Research plan

1、 Summary

The performance of target detector is very complex.One of the key challenges of target detection in depth network is that scale change is the difficulty of small target detection.In the coco data set, we found that: there are many small targets (41.43%), but the distribution is very uneven, nearly 50% of the images have no small targets.There is also the problem that even if the small goal is included by anchor, the small goal may not be matched, so the learning efficiency of the small goal is too low.An anchor fixed: feedback driven target detection anchor arbiter is proposed. Firstly, this arbiter model guides whether to correct the anchor by making full use of the scale loss of small target and anchor as the feedback information. By dynamically adjusting the size of the inappropriate anchor, the anchor makes the small target more compact and basically solves the problem caused by the inappropriate anchorBecause the IOU of GT box and anchor is too small, the problem of proposal loss provides more excellent anchors. At the same time, the group IOU balance strategy enables the classifier to obtain balanced training samples of different scales and improves the detection accuracy of small objects.

2、 Field development

1.1 Unbalanced summary of target detection

category	sketch
Imbalance of positive and negative samples	The foreground and background are unbalanced, and the number of different types of input bounding boxes in the foreground is unbalanced.

Scale imbalance	The scale of input image and bounding box is unbalanced, and the contribution of different feature layers to the final result is unbalanced.
Spatial imbalance	The contribution of different samples to regression loss is unbalanced, the distribution of positive sample IOU is unbalanced, and the position of target in image is unbalanced.
Imbalance of objective function	The contribution of different tasks (such as regression and classification) to global loss is unbalanced.

Table 1 Summary of target detection imbalance

1.2 anchor based Improvement

Based on the data level:

Stitcher: by setting a threshold value in each batch, large and medium-sized objects are transformed into small and medium-sized objects, which makes up for the number of small objects in the batch, makes the distribution of small objects more uniform, achieves the function of data expansion, and improves performance.

The work based on network level improvement includes:

FPN: adopt multi-scale feature fusion (before convolution operation in the current layer, add the sampling on the feature map of the previous layer and the feature map of the current layer, i.e. get the deep features by fusing the sampling on the feature map of the previous layer and the shallow features), so as to retain some detail information of the previous layer and improve the accuracy;

Each layer of FPN is predicted independently.

Tridentnet: use different size receptive fields for the same object to achieve data augmentation + share weight parameters to bring adaptability to various scales.

Adaptive Training Sample Selection：

Through ablation experiments, it is found that the fundamental reason why the performance of anchor free (ATSs) is better than that of anchor based (retinaet) is that the positive and negative sample selection strategies are different (retinaet defines the positive and negative samples through the IOU of anchor box and GT, and fcos defines the positive and negative samples through the location center).

1.3 **anchor free**Improvement based on

Corner (false free):

Cornernet = a set of points in the upper left corner and the lower right corner are used to replace the bounding box, and a new pooling method, cornernet pooling, is proposed.Center point (false free):

On the basis of cornernet, a center point prediction is added. The requirement to form an object is not only that two vertices can match, but also that the center of the frame with the meaning of the two center points should have corresponding center points, which can alleviate many strange false detections.Centernet also proposed the center

Pooling and cascade corner pooling to improve the accuracy of prediction for center points and vertices.

Reppoints (true free):

Directly predict 9 representative points (these vertices have no clear semantics), and then find the tightest frame surrounding the 9 points and calculate the loss with GT.Then loss will only return to those points that contribute to the generation of the box.

FSAF：

According to the loss of anchor free branch calculation, the feature graph is selected online. At the same time, it is proposed that the joint training of anchor based and anchor free can significantly improve the performance.

1.4 Common challenges

1. The anchor based method requires the anchor to cover all ground truth boxes and the IOU is relatively large, so it requires a lot of densely distributed

anchors, which leads to a serious imbalance between positive and negative samples.

2. Even if the small goal is included by anchor, the small goal may not be matched, which leads to the problem that the learning efficiency of the small goal is too low (see Figure 1).

Figure 1 the phenomenon of anchor mismatch

In Figure 1, the red box is gt; the green box is anchor; even if the GT is inside the anchor, because the IOU is too low, the anchor will still be misjudged as false positive.

3. Fast r-cnn baseline, in more than 50% iterations, the loss contribution of small objects is very low (less than 0.1 of the total loss), which will make the whole network tend to large and medium-sized objects, and the contribution of samples of different sizes to regression loss is unbalanced (see Figure 2).

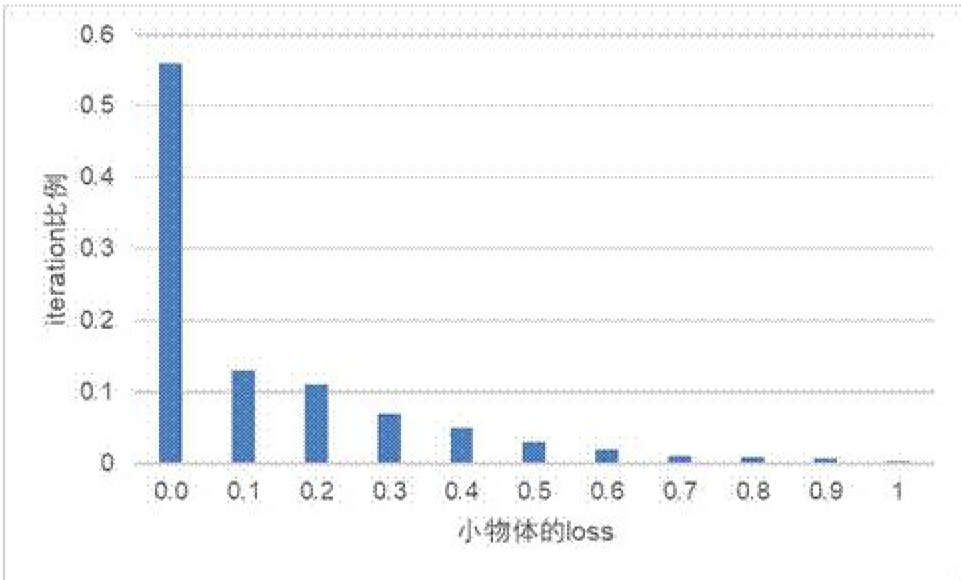

Fig. 2 loss distribution of small objects is uneven

1.5 Main contributions

In this paper, an anchor fitted: feedback driven target detection anchor arbiter model is proposed to solve the problem of small target learning and sample imbalance in target detection.The main contributions are as follows:

1.By making full use of the small target and anchor's scale loss as the feedback information to guide whether the anchor is modified to dynamically adjust the size of inappropriate anchor. As a hidden data enhancement method, it provides more excellent anchors.

2. The group Lou balance strategy ensures enough positive and negative samples with balanced proportion to participate in model training at each scale, so as to avoid the small gradient generated by simple samples being submerged by the large gradient generated by difficult samples.

3、**Motivation**

1. In the coco dataset, we found that there are many small targets (41.43%), but the distribution is very uneven, there are nearly 50% images without small targets (100-51.82 = 48.18).

2. Only 29.96% of them matched with small objects, while 44.49% matched with large objects.This means that each large object has 2.54 matching anchors, while each small object has only 1 Matching anchors.

3. In the above table, the anchor with the best matching of small and medium objects usually has a lower IOU value.The average maximum IOU of small objects is only 0.29, while the best matching anchor of medium and large objects is about twice that of small objects, 0.57 and 0.66 respectively.

	Object Count	Images	Total Object Area	Matched Anchors	Average matching anchors	Average max IoU
small	**41.43%**	**51.82%**	**1.23%**	**29.96%**	**1.00**	**0.29**
medium	34.32%	70.07%	10.18%	25.54%	1.03	0.57
large	24.24%	82.28%	88.59%	44.49%	2.54	0.66

Table 2 imbalance statistics of small objects

4. AP small is more than twice lower than AP large - > the learning of small objects is insufficient - > the learning of small objects needs to be strengthened.

AP	AP small	AP mid	AP large
36.7 %	21.1 %	39.9 %	48.1 %

Table 3 poor performance of small objects

4、 Main framework

4.1 General framework

1.Through the anchored module (red box), the anchor without the assign label is more suitable for the small target.

2. Group Lou balance sampling (green box) can avoid small gradients generated by simple samples being submerged by large gradients generated by difficult samples, so that the classifier will obtain balanced training samples of different sizes.

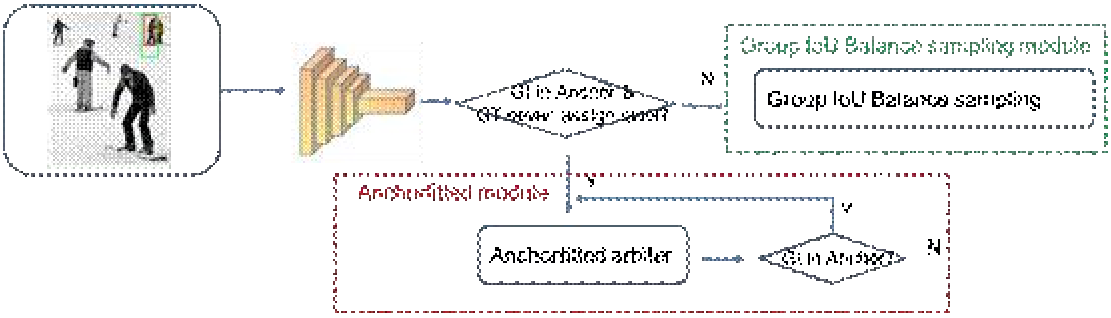

Figure 3 the overall framework of anchorfitted

4.2 Anchorfitted module

1. Introduce a consensus mechanism: if 1) GT is given a positive label for the first time; 2) GT is inside the anchor;To prevent the anchor from being missed, the anchor uses the anchored module

.2. Recursively scale the length and width of the anchor to 4 / 5, until the anchor no longer contains GT, and get the final anchor.

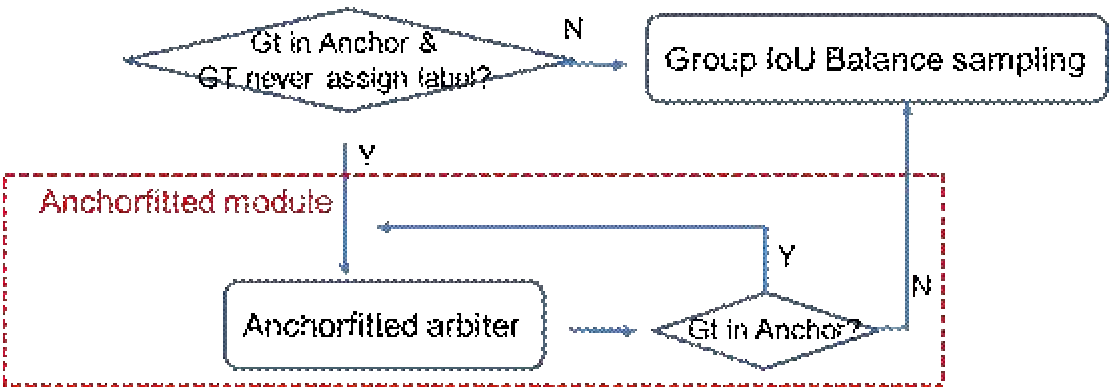

Figure 4 anchorfitted module framework

4.3 Group IoU Balance sampling

1 According to its scale, all the participating anchors are divided into several groups;

2. For each group: use IOU balanced sampling, keep the ratio of positive and negative samples at 3:1, and then

By dividing the value of IOU into k intervals, n negative samples are sampled in each interval, and the number of candidate samples in each interval is MK. the specific sampling formula is:

$$p_k = \frac{N}{K} * \frac{1}{M_k}, \quad k \in [0, K)$$

Figure 5 group IOU balance sampling formula

3. Through the uniform sampling on the IOU, the hard negative is evenly distributed on the IOU, so that the classifier will get the balanced training samples of different scales.

5、 Expected results

1. Anchorfitted feedback driver module solves the problem that the IOU of GT box and anchor is too small due to improper anchor, so proposal is lost, which can provide more excellent anchors.

2. A group IOU balance strategy is used to ensure a sufficient number of positive and negative samples with balanced proportion to participate in model training at each scale, so as to avoid the small gradient generated by simple samples being submerged by the large gradient generated by difficult samples.

6、 Summary

Anchorfitted = use anchor's scale loss to drive learning + group IOU balance sampling to balance learning.