Sparse Semi-DETR: Sparse Learnable Queries for Semi-Supervised Object Detection

Tahira Shehzadi\textsuperscript{1,2,3}, Khurram Azeem Hashmi\textsuperscript{1,2,3}, Didier Stricker\textsuperscript{1,2,3}, Muhammad Zeshan Afzal\textsuperscript{1,2,3} \\
\textsuperscript{1}DFKI, \textsuperscript{2}RPTU Kaiserslautern-Landau, \textsuperscript{3}MindGarage-RPTU

Abstract

In this paper, we address the limitations of the DETR-based semi-supervised object detection (SSOD) framework, particularly focusing on the challenges posed by the quality of object queries. In DETR-based SSOD, the one-to-one assignment strategy provides inaccurate pseudo-labels, while the one-to-many assignments strategy leads to overlapping predictions. These issues compromise training efficiency and degrade model performance, especially in detecting small or occluded objects. We introduce Sparse Semi-DETR, a novel transformer-based, end-to-end semi-supervised object detection solution to overcome these challenges. Sparse Semi-DETR incorporates a Query Refinement Module to enhance the quality of object queries, significantly improving detection capabilities for small and partially obscured objects. Additionally, we integrate a Reliable Pseudo-Label Filtering Module that selectively filters high-quality pseudo-labels, thereby enhancing detection accuracy and consistency. On the MS-COCO and Pascal VOC object detection benchmarks, Sparse Semi-DETR achieves a significant improvement over current state-of-the-art methods that highlight Sparse Semi-DETR’s effectiveness in semi-supervised object detection, particularly in challenging scenarios involving small or partially obscured objects.

1. Introduction

Semi-Supervised Object Detection (SSOD) aims to improve the effectiveness of fully supervised object detection through the integration of abundant unlabeled data [3, 12, 15, 19, 28, 39, 49–52, 58]. It has applications in diverse fields, ranging from autonomous vehicles [14, 20] to healthcare [31, 44], where obtaining extensive labeled datasets is often impractical or cost-prohibitive [1].

Several SSOD methods [3, 12, 15, 19, 28, 39, 49–52, 58] have been proposed. Two prevalent approaches in this domain are pseudo-labeling [27, 30, 39, 45, 50–52, 58] and consistency-based regularization [3, 12, 15, 19, 28, 49]. STAC [39] introduced a simple multi-stage SSOD training method with pseudo-labeling and consistency training, later simplified by a Teacher-Student framework for generating pseudo-labels [27]. Based on this framework, considerable research efforts have been directed towards enhancing the quality of pseudo-labels [51, 58]. These traditional SSOD methods are built upon conventional detectors like one-stage [33, 42] and two-stage [9, 35], which involve various manually designed components such as an-
chor boxes and non-maximum suppression (NMS). Employing object detection methods in SSOD poses several potential challenges that must be carefully dealt with to obtain reasonable performance. These factors include overfitting of the labeled data [37], pseudo-label noise [10], bias induced through label imbalance [17, 32], and poor detection performance on small objects [56]. Recently, DETR-based [2, 16, 18, 25, 38, 55, 59] SSOD methods [46, 56] remove the need for traditional components like NMS.

Even though DETR-based SSOD [46, 56] has progressed remarkably, state-of-the-art methods possess some limitations. (1) DETR-based SSOD methods perform poorly in the detection of small objects, as shown in Figure 4. This is because these methods don’t use multi-scale features [36] like Feature Pyramid Networks (FPN) [22], which play an important role in identifying smaller objects as in CNN-based SSOD methods [3, 12, 15, 19, 28, 49, 51, 52, 58]. Although recent advancements in DETR-based object detection [2, 16, 18, 25, 38, 55, 59] have improved the detection of small objects, their SSOD adaptation is still unable to cater this challenge effectively [56]. (2) SSOD approaches [51, 52, 56, 58] rely on handcrafted post-processing methods such as NMS [35]. This problem specifically appears in DETR-based SSOD when we use a large number of object queries and the one-to-many assignment strategy [56]. In DETR-based SSOD methods, this problem is partially solved using the one-to-one or hybrid (combination of one-to-one and one-to-many) assignment strategy. However, the hybrid assignment strategy is preferred because the one-to-one assignment strategy produces inaccurate pseudo-labels [46], thus resulting in inefficient learning. Although the number of duplicate bounding boxes is less in the hybrid strategy [56], the amount is high enough to impact object detection performance adversely, as depicted in Figure 5. (3) The pseudo-label generation produces both high and low-quality labels. The DETR-based SSOD methods lack an effective refinement strategy for one-to-many assignments, which is crucial for filtering out low-quality proposals.

To address the above mentioned issues, we propose enhancing the state-of-the-art DETR-based SSOD approach, namely ‘Sparse Semi-DETR’, presented in Figure 1 (b). Our approach involves expanding its architecture by integrating a couple of novel modules designed to mitigate the identified shortcomings. The key module among these is the **Query-Refinement module**, as depicted in Figure 2 and explained in Figure 3. This module significantly improves the quality of the queries and reduces their numbers. The proposed module uses the low-level features from the backbone and high-level features extracted directly from weakly augmented images using ROI alignment [13]. Fusing these features results in overcoming the first shortcoming, i.e., detecting small and obscured objects, as shown in Figure 4. The attention mechanism drives the aggregation of the features, resulting in refined, high-quality features to carry forward. To ensure the quality of the query features, the attention mechanism is accompanied by a query-matching strategy for filtering irrelevant queries. Thus, the Query Refinement Module not only improves the quality of the queries but also reduces their numbers, giving rise to efficient processing. This module results in significantly fewer overlapping proposals, improving the performance overall, thereby solving the second limitation. Besides, we introduce a **Reliable Pseudo-Label Filtering Module**, as illustrated in Figure 2, inspired by Hybrid-DETR [16] to address the third limitation. Employing this module significantly reduces the low-quality pseudo-labels. Therefore, it further reduces the amount of duplicate predictions that may still occur after the second stage of the hybrid assignment strategy. Our approach provides better results than previous SSOD methods, as shown in Figure 1 (c).

The key contributions of this work can be outlined as follows:

1. We present Sparse Semi-DETR, a novel approach in semi-supervised object detection, introducing two novel contributions. To our knowledge, we are the first to examine and propose query refinement and low-quality proposal filtering for the one-to-many query assignment strategy.
2. We introduce a novel query refinement module designed to improve object query features, particularly in complex detection scenarios such as identifying small or partially obscured objects. This enhancement not only boosts performance but also aids in learning semantic feature invariance among object queries.
3. We introduce a Reliable Pseudo-Label Filtering Module specifically designed to reduce the effect of noisy pseudo-labels. This module is designed to efficiently identify and extract reliable pseudo boxes from unlabeled data using augmented ground truths, enhancing the consistency of the learning process.
4. Sparse Semi-DETR outperforms current state-of-the-art methods on MS-COCO and Pascal VOC benchmarks. With only 10% labeled data from MS-COCO using ResNet-50 backbone, it achieves a 44.3 mAP, exceeding prior baselines by 0.8 mAP. Additionally, when trained on the complete COCO set with extra unlabeled data, it further improves, rising from 49.2 to 51.3 mAP.

2. Related Work

2.1. Object Detection

Object detection identifies and locates objects in images or videos. Deep learning-based object detection approaches are typically categorized into two primary groups: two-stage detectors [9, 35] and one-stage detectors [23, 26, 34, 42]. These methods depend on numerous heuristics, such as generating anchors and NMS. Recently, DEtection TRans-
former (DETR) [2] considers object detection as a set prediction problem, using transformer [43] to adeptly transform sparse object candidates [40] into precise target objects. Our Sparse Semi-DETR detects small or partially obscured objects in the DETR-based SSOD setting. Notably, our framework is compatible with various DETR-based detectors [4, 5, 7, 8, 24, 29, 47, 48, 59], offering flexibility in integration.

2.2. Semi-Supervised Object Detection

Most research in SSOD employs detectors categorized into three types: one-stage, two-stage, and DETR-based systems. One-stage STAC [39], an early SSOD, introduced a simple training strategy combining pseudo-labeling and consistency training, later streamlined by a student-teacher framework for easier pseudo-label generation [27]. DSL [3] introduced novel techniques including Adaptive Filtering, Aggregated Teacher, and uncertainty-consistency-regularization for improved generalization. Dense Teacher [57] introduced Dense Pseudo-Labels (DPL) for richer information and a region selection method to reduce noise.

Two-stage. Humble Teacher [41] uses soft labels and a teacher ensemble to boost pseudo-label reliability, matching other results. Instant-Teaching [58] creates pseudo annotations from weak augmentations, treating them as ground truth under strong augmentations with Mixup [54]. Unbiased Teacher [27] tackles class imbalance in pseudo-labeling with focal loss, focusing on underrepresented classes. Soft Teacher [51] minimizes incorrect foreground proposal classification by applying teacher-provided confidence scores to reduce classification loss. PseCo [19] enhances detector performance by combining pseudo-labeling with label and feature consistency methods, also using focal loss to address class imbalance.

DETR-based. Omni-DETR [46] is designed for omni-supervised detection and adapts to SSOD with a basic pseudo-label filtering method. It employs the one-to-one assignment strategy proposed in DETR [2], and encounters challenges when dealing with inaccurate pseudo-bounding boxes produced by the teacher network. These inaccuracies result in reduced performance, highlighting its limitations. Semi-DETR [56] adopts a stage-wise strategy, employing a one-to-many matching strategy in the first stage and switching to a one-to-one matching strategy in the second stage. This approach provides NMS-free end-to-end detection benefits but reduces performance compared to a one-to-many assignment strategy. Moreover, omni-DETR and Semi-DETR struggle to detect small or occluded objects. Our work introduces an advanced query refinement module that significantly refines object queries, enhancing training efficiency and performance and leading to the detection of small or densely packed objects in the DETR-based SSOD framework.

3. Preliminary

In DETR-based SSOD, one-to-one assignment strategy, denoted by \( \hat{\sigma}_{\text{one2one}} \), is achieved by applying Hungarian algorithm between the predictions made by the student model and the pseudo-labels provided by the teacher model as follows:

\[
\hat{\sigma}_{\text{one2one}} = \arg \min_{\sigma \in \xi_N} \sum_{j}^{N} \mathcal{L}_{\text{match}} \left( \hat{y}^s_j, \hat{y}^u_{\sigma(j)} \right)
\]

where \( \mathcal{L}_{\text{match}} \left( \hat{y}^s_j, \hat{y}^u_{\sigma(j)} \right) \) is the matching cost between the pseudo-labels \( \hat{y}^s_j \) generated by the teacher network and the predictions of the student network with index \( \sigma(j) \) and \( \xi_N \) is the permutation of \( N \) elements. Semi-DETR [56] addresses the issue of imprecise initial pseudo-labels by shifting from a one-to-one to a one-to-many assignment strategy, increasing the number of positive object queries to improve detection accuracy:

\[
\hat{\sigma}_{\text{one2many}} = \left\{ \arg \min_{\sigma \in C_N^M} \sum_{k}^{M} \mathcal{L}_{\text{match}} \left( \hat{y}^s_j, \hat{y}^u_{\sigma(k)} \right) \right\}_{j}^{\left| \hat{y}^l \right|}
\]

where \( C_N^M \) represents the combination of \( M \) and \( N \), denoting that a subset of \( M \) proposals is associated with each pseudo box \( \hat{y}^s \). Semi-DETR initially adopts a one-to-many assignment to improve label quality, then shifts to one-to-one assignment for an NMS-free model. This approach adopts a one-to-many assignment strategy aimed at boosting performance, but it’s less effective with small or occluded objects.

4. Sparse Semi-DETR

In semi-supervised learning, a collection of labeled data denoted as \( D_l \), where \( D_l = \{x^l_i, y^l_i\}_{i=1}^{N_l} \) is given, along with a set of unlabeled data represented as \( D_u \), where \( D_u = \{x^u_i\}_{i=1}^{N_u} \). Here, \( N_l \) and \( N_u \) correspond to the number of labeled and unlabeled data. The annotations \( y^u_i \) for the label data \( x^l \) contain object labels and bounding box information. The pipeline of the Sparse Semi-DETR framework is depicted in Figure 2. It introduces a Query Refinement Module for processing query features to enhance semantic representation for complex detection scenarios, such as identifying small or partially obscured objects. Additionally, we integrate a Reliable Pseudo-Label Filtering Module that selectively filters high-quality pseudo-labels, thereby enhancing detection accuracy. For comparison purposes, we employ DINO [55] with a ResNet-50 backbone. This section gives a detailed overview of the modules of Sparse Semi-DETR. We explain briefly our semi-supervised approach in Appendix A1.1.
with attentional operations is computationally demanding. Inspired by recent advancements in vision-based networks [11, 53], we introduce an innovative approach to work as works [11, 53], we introduce an innovative approach to address this, we firstly convert the query label features $F_{t1}$ from weakly augmented image $I$, also in the same dimension. Subsequently, feature extraction from the image backbones occurs. This results in the generation of features $F_{t2}$ for the student and $F_{t2}$ for the teacher network as $I^{k \times W_{2} \times 256}$. These features, encompassing both label and bounding box details, vary with the batch size, as indicated by $b$. The feature sets $W_{1}$ and $W_{2}$ differ in size, with $W_{2}$ being substantially larger than $W_{1}$. We provide a brief overview of each component of Query Refinement Module as illustrated in Figure 3.

**Query Refinement Module.** In our approach, we handle multi-scale features $F_{t1}$ and $F_{t2}$ with a focus on effective aggregation. The finer details are encapsulated within the features $F_{t1}$, while the features $F_{t2}$ encapsulate more abstract elements such as shapes and patterns. Simple aggregation of these features has been shown to degrade performance, as indicated in Table 5e). To solve this issue, we implement dual strategies to extract local and global information from high and low-resolution features. High-resolution features are crucial for detecting small objects. However, processing them with attentional operations is computationally demanding. To address this, we firstly convert the query label features $F_{t2} \in (\mathbb{R}^{b \times W_{2} \times 256})$ into $F'_{t2} \in (\mathbb{R}^{b \times W_{2} \times 16})$ by decreasing the channel dimension, and retaining the original resolution $b \times W_{2}$. Then, we apply attentional mechanism on $F'_{t2}$ to calculate the attentional weights $W_{k+q}$ in attention block as follows:

$$W_{k+q} = F'_{k} \cdot F'_{q}, \quad (3)$$

$$\bar{W}_{k+q} = \frac{\exp(W_{k+q})}{\sum_{l=1}^{L} \exp(W_{k+q})}, \quad (4)$$

where $W_{k+q}$ is the attentional weights of $F'_{k}$ and $F'_{q}$, and $\bar{W}_{k+q}$ is the normalized form of $W_{k+q}$. Using normalized attention weights, we compute the enhanced queries representation $Q$ as follows:

$$Q = \bar{W}_{k+q} \cdot F'_{q}, \quad (5)$$

now we find the similarity between the attentional $F'_{t2}$ features and $F_{t1}$ features to obtain $F'_{c_{s}} \in \mathbb{R}^{b \times W_{1} \times 16}$ from $Q \in \mathbb{R}^{b \times W_{2} \times 16}$ as follows:

$$F'_{c_{s}} = \frac{\sum_{l=1}^{n} P_{l}Q_{l}}{\sqrt{\sum_{l=1}^{n} P_{l}^{2}} \sqrt{\sum_{l=1}^{n} Q_{l}^{2}}} \quad (6)$$

where $P$ and $Q$ are $F_{t1}$ and attentional $F_{t2}$ features, respectively. Then, we concatenate $F'_{c_{s}}$ with $P$ to obtain refined query features. Interestingly, we observe a performance drop when our feature refinement strategy is applied to strongly augmented image features for the teacher network, as detailed in Table 5b. However, we achieve optimal results by concatenating strongly augmented image features and applying our refinement strategy to weakly augmented image features. Consequently, we proceed by concatenating the features $F_{s1}$ with $F_{c_{s}}$, thereby obtaining the query features $F'_{c_{s}}$. Furthermore, a Reliable Pseudo-label Filtering strategy is employed to filter low-quality pseudo-labels progressively during training.

![Figure 2. An overview of the Sparse Semi-DETR framework.](image-url)
The one-to-many training strategy, while effective, causes a loss of ground truths in the data using augmented ground truths. We employ m groups of refined query learning. This module is designed to efficiently identify and extract reliable pseudo boxes from unlabeled data. The filtering of pseudo boxes for small objects. For the best view, zoom in.

4.2. Reliable Pseudo-Label Filtering Module

The one-to-many training strategy, while effective, causes duplication prediction in the first stage. We introduce a pseudo-label filtering module to address this and improve the filtering of pseudo boxes rich in semantic content for refined query learning. This module is designed to efficiently identify and extract reliable pseudo boxes from unlabeled data using augmented ground truths. We employ m groups of ground truths \( \mathcal{g} = \{ \hat{g}_1, \hat{g}_2, \ldots, \hat{g}_m \} \) for one-to-many assignment strategy and select the top-k predictions as follows:

\[
\delta_{\text{one2many}} = \left\{ \arg \min_{\sigma_j \in C_N^M} \sum_k L_{\text{match}} \left( \hat{y}_j, \hat{g}_{\sigma_j(k)} \right) \right\}_{\hat{y}_j} \tag{9}
\]

where \( C_N^M \) represents the combination of M and N, denoting that a subset of M proposals is associated with each pseudo box \( \hat{g}_j \). Here, m is set to 6. Furthermore, we use the remaining predictions to filter out duplicates in the top-k predictions in the one-to-one matching branch. Through this improved selection scheme, we achieve a performance improvement of 0.4 mAP when m is set to 6, as shown in Table 5a. However, we observe no significant benefits when increasing m greater than 6, as detailed in Table 6a.

5. Experiments

5.1. Datasets

We evaluate our approach on the MS-COCO [21] and Pascal VOC [6] datasets, benchmarking it against current SOTA SSOD methods. Following [51, 56], Sparse Semi-DTER is evaluated in three scenarios: COCO-Partial. We use 1%, 5%, 10% of train2017 as label data and rest as unlabeled data. COCO-Full. We take train2017 as label data and unlabel2017 as unlabeled data. VOC. We take VOC2007 as label data and VOC2012 as unlabeled data. Evaluation metrics include \( AP_{50:95}, AP_{50}, \) and \( AP_{75} \). We provide complete implementation details for each experiment in Appendix A1.2.

5.2. Implementation Details

We set the quantity of DINO original object queries to 900. For the setting hyperparameters, following [56]: (1) In the COCO-Partial setting, we set the training iterations to 120k with a labeled to unlabeled data ratio of 1:4. The first 60k iterations adopt a one-to-many assignment strategy. (2) In the COCO-Full setting, Sparse Semi-DTER is trained for 240k iterations with labeled to unlabeled data ratio of 1:1. The first 120k iterations adopt a one-to-many assignment strategy. (3) In the VOC setting, we train the network for 60k iterations with a labeled to unlabeled data ratio of 1:4. The first 40k iterations adopt a one-to-many assignment strategy. For all experiments, the filtering threshold \( \sigma \) value is 0.4. We set the value of m to 6 and the value of k to 4.

6. Results and Comparisons

We evaluate Sparse Semi-DTER and compare it against current SOTA SSOD methods. Our results demonstrate the superior performance of Sparse Semi-DTER in these aspects: (1) its effectiveness compared to both one-stage and two-stage detectors, (2) its comparison with traditional DETR-based detectors, and (3) its exceptional proficiency in accurately
detecting small and partially occluded objects. We provide more results details in Appendix A.1.3.

**COCO-Partial benchmark.** Sparse Semi-DETR outperforms the current SSOD methods in COCO-Partial across all experiment settings, as demonstrated in Table 1. (1) We compare our method to both one-stage and two-stage SSOD. Sparse Semi-DETR surpasses Dense Teacher by 8.52, 7.79, 7.17 mAP on 1%, 5%, and 10% label data. It also outperforms PseCo by 8.47, 8.30, 8.24 mAP on 1%, 5%, and 10% label data. Sparse Semi-DETR’s superior performance as a semi-supervised object detector is achieved without needing hand-crafted components commonly used in two-stage and one-stage detectors. (2) When compared to DETR-based detectors, Sparse Semi-DETR outperforms omni-DETR by 3.30, 3.10, and 3.00 mAP and beats Semi-DETR by 0.40, 0.70, 0.80 mAP on 1%, 5%, and 10% label data. (3) Sparse Semi-DETR’s exceptional proficiency in precisely detecting small and partially obscured objects is a standout feature. In Figure 4, we visually compare Sparse Semi-DETR with the two preceding approaches using the COCO 10% labeled dataset. These results demonstrate the impressive capabilities of Sparse Semi-DETR, particularly in its ability to identify small objects and objects concealed by obstacles, as highlighted in the third-row images by the white arrows. These results confirm that each component of Sparse Semi-DETR enhances our model’s performance.

**6.1. Ablation Studies**

This section ablates the key design choices of Sparse Semi-DETR. The experiments detailed in this section are executed on the MS COCO dataset with 10% label data, employing DINO as the primary detector.

**Effect of Individual Component.** We conduct three experiments to assess the efficacy of each module of Sparse Semi-DETR. The first experiment involves the Query Refinement Module. In Table 5b, we explore various QR combinations to determine the most effective combination for our model. The results show that the Query Refinement Module significantly improves the performance of Sparse Semi-DETR.

**Effect of Query Refinement Module.** We examine the impact of the Query Refinement (QR) Module. In Table 5b, we explore various QR combinations to determine the most effective combination for our model.

**Pascal VOC benchmark.** Sparse Semi-DETR exhibits a remarkable performance boost on the Pascal VOC benchmark, as shown in Table 3. It surpasses the supervised baseline by 5.1 AP and by 5.9 AP on AP$_{50:95}$. Furthermore, it outperforms all previously single-stage, two-stage, and DETR-based SSOD methods by a significant margin.

**Figure 4.** Visual comparison of Sparse Semi-DETR with the two previous approaches on the COCO 10% label dataset. These results highlight Sparse Semi-DETR’s capabilities, particularly in identifying small objects and those obscured by obstacles (as indicated by white arrows) in the third-row images. For optimal clarity and detail, please zoom in.
Table 1. Comparing Sparse Semi-DETR with other approaches on COCO-Partial setting. The results are the average across all five folds. Under the COCO-partial setting, FCOS serves as the baseline for one-stage detectors, Faster RCNN for two-stage detectors, and DINO for transformer-based end-to-end detectors.

Table 2. Experimental results on COCO-partial settings for small, medium, and large objects. The results shown are the average across all five folds. We reproduce Semi-DETR results using their source code.

Table 3. Experimental results on Pascal VOC protocol. Here, FCOS, Faster RCNN, and DINO are the supervised baselines.

Table 4. Comparing Sparse Semi-DETR with other approaches on COCO-Full. Note that I denotes 30k training iterations, while an N× signifies N times 30k iterations.
Table 5. Ablations for the proposed Sparse Semi-DETR on COCO 10% Label dataset. (a) We analyze the effectiveness of each module of Sparse Semi-DETR. (b) We experiment with different QR combinations to identify the optimal design, applying QR selectively on the student, the teacher, or both. (c) We analyze the effect of using MLP layers for $F_1$ and $F_s$. Empirical observation reveals that we do not require MLP as in [56]. (d) We observe the effect of the attentional block in the Query Refinement module by replacing it with simple concatenation or cosine similarity. (e) We vary the type of object queries fed to the decoder with the decoder’s original queries. The best results are highlighted.

Table 6. Ablations for the Reliable Pseudo-Label Filtering Module on COCO 10% Label dataset. We study the impact of various parameters as augmented ground truth (m), top pseudo-labels selection (k) and filtering threshold ($\sigma$).

Table 7. Evaluating One-to-Many Assignment Strategy.

7. Conclusion

In conclusion, we successfully address the inherent limitations of DETR-based semi-supervised object detection frameworks by introducing Sparse Semi-DETR. This novel solution effectively tackles overlapping predictions and the detection of small objects. Sparse Semi-DETR incorporates a Query Refinement Module to enhance object query quality, mainly benefitting the detection of small and partially obscured objects. Besides, it also introduces a Reliable Pseudo-Label Filtering Module to filter out low-quality pseudo-labels selectively, thereby enhancing overall detection accuracy and consistency with the remaining high-quality labels. Our method outperforms existing SSOD approaches, with extensive experiments demonstrating its effectiveness.

Ethical considerations. We study semi-supervised models, and agree that standard ethical considerations for visual recognition are applicable to our work.

Acknowledgements

This work was in parts supported by the EU Horizon Europe Framework under grant agreements 101135724 (LUMINOUS) and 101092312 (AIRISE) and by the Federal Ministry of Education and Research of Germany (BMBF) under grant 01QE2227C (HERON).
References


[38] Tahira Shehzadi, Khurram Azeem Hashmi, Didier Stricker, and Muhammad Zeshan Afzal. Object detection with transformers: A review, 2023. 2


