BMVC 2024

Code Paper

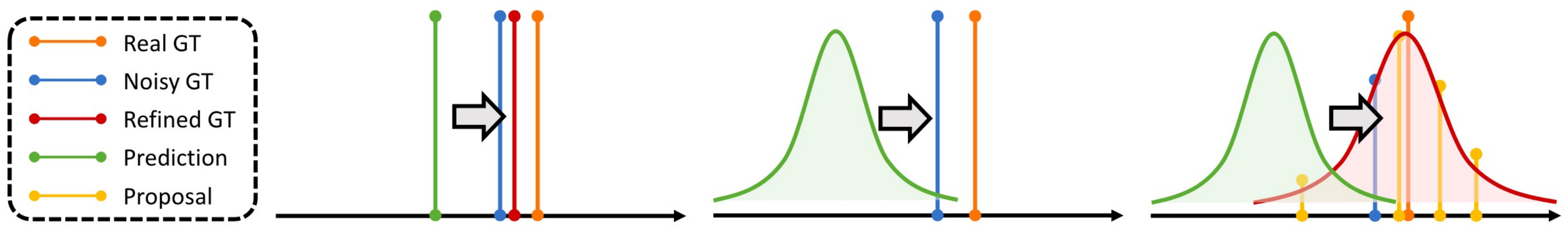
Distribution-Aware Calibration for Object Detection with Noisy Bounding Boxes

Donghao Zhou[1], Jialin Li[2], Jinpeng Li[1], Jiancheng Huang[3], Qiang Nie[4,2], Yong Liu[2†], Bin-Bin Gao[2], Qiong Wang[5], Pheng-Ann Heng[1], Guangyong Chen[6†]

[1] CUHK [2] Tencent YouTu Lab [3] SIAT, CAS [4] HKUST(GZ) [5] CV2R-Lab, SIAT, CAS [6] Zhejiang Lab

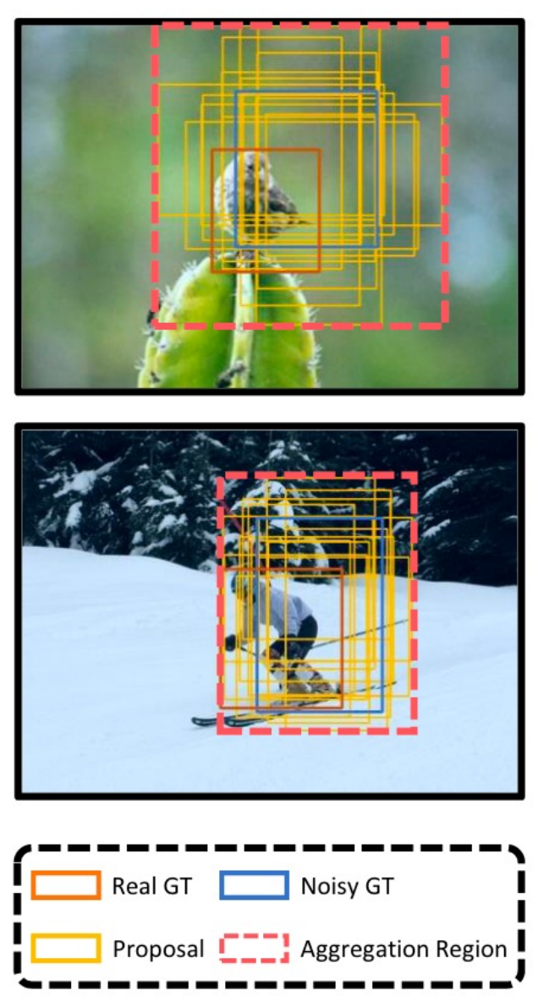
Task and Motivations

In this work, we focus on **object detection with noisy bounding boxes** (*i.e.*, the location of box annotations is inaccurate). (a) Encountering this task, existing solutions still exhibit drawbacks. (b) we observed that the spatial distribution of proposals can act as a statistical prior for the potential locations of objects.



Improvement	Box Correction	Confidence Estimation	DISCO (Ours)
Classification	✓	×	✓
Localization	✓	×	✓
Interpretability	×	✓	✓

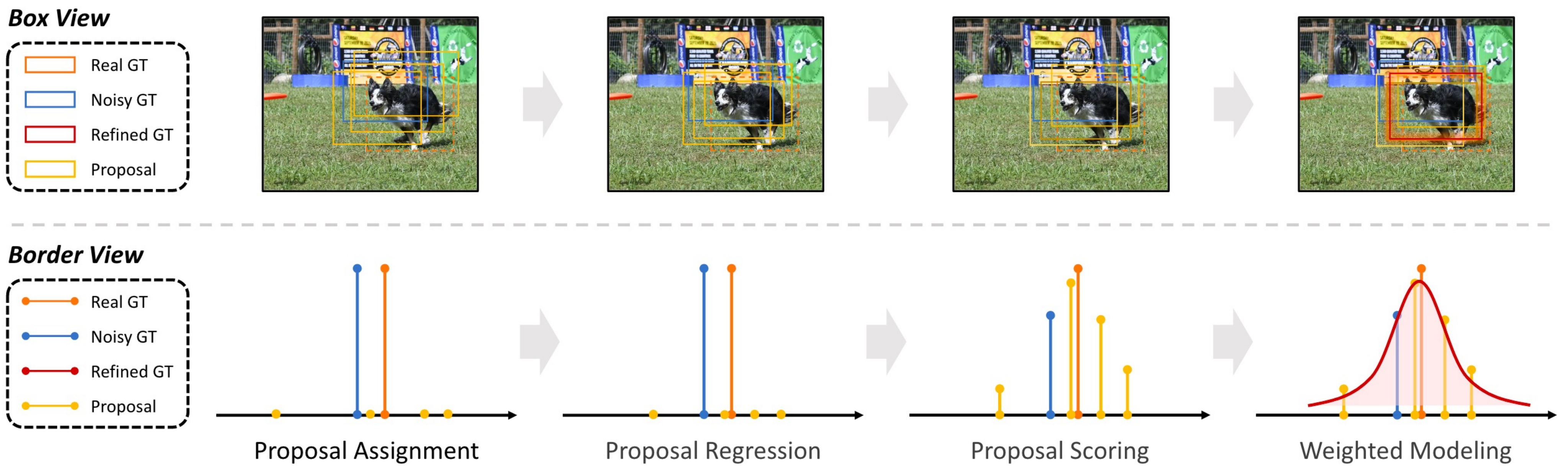
(a)



(b)

Methodology

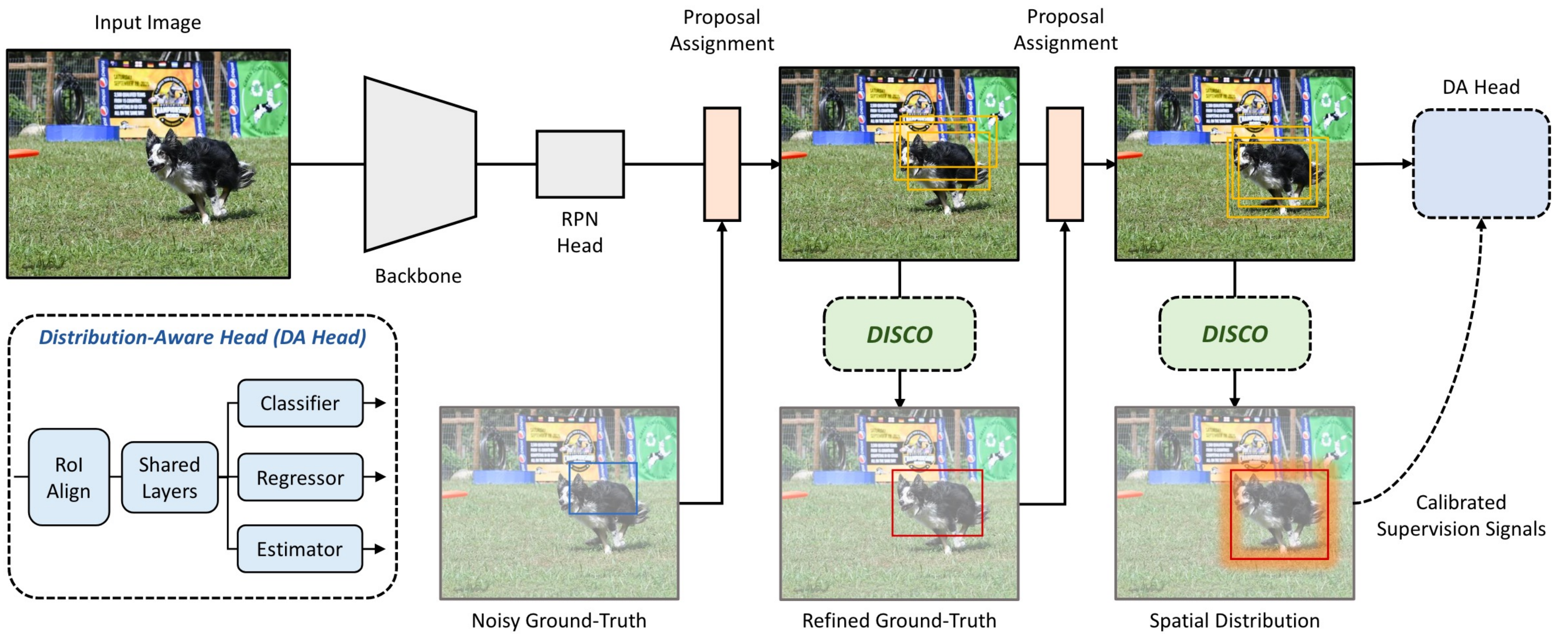
Based on our observation, we propose **DIStribution-aware CalibratiOn (DISCO)** to model the spatial distribution of proposals for calibrating supervision signals. First, **spatial distribution modeling** is performed to statistically extract the potential locations of objects.



Then, we develop three distribution-aware techniques to collaborate with spatial distribution modeling:

1. **Distribution-aware proposal augmentation (DA-Aug)**
2. **Distribution-aware box refinement (DA-Ref)**
3. **Distribution-aware confidence estimation (DA-Est)**

They are proposed to improve improve *classification*, *localization*, and *interpretability*, respectively (refer to our paper for more details).



Experiments

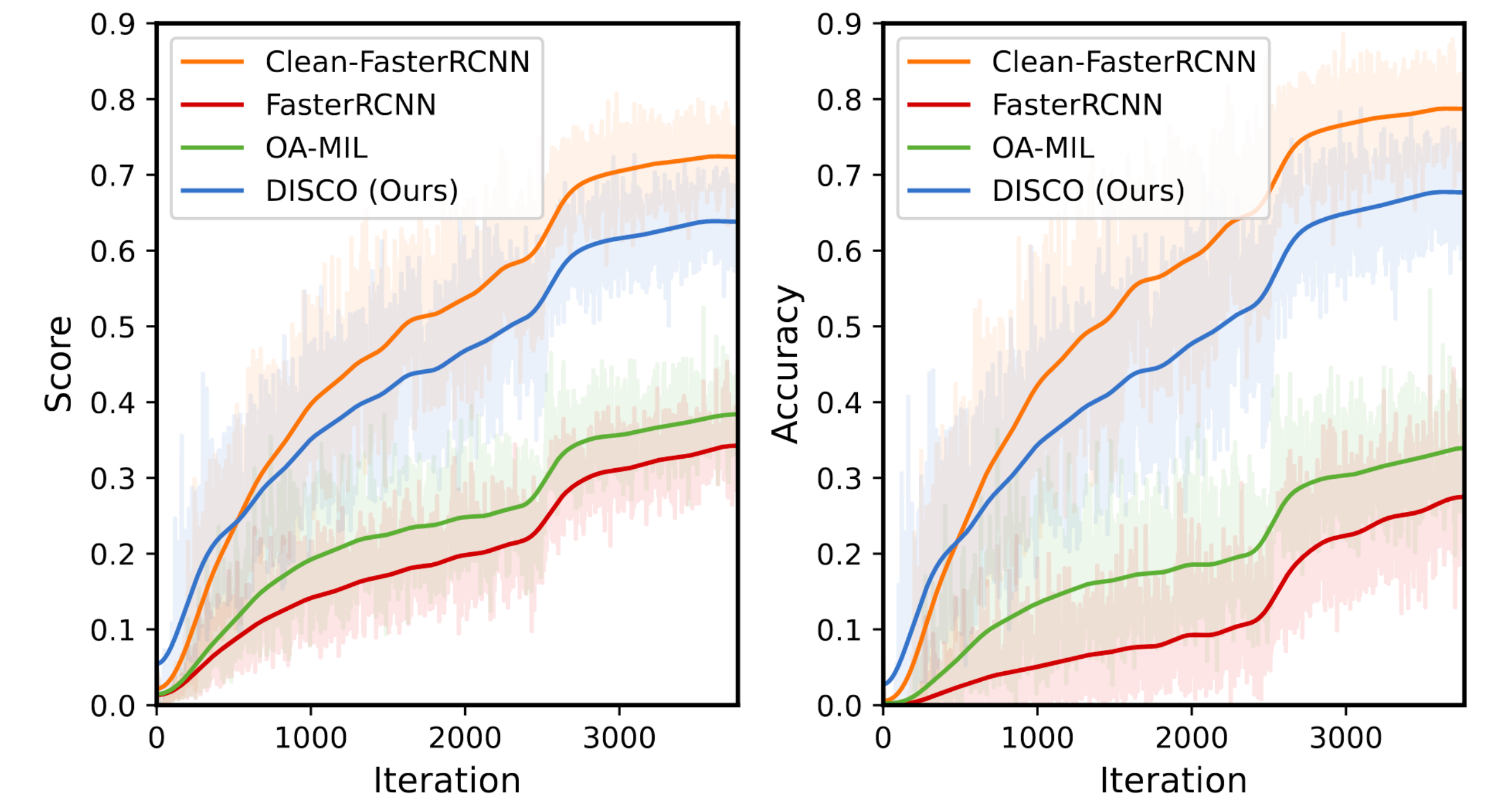
Our DISCO can achieve **SOTA results** on two large-scale datasets at various noisy ratio.

Method	VOC				COCO											
	Noise Level				20% Noise Level						40% Noise Level					
	10%	20%	30%	40%	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
Clean-FasterRCNN [28]	77.2	77.2	77.2	77.2	37.9	58.1	40.9	21.6	41.6	48.7	37.9	58.1	40.9	21.6	41.6	48.7
FasterRCNN [28]	76.3	71.2	60.1	42.5	30.4	54.3	31.4	17.4	33.9	38.7	10.3	28.9	3.3	5.7	11.8	15.1
RetinaNet [21]	71.5	67.5	57.9	45.0	30.0	53.1	30.8	17.9	33.7	38.2	13.3	33.6	5.7	8.4	15.9	18.0
Co-teaching [11]	75.4	70.6	60.9	43.7	30.5	54.9	30.5	17.3	34.0	39.1	11.5	31.4	4.2	6.4	13.1	16.4
SD-LocNet [37]	75.7	71.5	60.8	43.9	30.0	54.5	30.3	17.5	33.6	38.7	11.3	30.3	4.3	6.0	12.7	16.6
FreeAnchor [38]	73.0	67.5	56.2	41.6	28.6	53.1	28.5	16.6	32.2	37.0	10.4	28.9	3.3	5.8	12.1	14.9
KL Loss [14]	75.8	72.7	64.6	48.6	31.0	54.3	32.4	18.0	34.9	39.5	12.1	36.7	3.7	6.2	13.0	17.4
OA-MIL [22]	77.4	74.3	70.6	63.8	32.1	**55.3**	33.2	18.1	35.8	**41.6**	18.6	42.6	12.9	9.2	19.0	26.5
DISCO (Ours)	**77.5**	**75.3**	**72.1**	**68.7**	**32.3**	54.7	**34.5**	**18.7**	**35.8**	41.2	**21.2**	**45.7**	**16.9**	**11.4**	**24.7**	**27.8**

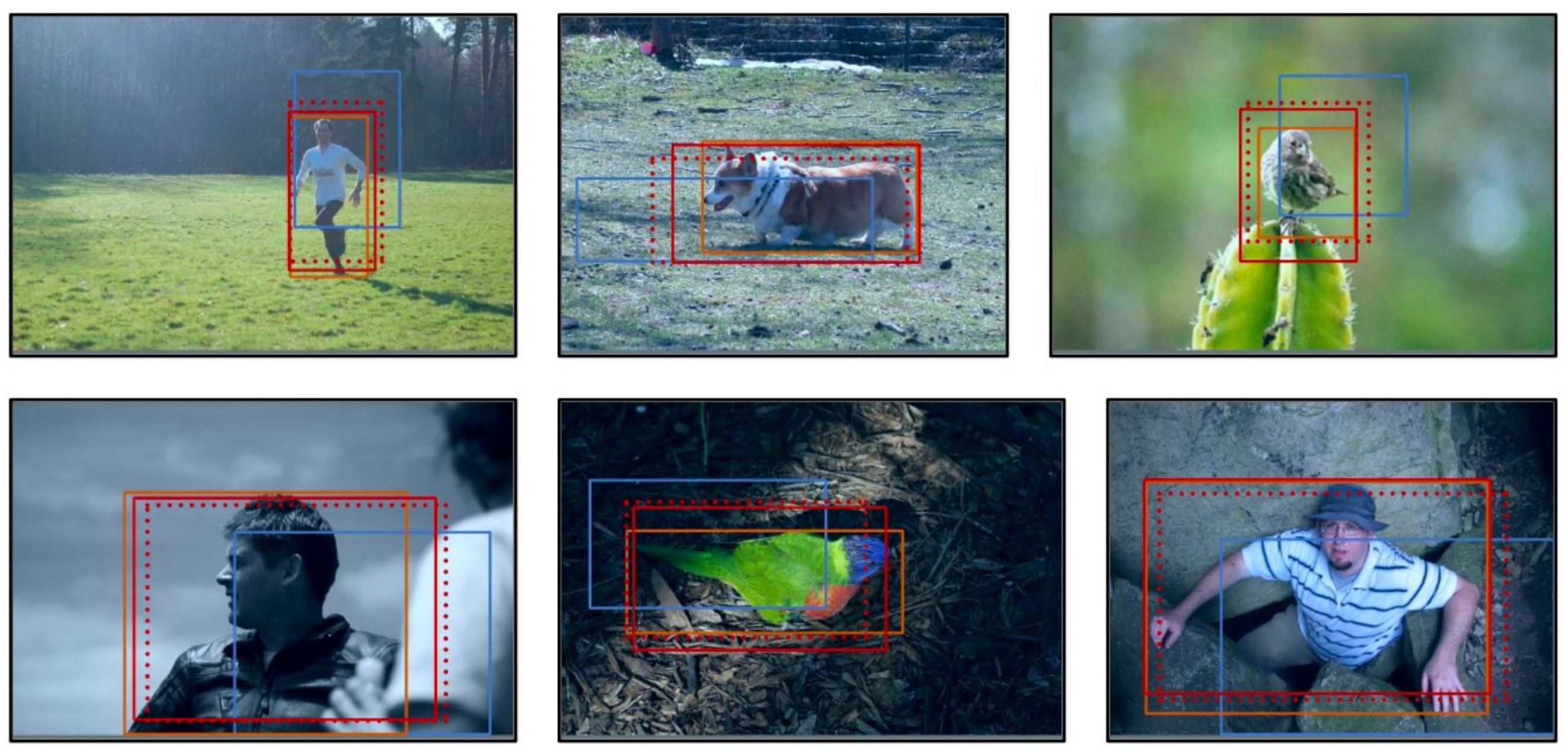
The proposed components of DISCO can **all significantly help in detection performance**.

Component			Category																				All
DA-Aug	DA-Ref	DA-Est	Aero	Bicy	Bird	Boat	Bot	Bus	Car	Cat	Cha	Cow	Dtab	Dog	Hors	Mbik	Pers	Plnt	She	Sofa	Trai	Tv	
✓			49.4	69.5	47.4	32.1	35.2	62.3	64.1	60.3	31.9	55.1	41.5	61.8	54.3	56.8	58.7	22.5	48.6	49.7	49.8	51.3	50.1
	✓		56.0	70.6	56.0	38.5	33.0	64.7	74.8	77.4	32.2	58.5	42.4	72.1	65.6	64.8	62.5	23.4	51.2	51.5	65.7	50.5	55.6
	✓	✓	61.2	74.4	59.7	43.1	37.0	69.4	75.2	73.3	34.8	64.1	54.5	74.1	71.7	66.0	66.7	28.7	54.1	55.4	70.5	60.2	59.7
✓	✓		69.9	**77.1**	68.2	**47.2**	49.9	70.9	80.6	80.8	43.0	76.4	60.0	**82.6**	81.0	74.4	73.4	39.2	62.7	64.3	67.9	68.6	66.9
✓	✓	✓	**71.5**	76.9	**71.5**	45.6	**52.2**	**76.1**	**81.2**	**83.2**	**43.4**	**79.8**	**60.3**	81.5	**82.9**	**75.4**	**73.6**	**40.6**	**64.4**	**68.2**	**76.8**	**70.0**	**68.7**

DISCO can **offer superior improvement for classification**, even approaching the results of training with clean annotations.



DISCO can **gradually refine noisy bounding boxes** during training.



DISCO can also **output reasonable variances** for each border of box prediction.