontal flip	Scale/solarize/brightness/contrast/sharpness jitters, translation, rotate, shift, cutout

model(s) when weak & strong augmentations of images are used as inputs (see Fig. 1 & Table 1). While these methods enforce the consistency (or agreement) of Teacher-Student predictions to train the student model, the classical *exponential moving average* (EMA) strategy is adopted to evolve the teacher model progressively. This strategy gradually refines the weights of the teacher model to improve the accuracy of pseudo-labels & resiliency to noisy weights of the student model [24] as well as alleviate the adverse effects of pseudo-labeling bias [15].

Despite the progress made in the SSOD, training Teacher-Student models still faces three major challenges. The first challenge pertains to applying the classical EMA strategy using manually defined smoothing coefficients. It can lead to two potential issues when refining the weights of the teacher model: i) insensitivity to important changes in weights of the student model due to excessive reliance on the initialized detector, and ii) sub-optimal performance because constant coefficients may not be effective for all refinement steps, resulting in weaker pseudo-labels. To address these limitations, we propose a novel *training-based model refinement* (TMR) stage that adaptively refines the weights of Teacher-Student models (see Fig. 1). This stage is added to the commonly used training stages of i) pre-training on limited labeled data (or Burn-In stage), and ii) *semi-supervised learning* (SSL) stage in SSOD methods. Inspired by *meta-transfer learning* (MTL) [23], we optimize the lightweight scaling operation corresponding to learnable parameters of Teacher-Student models to effectively aggregate information from labeled and unlabeled data. The models' weights can then be dynamically refined using the introduced update rules to ensure the teacher model is up-to-date and reduce the effect of noisy pseudo-labels on the student model without the risk of overfitting or forgetting the patterns learned from unlabeled data.

The second challenge is the consensus problem (i.e., losing the distinctiveness of Teacher-Student models) at the latter stages of the training procedure when two models become almost identical. This is also derived from the classical EMA strategy, leading to teacher weights being close to student ones as training progresses. Accordingly, both models generate similar predictions and make it difficult for

the teacher model to extract helpful information from unlabeled data for supervising the student model. To alleviate this issue, we propose a simple yet effective *representation disagreement* (RD) strategy that incorporates the asymmetric *Kullback-Leiber* (KL) divergence between the semantic representation of Teacher-Student models to prevent early convergence. This strategy aims to increase model divergence by encouraging the student model to explore more robust representations, learn complementary information, and reduce the memorization effect of easy samples.

The last challenge is to address noisy/misleading pseudo-labels, which can impede accurate model optimization leading to ineffective learning from unlabeled data and slow convergence. As shown in Table 1, extensive efforts have been made to provide more reliable pseudo-labels. Although most of these methods rely on the well-established Faster-RCNN detector [20], it can be optimal solely for detecting objects at a single-quality level due to the adversarial nature of producing noisy boxes or assembling inadequate positive proposals. Following the Humble-Teacher [24], we integrate cascade regression into our baselines (i.e., [15,16]) to generate more reliable pseudo-labels while reducing the overfitting problem. Note that we consider this integration *not as our contribution* but rather as a means to highlight the versatility of our proposed TMR-RD approach with various base detectors. To conduct a fair comparison, our empirical experiments (Sec. 4) involve using two SSOD baselines with or without cascade regression.

The main contributions are summarized as follows: (1) A novel TMR stage is proposed to dynamically refine the weights of Teacher/Student models in SSOD frameworks. It can learn the lightweight scaling operation corresponding to the learnable parameters of the Teacher-Student models, enabling fast convergence without the risk of overfitting or forgetting patterns learned from unlabeled data. To the best of our knowledge, this is the first work on the dynamic refinement of Teacher-Student learning, with the potential to inspire further research in its applications (e.g., knowledge exchange, domain adaptation, etc.). (2) A simple yet effective RD strategy is proposed to alleviate the consensus problem of Teacher-Student models with progress in training. This strategy prevents the models from converging

ing too early, allowing better generalization through more exploring underlying patterns in unlabeled data. (3) Extensive empirical evaluations and ablation analyses demonstrate superior performance and generalizability of the proposed approach, which can be incorporated into existing SSOD methods to boost their performance.

2. Related Work

2.1. Classical Exponential Moving Average (EMA)

While existing SSOD methods tackle various challenges (e.g., efficient training [7, 22], detection discrepancies [30], localization certainty [11], prediction consistency [9], object size [6], data augmentation [8], confusing samples [4], and reliability of pseudo-labels [18, 24, 29, 35]), they mostly employ the classical EMA strategy [25] to refine the teacher’s model weights (see Table 1). This strategy updates this model by averaging its weights to ensure more accurate predictions from the teacher model than the student one. It can minimize the adverse effects of imbalanced and noisy labels through mutually reinforcing pseudo-labeling and the detection training steps by

$$\begin{aligned} \theta_t^n &\leftarrow \alpha \theta_t^{n-1} + (1 - \alpha) \theta_s^{n-1} \\ \Rightarrow \theta_t^n &\leftarrow \alpha^n \theta_t^0 + (1 - \alpha) \sum_{k=0}^{n-1} \alpha^{n-1-k} \theta_s^k \end{aligned} \quad (1)$$

in which θ_t^0 , θ_t^n , and θ_s^{n-1} represent the initialized detector (trained on limited labeled data) at time step $n = 0$, the parameters of the teacher model at n -th time step, and the parameters of the student model at $(n - 1)$ -th time step, respectively. In addition, α is a manually-defined smoothing coefficient (i.e., EMA decay) often set to 0.999 in SSOD methods so that the teacher model can benefit from a long memory while it assumes the student model improves slowly. The limitations of this strategy (also presented in Sec. 1) have recently been discussed in the Diverse-Learner [12], which uses two Teacher-Student pairs and multi-threshold classification loss to alleviate the associated drawbacks, maintain distinctiveness between models, and improve pseudo-labels. However, it still uses the classical EMA strategy to refine its teacher models. In this paper, we propose a novel approach for refining Teacher-Student models by optimizing the lightweight scaling operation corresponding to the learnable model parameters, addressing the limitations of the classical EMA strategy.

2.2. Model Robustness with Reliable Pseudo-labels

Recent advancements have been primarily directed at enhancing the reliability of pseudo-labels (see Table 1). These improvements encompass several approaches, such as: i) selecting pseudo-boxes with higher scores [29], ii) employing instant pseudo-labeling and simultaneous training of two models [35], iii) utilizing soft pseudo-labels and

teacher ensembles [24], iv) selecting the most optimal labeled examples [18], v) employing a memory module and a two-stage pseudo-label filtering [32], vi) using a redistributed mean teacher and proposal self-assignment [3], vii) applying prediction-guided label assignment and positive-proposal consistency voting [10], viii) utilizing Teacher-Student mutual learning [15], ix) assessing relative model uncertainties [16], x) developing two differently structured pseudo-label polishing networks [33], and xi) directly generating proposals for consistency regularization between perturbed image pairs [26]. These methods aim to train robust models by imposing an agreement strategy (or consistency) between the predictions of Teacher-Student models using labeled and unlabeled data. However, relying on similar predictions and classical EMA can lead to losing model distinctiveness and early convergence of models. To address this, we introduce a simple yet effective RD strategy, which helps maintain divergence between the models, enabling the exploration of additional patterns in unlabeled data and promoting model generalization.

Aside from that, existing methods typically use a confidence threshold to determine the number of bounding boxes. Decreasing this threshold increases the information mined but can introduce noisy pseudo-labels, reducing detection performance. Conversely, using a higher threshold produces a limited number of high-quality boxes, ignoring some objects during the SSL stage. In addition to the Faster-RCNN detector, we assess our TMR-RD approach with the Faster-RCNN equipped with cascade regression, which relies on a multi-stage architecture featuring specialized regressors for precise localization. This evaluation serves to highlight the versatility of our approach, indicating its potential to improve the overall performance and generalization of existing methods.

3. Proposed Approach: TMR-RD

In this section, we present the TMR-RD approach, which can be integrated into most existing SSOD methods employing Teacher-Student models and the classical EMA strategy. We utilize the well-designed Unbiased-Teacher-v1 [15] and Unbiased-Teacher-v2 [16], both with and without cascade regression, as the baseline SSOD methods to demonstrate our approach’s effectiveness. For the sake of generality, we will describe the integration of our approach into the Unbiased-Teacher-v2 (named TMR-RD-v2) with the base detector using cascade regression, as shown in Fig. 2. Still, our approach can be readily adapted to the Unbiased-Teacher-v1 without the additional branch of localization uncertainty (named TMR-RD-v1) and work with the Faster-RCNN detector (single-stage regression). An overview of our approach is shown in Algorithm 1. The experimental results for various models are provided in Sec. 4.

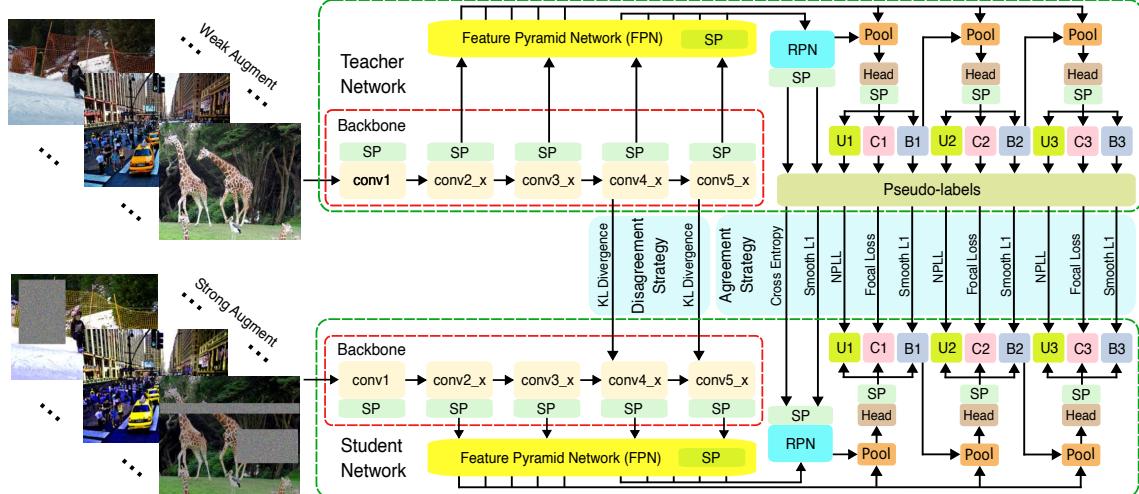


Figure 2. Overview of integrating the TMR-RD approach and cascade regression into the baseline Unbiased-Teacher-v2 [16]. SPs, U, C, B, and NPLL represent scaling operation parameters, localization uncertainty, ROI classification, boundary distance, and negative power log-likelihood loss, respectively. For comprehensive and fair comparisons, we also evaluate our models w/o the uncertainty prediction head (e.g., Unbiased-Teacher-v1 [15]) and w/o cascade regression (i.e., Faster-RCNN).

3.1. Training-based Model Refinement (TMR)

The proposed TMR stage is inspired by the MTL [23] that adapts large-scale trained base classification models to new tasks with limited data using two lightweight neuron operations of scaling and shifting. However, the proposed TMR is designed following the EMA equation (see Eq. 1) that merely requires scaling coefficients to refine the model weights. The proposed approach comprises three stages of Burn-In, SSL, and TMR (see Fig. 1), of which the first is the same as that in the baseline methods [15, 16]. The Burn-In stage provides a good initialization by training the base object detector on the available labeled data. Then, the initialized detector is duplicated into the Teacher-Student models so that pseudo-labels are generated by the Teacher and used to train the Student during the SSL stage. However, the weights of the Teacher are not updated using the classical EMA. Instead, we freeze neuron weights from the SSL stage after completing N iterations and initiate the TMR stage to learn a set of lightweight scaling operations associated with the trained parameters of the Teacher-Student models. By completing N' iterations, the proposed TMR can promote the progressive transfer of knowledge from the continually learning student model to the teacher one, thereby aggregating information more effectively and improving the generation of pseudo-labels.

In the TMR stage, the scaling operation is denoted as Ω_i , where $i \in (t, s)$ refers to the teacher or student model. The MTL [23] defines the scaling and shifting operations exclusively for the feature extractor (i.e., backbone layers) and updates them after optimizing a temporal classifier (i.e., as per a new task). However, we apply scaling operations to all frozen neuron weights (see Fig. 2) of Teacher/Student by

$S(\hat{x}; \theta_i; \Omega_i) = (\theta_i \odot \Omega_i) \hat{x}_k$, where $\hat{\mathcal{D}} = \{\hat{x}_k, \hat{y}_k\}_{k=1}^{N_{sup}}$ and \odot are the strongly-augmented labeled data and the element-wise multiplication, respectively. Then, the TMR loss is defined as

$$\mathcal{L}_{\text{TMR}} = \lambda_t \mathcal{L}_{sc}([\theta_t; \Omega_t]) + \lambda_s \mathcal{L}_{sc}([\theta_s; \Omega_s]), \quad (2)$$

$$\begin{aligned} \mathcal{L}_{sc}([\theta_i; \Omega_i]) = & \sum_{j \in \hat{\mathcal{D}}} \mathcal{L}_{cls}^{\text{rpn}} \left(\mathcal{C}_{[\theta_i; \Omega_i]}^0(\hat{x}_j^0), \hat{y}_j^0 \right) \\ & + \mathcal{L}_{reg}^{\text{rpn}} \left(\mathcal{B}_{[\theta_i; \Omega_i]}^0(\hat{x}_j^0), \hat{y}_j^0 \right) + \mathcal{L}_{cls}^{\text{roi}} \left(\mathcal{C}_{[\theta_i; \Omega_i]}^1(\hat{x}_j^1), \hat{y}_j^1 \right) \\ & + \mathcal{L}_{reg}^{\text{roi}} \left(\mathcal{B}_{[\theta_i; \Omega_i]}^1(\hat{x}_j^1), \hat{y}_j^1 \right) + \mathcal{L}_{cls}^{\text{roi}} \left(\mathcal{U}_{[\theta_i; \Omega_i]}^1(\hat{x}_j^1), \hat{y}_j^1 \right) \\ & + \mathcal{L}_{reg}^{\text{roi}} \left(\mathcal{C}_{[\theta_i; \Omega_i]}^2(\hat{x}_j^2), \hat{y}_j^2 \right) + \mathcal{L}_{reg}^{\text{roi}} \left(\mathcal{B}_{[\theta_i; \Omega_i]}^2(\hat{x}_j^2), \hat{y}_j^2 \right) \\ & + \mathcal{L}_{reg}^{\text{roi}} \left(\mathcal{U}_{[\theta_i; \Omega_i]}^2(\hat{x}_j^2), \hat{y}_j^2 \right) + \mathcal{L}_{cls}^{\text{roi}} \left(\mathcal{C}_{[\theta_i; \Omega_i]}^3(\hat{x}_j^3), \hat{y}_j^3 \right) \\ & + \mathcal{L}_{reg}^{\text{roi}} \left(\mathcal{B}_{[\theta_i; \Omega_i]}^3(\hat{x}_j^3), \hat{y}_j^3 \right) + \mathcal{L}_{reg}^{\text{roi}} \left(\mathcal{U}_{[\theta_i; \Omega_i]}^3(\hat{x}_j^3), \hat{y}_j^3 \right), \end{aligned} \quad (3)$$

where \mathcal{L}_{sc} , \mathcal{C}^k , \mathcal{B}^k , and \mathcal{U}^k represent the scaling operation loss, a classifier at stage k , a boundary regressor at stage k , and localization uncertainty at stage k , respectively. Each regression stage is optimized for an IoU threshold τ^k , respectively. In addition, rpn and roi refer to the RPN and ROI-Head branches, respectively. As shown in Fig. 2, we utilize the Cross-Entropy, Smooth-L1, Focal loss, and negative power log-likelihood loss (NPLL) as in the baseline methods [15, 16] for our models to provide fair comparisons. Next, the scaling operation weights associated with learnable parameters of the Teacher/Student are updated by

$$\Omega_i =: \Omega_i - \gamma \nabla_{\Omega_i} \mathcal{L}_{\text{TMR}}, \quad (4)$$

in which γ denotes the learning rate. Optimizing the lightweight scaling operation while keeping the large-scale

trained weights of Teacher-Student models unchanged allows for fast convergence while reducing the overfitting risk. After N' iterations, the models are refined using our updating rules as

$$\theta_t^n = \frac{\Omega_t}{\Omega_t + \Omega_s} \theta_t^{n-1} + \frac{\Omega_s}{\Omega_t + \Omega_s} \theta_s^{n-1}, \quad (5)$$

$$\theta_s^n = (1 - \frac{\Omega_t}{\Omega_t + \Omega_s}) \theta_t^{n-1} + (1 - \frac{\Omega_s}{\Omega_t + \Omega_s}) \theta_s^{n-1}. \quad (6)$$

These rules imply that the scaling operation weights stay within the permissible range of zero to one and selectively refine models by the ability to adjust or forget inaccurate model weights more effectively. Besides updating the teacher weights, the student weights are also slightly refined to reduce the impact of noisy pseudo-labels from training this model with potentially misleading pseudo-labels in the SSL stage. The SSL and TMR stages are alternatively continued until the end of the training procedure. This approach makes it possible to efficiently learn EMA coefficients on-the-fly, improving the performance of SSOD methods.

3.2. Representation Disagreement (RD)

The disagreement strategy idea lies in the principles of Co-training [2], where the effectiveness of an ensemble can be improved by keeping divergent classifiers. It has then been extended to train deep networks in the presence of noisy labels (e.g., [17, 31]). These scenarios involve simultaneous training of two deep networks based on the cross-update principle implied by the culture-evolving hypothesis [1], where a network can improve its learning capability when it is aided by signals generated by another network. Although two networks with distinct learning capabilities can distinguish different error categories at the beginning of the training phase, they will progressively converge toward

Algorithm 1 : Proposed TMR-RD Approach

INPUT: Labeled images \mathcal{D}_{sup} , Unlabeled images $\mathcal{D}_{\text{unsup}}$, Detection networks (w/o cascade reg.): Teacher Net($\cdot; \theta_t$) & Student Net($\cdot; \theta_s$)
OUTPUT: Trained Net($\cdot; \theta_t$) & Net($\cdot; \theta_s$)

```

Freeze  $\Omega_t$  ▷ Burn-In Stage
for samples in  $\mathcal{D}_{\text{sup}}$  do
| Evaluate  $\mathcal{L}_{\text{sup}}$  following [15, 16]
| Optimize  $\theta_t$ 
end
Initialize Net( $\cdot; \theta_s$ ) by setting  $\theta_s \leftarrow \theta_t$  ▷ SSL Stage
while not done do
| Freeze  $\Omega_t$ ,  $\Omega_s$ , and  $\theta$ 
| for  $N$  iterations do
| | Samples from  $\mathcal{D}_{\text{unsup}}$  and  $\hat{\mathcal{D}}_{\text{unsup}}$  ▷ Weak & Strong Aug.
| | Evaluate  $\mathcal{L}_{\text{unsup}}$  as in [16] for TMR-RD-v2 (or [15] for TMR-RD-v1)
| | Evaluate  $\mathcal{L}_{\text{RD}}$  using Eq. 7
| | Optimize  $\theta_s$  using Eq. 8
| end
| Freeze  $\theta_t$  and  $\theta_s$  ▷ TMR Stage
| for  $N'$  iterations do
| | Samples from  $\hat{\mathcal{D}}_{\text{sup}}$  ▷ Strong Aug.
| | Evaluate  $\mathcal{L}_{\text{TMR}}$  using Eq. 2 (single-stage regression of Eq. 3)
| | Optimize  $\Omega_t$  and  $\Omega_s$  using Eq. 4
| end
| Refine  $\theta_t$  and  $\theta_s$  using Eq. 5 and Eq. 6
end

```

being close to each other, known as the consensus problem. Hence, the disagreement strategy seeks to alleviate this issue and boost the performance by keeping two networks diverged within the training epochs or slowing the consensus rate between two networks as the number of epochs increases.

Similarly, the lack of distinctiveness is a common problem in self-training SSOD methods employing the EMA strategy, as the weights of the Teacher-Student models become almost identical towards the latter stages of training [12]. To alleviate this problem, we introduce the simple yet effective RD strategy during the SSL stage that encourages the student model to explore further underlying patterns in unlabeled data. It is also motivated by JoCoR [27] and DML [34] methods used for the weakly-supervised learning and knowledge distillation. However, our strategy relies on the representation space to keep the models diverged in contrast with the DML and JoCoR aimed at reducing the diversity between ensemble networks and minimizing the KL divergence between the probabilistic outputs of networks. For the proposed RD strategy, we first compute the probability distributions of semantic representations $f_i \in \mathbb{R}^{C \times H \times W}$ from Teacher/Student models using $p_t = \text{softmax}(f_t(\bar{x}^u))$ and $p_s = \text{softmax}(f_s(\hat{x}^u))$, where \bar{x}^u and \hat{x}^u are the weakly-augmented and strongly-augmented samples from unlabeled data, respectively. Then, the asymmetric KL divergence is computed between the probability distributions based on the supervision of the teacher model as

$$\mathcal{L}_{\text{RD}} = KL(p_s(\hat{x}^u) || p_t(\bar{x}^u)). \quad (7)$$

Following that, the learnable parameters for the student model are updated during the SSL stage by

$$\theta_s \leftarrow \theta_s + \xi \frac{\partial(\lambda_u \mathcal{L}_{\text{unsup}} - \lambda_d \mathcal{L}_{\text{RD}})}{\partial \theta_s} \quad (8)$$

where the learning rate is denoted by ξ , while λ_u , and λ_d control the contribution of the unsupervised and representation disagreement losses, respectively. Here, $\mathcal{L}_{\text{unsup}}$ represents the unsupervised loss of the student model using $\{\hat{x}_k^u, \hat{y}_k\}_{k=1}^{N_{\text{unsup}}}$, in which \hat{y} denotes pseudo-labels from the teacher model. To ensure valid comparisons, we adopted $\mathcal{L}_{\text{unsup}}$ similar to that of our baseline methods [15, 16], both with and without incorporation of cascade regression.

4. Empirical Experiments

This section includes implementation details of the proposed approach, presents state-of-the-art (SOTA) comparisons using the MS-COCO [14] and Pascal VOC [5] datasets, and provides systematic ablation analyses. The COCO dataset consists of the COCO-standard (comprising the train2017 and val2017 splits with $\sim 118k$ and $\sim 5k$ labeled images, respectively) and COCO-additional ($\sim 123k$ unlabeled images) sets. For

the MS-COCO dataset, we evaluated our TMR-RD using two experimental settings, including partially-labeled and fully-labeled data. The partially-labeled data setting involves randomly selecting 1%, 5%, and 10% of the `train2017` split as labeled training data, while the remaining data was treated as unlabeled training data. For the fully-labeled data setting, the entire `train2017` split and the COCO-additional set are utilized as labeled data and unlabeled data, respectively. The evaluations are performed on the `val2017` set using the *mean average precision* (mAP) metric. For the Pascal VOC dataset, the training was conducted using `VOC07-trainval` ($\sim 5k$ images) and `VOC12-trainval` ($\sim 11.5k$ images) as the labeled set and unlabeled set, respectively. The models were then evaluated on the `VOC07-test` set ($\sim 5k$ images) using AP₅₀ and AP_{50:95} (denoted as mAP) metrics.

4.1. Implementation Details

We employ the Faster-RCNN [20] with and without cascade regression implemented in Detectron2 [28] as our base detection framework. The backbones consist of ResNet-50-FPN architecture [13] initialized with the pre-trained Image-Net [21] weights. The Burn-In and SSL (without classical EMA) stages followed the baselines of Unbiased-Teacher-v1 [15] and Unbiased-Teacher-v2 [16] with pre-training for 2k iterations for the COCO-standard, 90k for the COCO-additional, and 30k for the Pascal VOC. Then, the trained detectors were duplicated as Teacher-Student models for further training.

We experimentally adopted cyclic SSL and TMR stages with $N = 4k$ iterations dedicated to the SSL and $N' = 2k$ iterations devoted to the TMR (see Sec. 4.3). For partially-labeled data setting, the models were trained for 269k iterations (i.e., 178k for SSL (as in [15,16]) & 89k for TMR) and 536k iterations (i.e., 356k for SSL & 178k for TMR) using the base detectors of Faster-RCNN without and with cascade regression, respectively. In the fully-labeled data setting, we trained the models with Faster-RCNN without and with cascade regression detectors for 495k and 900k iterations, respectively. Specifically, we used 270k iterations for SSL (as in [15,16]) and 135k iterations for TMR using the Faster-RCNN detector, while with cascade regression, we used 540k iterations for SSL and 270k iterations for TMR. The batch size of 64 was used in training, where 32 labeled & 32 unlabeled images were randomly selected for each batch. Following the SSL iterations in baselines [15,16], the models were trained on the Pascal VOC dataset for 255k iterations (i.e., 150k for SSL & 75k for TMR) using the base detector of Faster-RCNN. For Faster-RCNN with cascade regression, the models were trained for 480k iterations (i.e., 300k for SSL & 150k for TMR) using a batch size of 32 (16 labeled images and 16 unlabelled images).

The implementations were performed on 16 synchro-

nized Nvidia Tesla V100 GPUs with 16GB RAM. The cascade regression utilizes three detection stages with IoU thresholds $\tau = \{0.5, 0.6, 0.7\}$ for generating high-quality pseudo-labels. The loss coefficients were set to $\lambda_t = 1$, $\lambda_s = 4$, $\lambda_u = 4$, while the RD loss coefficient was set to $\lambda_d = 0.5$ or $\lambda_d = 1$ using the base detector of Faster-RCNN without or with cascade regression, respectively. Moreover, the semantic representations were extracted from the last two levels of the backbones, i.e., $f_i = [\text{conv4_}x, \text{conv5_}x]$ (typically responsible for capturing complex patterns and representations that are relevant to final predictions). The optimizer, learning rates, data augmentations, and other hyperparameters were applied similarly to those in the baselines [15, 16].

4.2. State-of-the-art Comparison

In this section, we compare the proposed TMR-RD approach integrated into Unbiased-Teacher-v1 [15] (referred to as TMR-RD-v1) and Unbiased-Teacher-v2 [16] (referred to as TMR-RD-v2) with 13 recent SSOD methods, namely ISMT [30], Instant-Teaching [35], CAPL [11], SED [6], Humble-Teacher [24], Soft-Teacher [29], MA-GCP [9], Unbiased-Teacher-v1 [15] (baseline 1), Unbiased-Teacher-v2 [16] (baseline 2), Active-Teacher [18], Diverse-Learner [12], Pseudo-Polish [33], and De-biased Teacher [26]. Table 2 and Table 3 provide the results on the MS-COCO and Pascal VOC datasets, respectively. To ensure fair comparisons, we compare the two versions of the proposed TMR-RD using the Faster-RCNN against SOTA methods. Additionally, we examine improvements resulting from cascade regression compared to their baselines, aimed at generating more reliable pseudo-labels. We highlight the generalization capability of our approach through extensive empirical experiments, with and without cascade regression.

COCO-standard: First, we compare the proposed approach with existing SOTA methods partially trained on the COCO-standard set. As shown in Table 2, our TMR-RD-v2 considerably outperforms the Diverse-Learner [12], Pseudo-Polish [33], and De-biased Teacher [26] by average margins of 2.92, 2.69, and 2.97 mAP across all labeling ratios, respectively. In particular, our TMR-RD demonstrates superior performance over the Unbiased-Teacher-v1 [15] and Unbiased-Teacher-v2 [16] baselines by an average of 2.07 and 2.23 mAP, respectively. Furthermore, integrating the proposed approach with cascade regression surpasses the performance of both TMR-RD-v1 and TMR-RD-v2 by average margins of 2.64 and 2.88 mAP, respectively.

COCO-additional: We then evaluate the performance of the proposed approach fully trained on supervised data to determine the extent of further improvement by incorporating additional unlabeled data. According to Table 2, our TMR-RD-v2 provides an mAP improvement of 2.1, 2.1, and 2.2 compared to the Unbiased-Teacher-v2 [16],

Table 2. State-of-the-art comparison on the MS-COCO dataset under the partially-labeled data (COCO-standard) and fully-labeled data (COCO-additional) settings. The proposed approach, referred to as TMR-RD-v1 and TMR-RD-v2, is integrated into [15] and [16], respectively, with base detectors shown as Faster-RCNN with or without cascade regression.

Methods	1% labeled samples	5% labeled samples	10% labeled samples	Fully-labeled samples
Supervised	9.05	18.47	23.86	40.2
Supervised (with Cascade regression)	10.86	19.44	25.18	42.1
ISMT [30]	18.88	26.37	30.53	39.6
Instant-Teaching [35]	18.05	26.75	30.40	40.2
CAPL [11]	19.02	28.40	32.23	43.3
SED [6]	-	29.01	34.02	41.5
Humble-Teacher [24]	16.96	27.70	31.61	42.4
Soft-Teacher [29]	20.46	30.74	34.04	44.5
MA-GCP [9]	21.30	31.67	35.02	45.9
Unbiased-Teacher-v1 [15]	20.75	28.27	31.50	41.3
Unbiased-Teacher-v2 [16]	25.40	31.85	35.08	44.8
Active-Teacher [18]	22.20	30.07	32.58	-
Diverse-Learner [12]	23.72	31.92	34.61	44.8
Pseudo-Polish [33]	23.55	32.10	35.30	-
De-biased Teacher [26]	22.50	32.10	35.50	44.7
TMR-RD-v1 (baseline [15]: Faster-RCNN)	22.23	30.60	33.92	43.2
TMR-RD-v2 (baseline [16]: Faster-RCNN)	26.91	34.37	37.74	46.9
TMR-RD-v1 (baseline [15] + cascade regression)	24.39	33.41	36.87	46.6
TMR-RD-v2 (baseline [16] + cascade regression)	29.16	37.58	40.93	50.4

Table 3. State-of-the-art comparison on the Pascal VOC.

Method	AP ₅₀	AP _{50:95}
ISMT [30]	77.23	46.23
Unbiased-Teacher-v1 [15]	77.37	48.69
Instant-Teaching [35]	78.30	48.70
Humble-Teacher [24]	80.94	53.04
Soft-Teacher [29]	80.32	-
CAPL [11]	79.0	54.60
SED [6]	80.60	-
MA-GCP [9] [9]	81.72	-
Unbiased-Teacher-v2 [16]	81.29	56.87
Pseudo-Polish [33]	82.50	52.40
De-biased Teacher [26]	81.50	-
TMR-RD-v1 (baseline [15]: Faster-RCNN)	79.83	51.96
TMR-RD-v2 (baseline [16]: Faster-RCNN)	83.66	60.23
TMR-RD-v1 (baseline [15] + cascade regression)	82.68	55.08
TMR-RD-v2 (baseline [16] + cascade regression)	85.92	63.48

Diverse-Learner [12], and De-biased Teacher [26], respectively. Moreover, the proposed TMR-RD-v1 and TMR-RD-v2 outperform their corresponding supervised models by 3 and 6.7 mAP, respectively. Further, the utilization of cascade regression demonstrates superior performance, surpassing the mAP of TMR-RD-v1 and TMR-RD-v2 by 3.4 and 3.5, respectively. These comparisons highlight the effectiveness of the proposed approach with accessible large amounts of labeled and unlabeled data, in addition to partially labeled datasets where labeled data is limited.

Pascal VOC: At last, the proposed TMR-RD demonstrates its superiority over existing SSOD methods, as shown in Table 3. For example, TMR-RD-v2 outperforms De-biased Teacher [26], MA-GCP [9], and Pseudo-Polish [33] by margins of up to 2.16, 1.94, and 1.16 in the AP₅₀ metric, respectively. Also, it surpasses Unbiased-Teacher-v1 [15] and Unbiased-Teacher-v2 [16] by 3.36 and 3.27

mAP, respectively. Additionally, the integration of cascade regression yields superior results, surpassing TMR-RD-v2 and TMR-RD-v1, leading to improvements of 3.25 and 3.12 mAP, respectively.

4.3. Ablation Analysis

A systematic ablation analysis of the proposed approach, integrated into the Unbiased-Teacher-v1 [15] and Unbiased-Teacher-v2 [16], has been conducted using a 5% labeled COCO-standard and Pascal VOC datasets. The analysis includes 12 configurations (see Table 4 & Table 5), including: i) baseline [15] with the Faster-RCNN and classical EMA (A1), ii) A1 equipped with TMR stage (A11), iii) A11 equipped with RD strategy (A12), iv) baseline [16] with the Faster-RCNN and classical EMA (A2), v) A2 equipped with TMR stage (A21), vi) A21 equipped with RD strategy (A22), vii) baseline [15] equipped with cascade regression and classical EMA (A3), viii) A3 equipped with TMR stage (A31), ix) A31 equipped with RD strategy (A32), x) baseline [16] equipped with cascade regression and classical EMA (A4), xi) A4 equipped with TMR stage (A41), and xii) A41 equipped with RD strategy (A42). These analyses validate the effectiveness of our TMR stage and RD strategy.

Effect of TMR stage: We first ablate the impact of the proposed TMR stage integrated into both baselines with or without cascade regression. According to the results of A11, A21, A31, and A41 in Table 4 and Table 5, the baselines of [16] and [15] with the Faster-RCNN are notably outperformed by the proposed TMR stage, with average mAP improvements of 2.12 and 2.03, respectively. Moreover, incorporating our TMR stage considerably en-

Table 4. Ablation study of the proposed TMR-RD approach on the COCO-standard set. Classical EMA and cascade regression are denoted as cEMA and c-regress, respectively.

Abl.	A1	A2	cEMA	TMR	RD	c-regress.	mAP
A1 [15]	✓	✗	✓	✗	✗	✗	28.27
A11	✓	✗	✗	✓	✗	✗	29.93
A12	✓	✗	✗	✓	✓	✗	30.60
A2 [16]	✗	✓	✓	✗	✗	✗	31.85
A21	✗	✓	✗	✓	✗	✗	33.63
A22	✗	✓	✗	✓	✓	✗	34.37
A3	✓	✗	✓	✗	✗	✓	30.19
A31	✓	✗	✗	✓	✗	✓	32.66
A32	✓	✗	✗	✓	✓	✓	33.41
A4	✗	✓	✓	✗	✗	✓	34.05
A41	✗	✓	✗	✓	✗	✓	36.74
A42	✗	✓	✗	✓	✓	✓	37.58

Table 5. Ablation study of the proposed TMR-RD approach on the Pascal VOC dataset.

Abl.	A1	A2	cEMA	TMR	RD	AP ₅₀	AP _{50:95}
A1 [15]	✓	✗	✓	✗	✗	77.37	48.69
A11	✓	✗	✗	✓	✗	79.21	51.09
A12	✓	✗	✗	✓	✓	79.83	51.96
A2 [16]	✗	✓	✓	✗	✗	81.29	56.87
A21	✗	✓	✗	✓	✗	82.94	59.32
A22	✗	✓	✗	✓	✓	83.66	60.23

hances the performance of these baselines by 2.69 and 2.47 mAP, respectively, when employing the base detector with cascade regression. These results further demonstrate the versatility of our method and its ability to improve overall performance and generalization of existing methods. We also validated the efficacy of the TMR stage by visually analyzing the mAP improvement of baseline [16] in Fig. 3(a). While our model (shown in green) was trained with fewer SSL steps (For comparability purposes considering our cyclic SSL and TMR stages), our method effectively refines the models to transfer reliable knowledge while preventing them from misleading information and converging too early (unlike the baseline [16] shown in red).

Effect of RD Strategy: The results of experiments A12, A22, A32, and A42 provide compelling support for the usefulness of the proposed RD strategy. These results demonstrate that the performance of [16] and [15] baselines utilizing the Faster-RCNN detector are surpassed by employing the proposed RD strategy with average mAP improvements of 0.82 and 0.77, respectively. Notably, the integration of this strategy led to satisfactory effects in the performance of the baseline [16] as shown in Fig. 3(a). It validates the complementary effect of this strategy and demonstrates how it promotes the mutual learning of Teacher-Student models by reinforcing model distinctiveness.

Effect of SSL and TMR iterations: We conducted additional ablations, as illustrated in Fig. 3(b), to analyze the impacts of various SSL and TMR iterations on model performance using baseline [15]. The best result (blue curve)

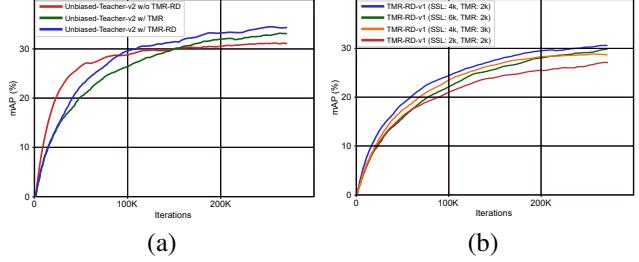


Figure 3. Ablation study of mAP improvement in the case of COCO-standard 5% labeled data. (a): Integration of the TMR stage and RD strategy into Unbiased-Teacher-v2 [16] for effectiveness evaluation. (b): Exploration of the impact of varying SSL and TMR stages using Unbiased-Teacher-v1 [15]. Best viewed in color and zoom-in.

was achieved when employing cyclic SSL and TMR stages with $N = 4k$ and $N' = 2k$ iterations. The red curve highlights the necessity for models to undergo sufficient evolution during the SSL stage before each subsequent TMR refinement while increasing the SSL stage (green curve) can result in a delay in model refinement and slight performance degradation, as the teacher model is expected to generate more accurate predictions than the student model [24]. Also, increasing the number of TMR iterations with partially labeled data can impede the subsequent joint evolution of models using unlabeled data (orange curve).

5. Conclusion

In this paper, we presented a novel training-based model refinement stage and a representation disagreement strategy for existing SSOD frameworks. To address the limitations of the classical EMA strategy, our proposed model refinement stage learns lightweight scaling operation parameters to dynamically refine the weights of Teacher-Student models, ensuring that overfitting and forgetting learned patterns from unlabeled data are avoided. Moreover, we introduce a simple yet effective representation disagreement strategy to alleviate the consensus of Teacher-Student models that arise with the progression of model training. This strategy promotes model distinctiveness and prevents the models from converging too early, encouraging the student model to explore more underlying patterns in unlabeled data. Our approach can be integrated into existing SSOD methods to improve performance and generalization, transfer reliable knowledge from the student to the teacher, and prune noisy student weights. Extensive experiments demonstrate notable performance improvements with different baseline detectors, as well as the potential to integrate the proposed approach into existing SSOD methods.

Acknowledgments. We thank the fRI Research - MPB Ecology Program for funding and the Digital Research Alliance of Canada for providing the computational resources for our experiments.

References

[1] Yoshua Bengio. *Evolving culture versus local minima*, pages 109–138. 2014. 5

[2] Avrim Blum and Tom Mitchell. Combining labeled and unlabeled data with co-training. In *Proc. Annual Conf. Computational Learning Theory*, page 92–100, 1998. 5

[3] Binbin Chen, Weijie Chen, Shicai Yang, Yunyi Xuan, Jie Song, Di Xie, Shiliang Pu, Mingli Song, and Yueteng Zhuang. Label matching semi-supervised object detection. In *Proc. CVPR*, pages 14361–14370, 2022. 2, 3

[4] Changrui Chen, Kurt Debbattista, and Jungong Han. Semi-supervised object detection via virtual category learning. In *Proc. ECCV*, 2022. 2, 3

[5] Mark Everingham, Luc Van Gool, Christopher K. I. Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, 88(2):303–338, 2010. 5

[6] Qiushan Guo, Yao Mu, Jianyu Chen, Tianqi Wang, Yizhou Yu, and Ping Luo. Scale-equivalent distillation for semi-supervised object detection. In *Proc. CVPR*, pages 14502–14511, 2022. 2, 3, 6, 7

[7] Jisoo Jeong, Seungeui Lee, Jeesoo Kim, and Nojun Kwak. Consistency-based semi-supervised learning for object detection. In *Proc. NIPS*, 2019. 2, 3

[8] JongMok Kim, JooYoung Jang, Seunghyeon Seo, Jisoo Jeong, Jongkeun Na, and Nojun Kwak. MUM: Mix image tiles and unmix feature tiles for semi-supervised object detection. In *Proc. CVPR*, pages 14492–14501, 2022. 2, 3

[9] Aoxue Li, Peng Yuan, and Zhenguo Li. Semi-supervised object detection via multi-instance alignment with global class prototypes. In *Proc. CVPR*, pages 9799–9808, 2022. 2, 3, 6, 7

[10] 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 *Proc. ECCV*, pages 457–472, 2022. 2, 3

[11] Hengduo Li, Zuxuan Wu, Abhinav Shrivastava, and Larry S. Davis. Rethinking pseudo labels for semi-supervised object detection. In *Proc. AAAI*, pages 1314–1322, 2022. 2, 3, 6, 7

[12] Linfeng Li, Minyue Jiang, Yue Yu, Wei Zhang, Xiangru Lin, Yingying Li, Xiao Tan, Jingdong Wang, and Errui Ding. Diverse learner: Exploring diverse supervision for semi-supervised object detection. In *Proc. ECCV*, pages 640–655, 2022. 2, 3, 5, 6, 7

[13] Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proc. CVPR*, 2017. 6

[14] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft COCO: Common objects in context. In *Proc. ECCV*, pages 740–755, 2014. 5

[15] 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 *Proc. ICLR*, 2021. 2, 3, 4, 5, 6, 7, 8

[16] Yen-Cheng Liu, Chih-Yao Ma, and Zsolt Kira. Unbiased teacher v2: Semi-supervised object detection for anchor-free and anchor-based detectors. In *Proc. CVPR*, pages 9809–9818, 2022. 2, 3, 4, 5, 6, 7, 8

[17] Eran Malach and Shai Shalev-Shwartz. Decoupling “when to update” from “how to update”. In *Proc. NIPS*, page 961–971, 2017. 5

[18] Peng Mi, Jianghang Lin, Yiyi Zhou, Yunhang Shen, Gen Luo, Xiaoshuai Sun, Liujuan Cao, Rongrong Fu, Qiang Xu, and Rongrong Ji. Active teacher for semi-supervised object detection. In *Proc. CVPR*, pages 14462–14471, 2022. 2, 3, 6, 7

[19] Guo-Jun Qi and Jiebo Luo. Small data challenges in big data era: A survey of recent progress on unsupervised and semi-supervised methods. *IEEE TPAMI*, 44(4):2168–2187, 2022. 1

[20] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. In *Proc. NIPS*, 2015. 2, 6

[21] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. ImageNet large scale visual recognition challenge. *IJCV*, 115(3):211–252, 2015. 6

[22] Kihyuk Sohn, Zizhao Zhang, Chun-Liang Li, Han Zhang, Chen-Yu Lee, and Tomas Pfister. A simple semi-supervised learning framework for object detection. In *arXiv:2005.04757*, 2020. 2, 3

[23] Qianru Sun, Yaoyao Liu, Tat-Seng Chua, and Bernt Schiele. Meta-transfer learning for few-shot learning. In *Proc. CVPR*, 2019. 2, 4

[24] Yihe Tang, Weifeng Chen, Yijun Luo, and Yuting Zhang. Humble teachers teach better students for semi-supervised object detection. In *Proc. CVPR*, pages 3132–3141, 2021. 2, 3, 6, 7, 8

[25] Antti Tarvainen and Harri Valpola. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. In *Proc. NIPS*, 2017. 3

[26] Kuo Wang, Jingyu Zhuang, Guanbin Li, Chaowei Fang, Lechao Cheng, Liang Lin, and Fan Zhou. De-biased teacher: Rethinking iou matching for semi-supervised object detection. *Proc. AAAI*, 2023. 2, 3, 6, 7

[27] Hongxin Wei, Lei Feng, Xiangyu Chen, and Bo An. Combating noisy labels by agreement: A joint training method with co-regularization. In *Proc. CVPR*, 2020. 5

[28] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. Detectron2. <https://github.com/facebookresearch/detectron2>, 2019. 6

[29] 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 *Proc. ICCV*, pages 3060–3069, 2021. 2, 3, 6, 7

[30] Qize Yang, Xihan Wei, Biao Wang, Xian-Sheng Hua, and Lei Zhang. Interactive self-training with mean teachers for semi-supervised object detection. In *Proc. CVPR*, pages 5937–5946, 2021. 2, 3, 6, 7

[31] Xingrui Yu, Bo Han, Jiangchao Yao, Gang Niu, Ivor Tsang, and Masashi Sugiyama. How does disagreement help gen-

eralization against label corruption? In *Proc. ICML*, pages 7164–7173, 2019. 5

[32] Fangyuan Zhang, Tianxiang Pan, and Bin Wang. Semi-supervised object detection with adaptive class-rebalancing self-training. In *Proc. AAAI*, pages 3252–3261, 2022. 2, 3

[33] Lei Zhang, Yuxuan Sun, and Wei Wei. Mind the gap: Polishing pseudo labels for accurate semi-supervised object detection. *Proc. AAAI*, 2023. 2, 3, 6, 7

[34] Ying Zhang, Tao Xiang, Timothy M. Hospedales, and Huchuan Lu. Deep mutual learning. In *Proc. CVPR*, 2018. 5

[35] Qiang Zhou, Chaohui Yu, Zhibin Wang, Qi Qian, and Hao Li. Instant-teaching: An end-to-end semi-supervised object detection framework. In *Proc. CVPR*, pages 4081–4090, 2021. 2, 3, 6, 7