DEIM: DETR with Improved Matching for Fast Convergence

Project Page: https://www.shihuahuang.cn/DEIM/

Shihua Huang, Zhichao Lu, Xiaodong Cun, Yongjun Yu, Xiao Zhou, Xi Shen†
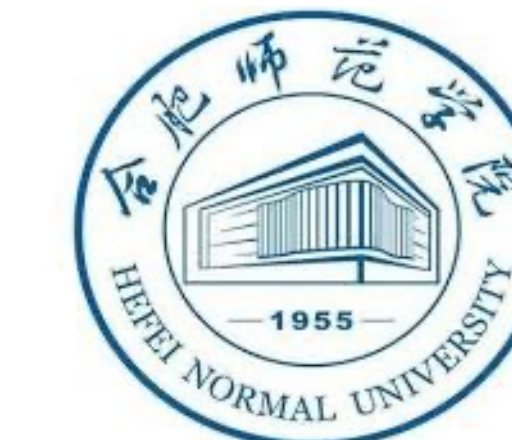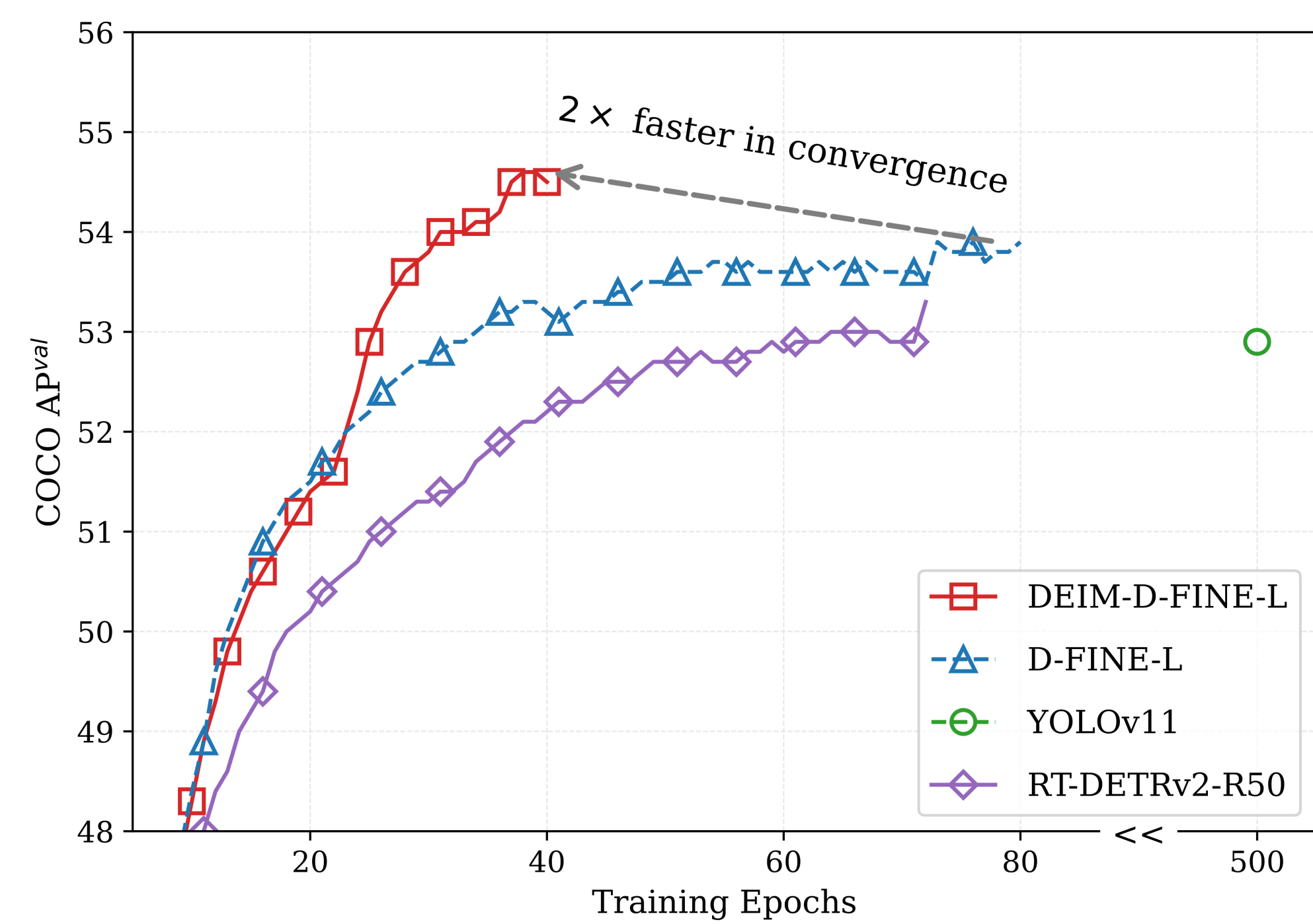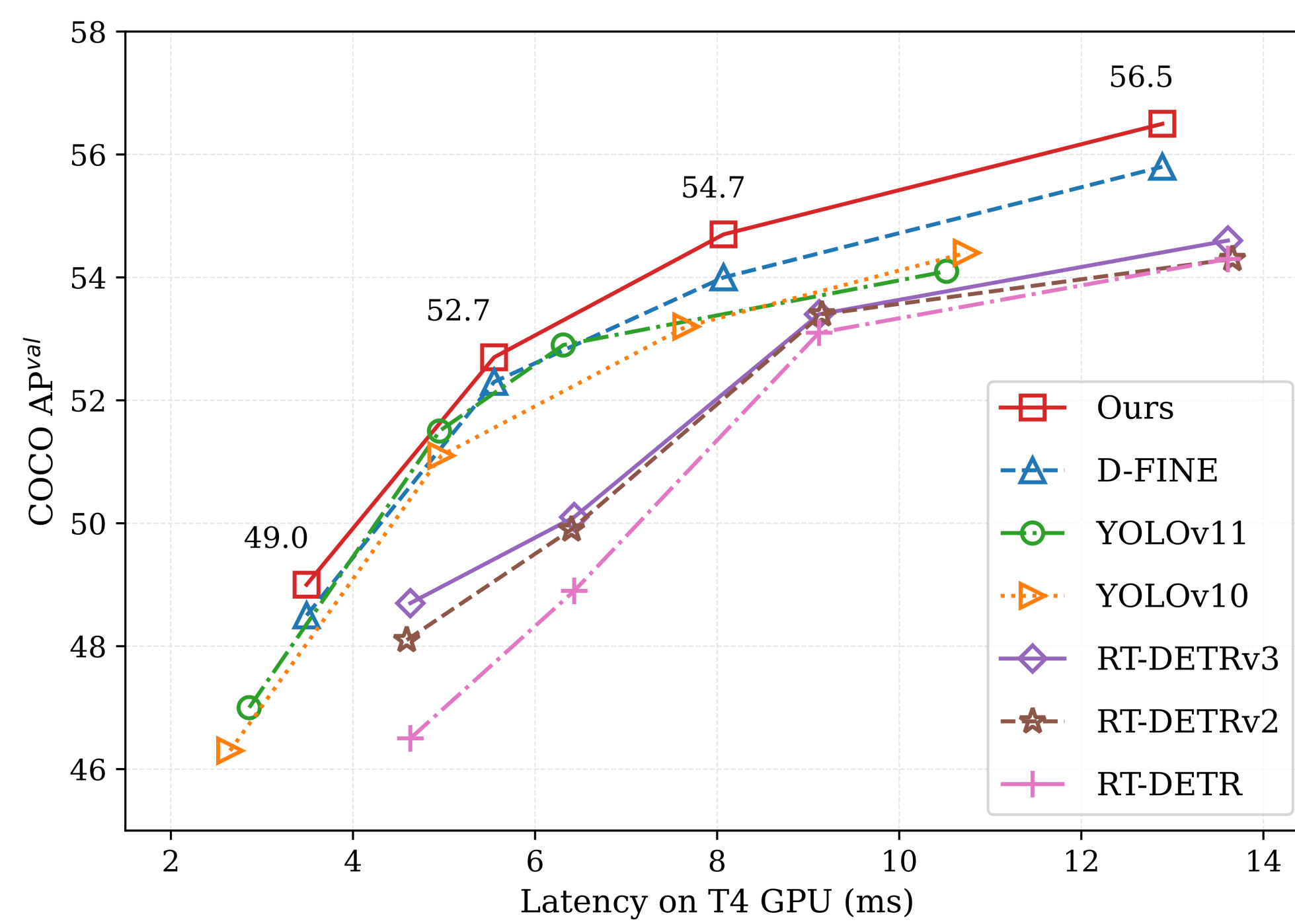
Highlight

DEIM is an advanced training framework designed to enhance the matching mechanism in DETRs, enabling **faster convergence** and **improved performance**.
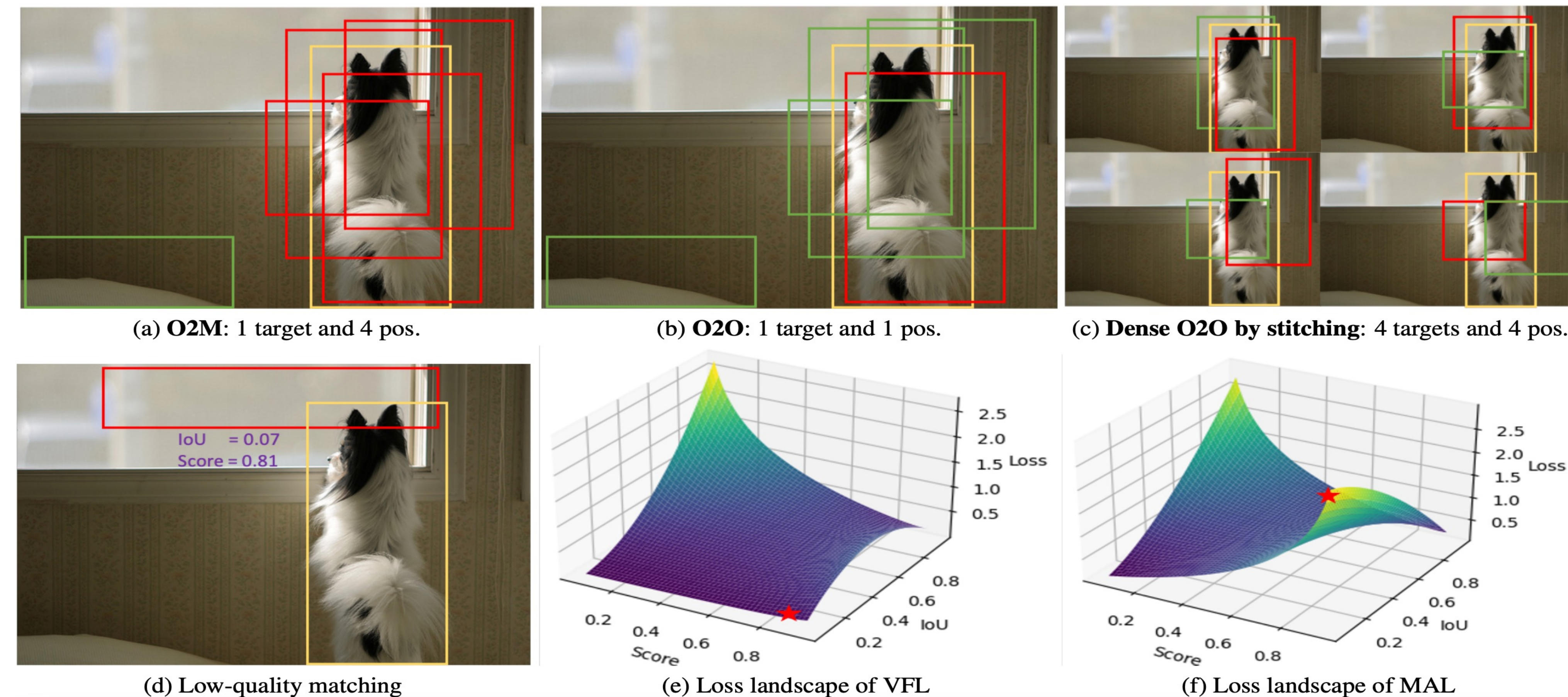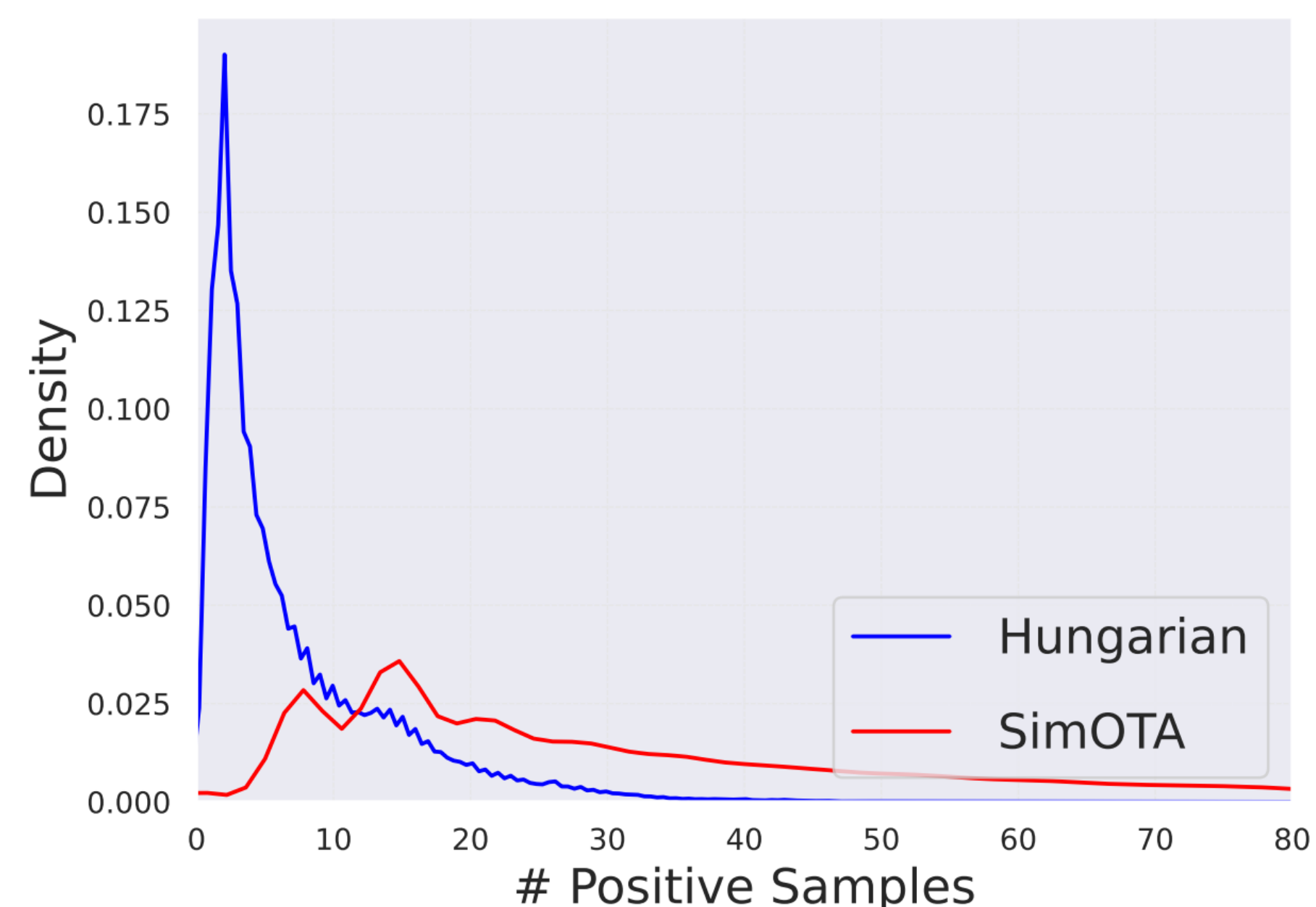

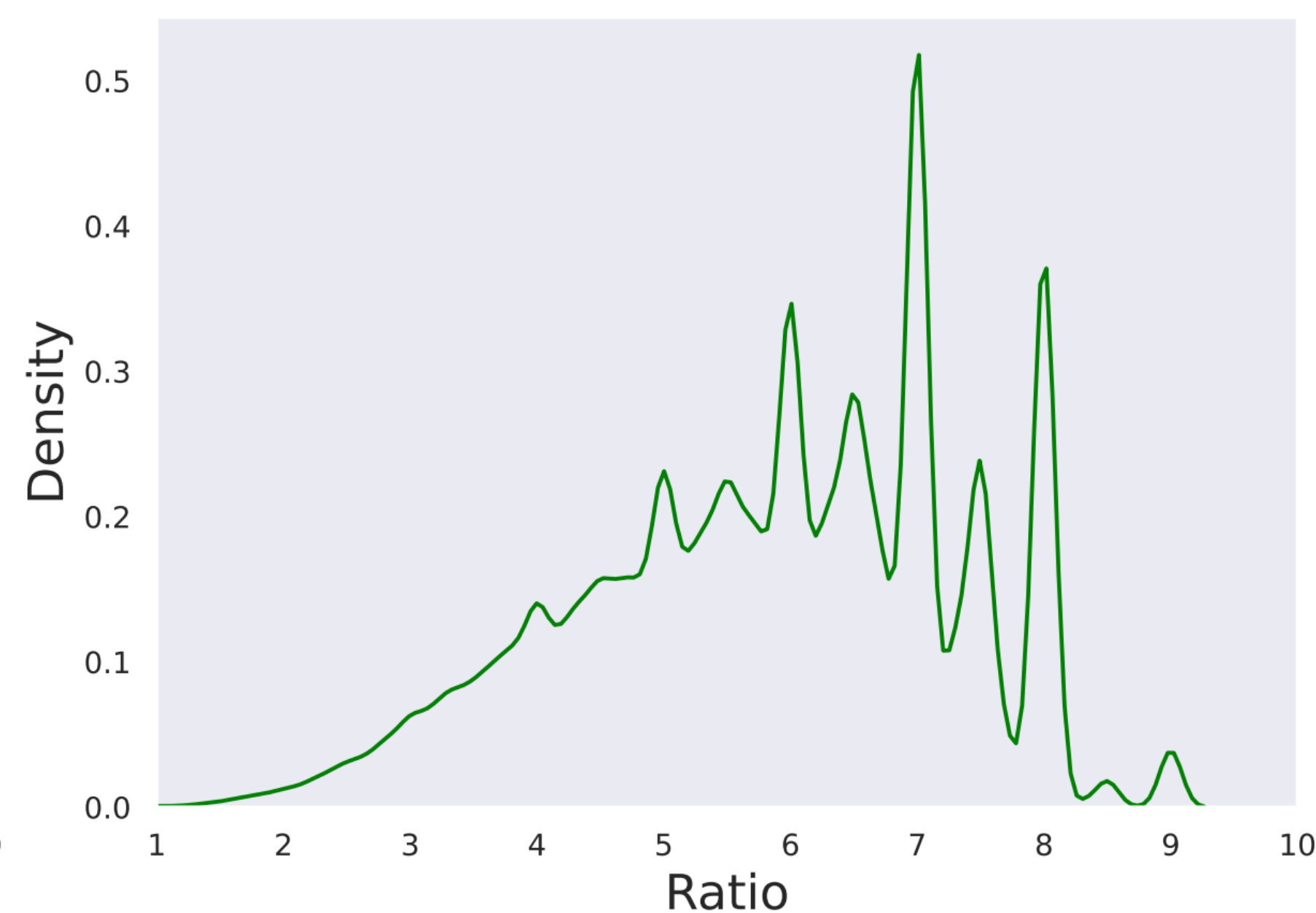**Faster Convergence**

Improved Performance

Key Idea: Dense O2O (one-to-one) with MAL (Matchability-Aware Loss)

(a) **O2M**: 1 target and 4 pos.

(b) **O2O**: 1 target and 1 pos.

(c) **Dense O2O by stitching**: 4 targets and 4 pos.

(d) Low-quality matching

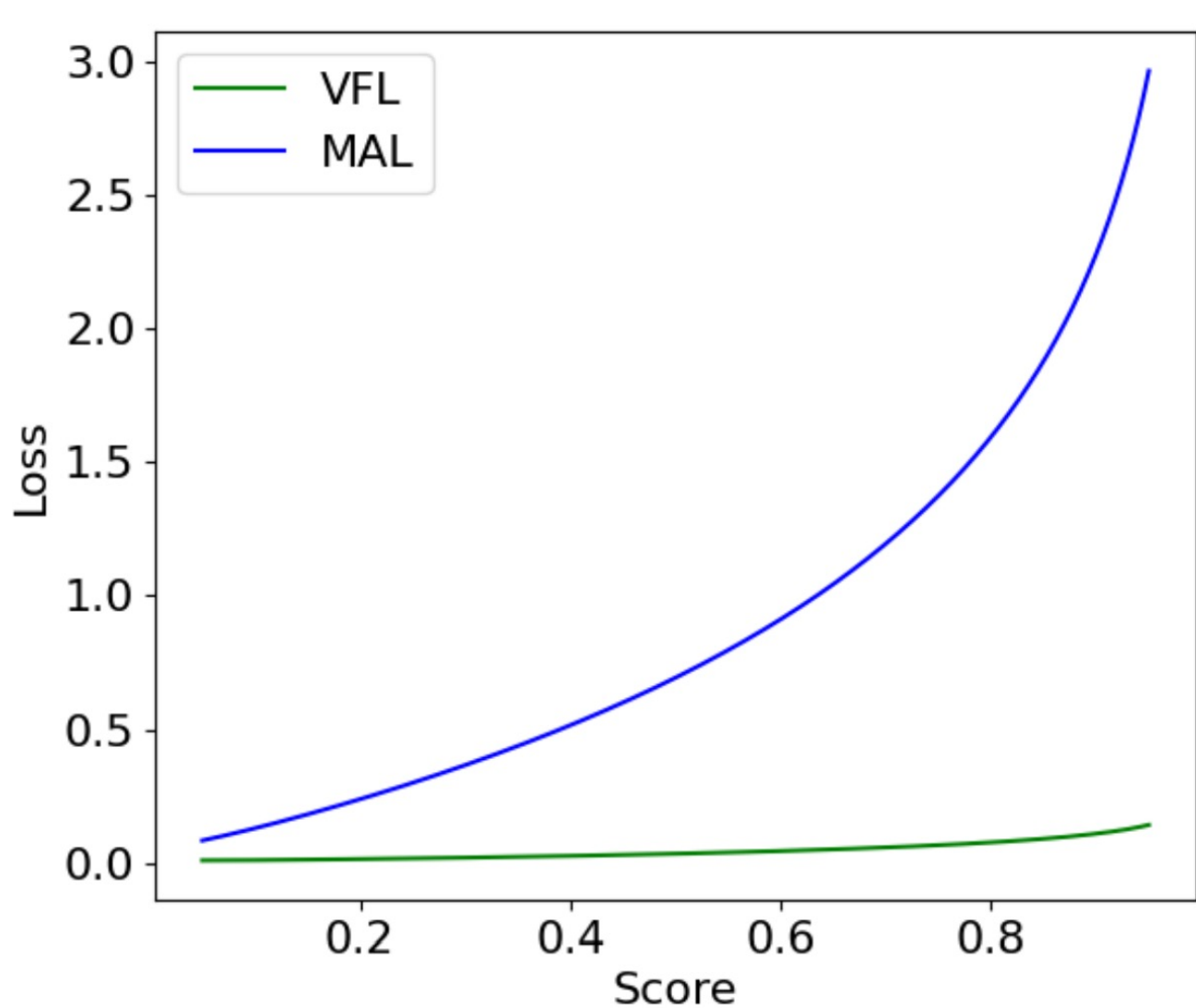(e) Loss landscape of VFL

(f) Loss landscape of MAL

Analysis

(a) Matching distribution

(b) Ratios between O2M and O2O

O2M (One-to-many, such as SimOTA) has more matched anchors than O2O (such as Hungarian)

(a) **Low quality**: IoU = 0.05

(b) **High quality**: IoU = 0.95

MAL embeds matching quality into the loss, making it more sensitive to low-IoU matches.

$$\text{MAL}(p, q, y) = \begin{cases} -q^\gamma \log(p) + (1 - q^\gamma) \log(1 - p) & y = 1 \\ -p^\gamma \log(1 - p) & y = 0 \end{cases}$$

Experiments on COCO Val 2017

Model	#Epochs	#Params	GFLOPs	Latency (ms)	AP^{val}	AP^{val}_{50}	AP^{val}_{75}	AP^{val}_S	AP^{val}_M	AP^{val}_L
YOLO-based Real-time Object Detectors										
YOLOv8-L [12]	500	43	165	12.31	52.9	69.8	57.5	35.3	58.3	69.8
YOLOv8-X [12]	500	68	257	16.59	53.9	71.0	58.7	35.7	59.3	70.7
YOLOv9-C [34]	500	25	102	10.66	53.0	70.2	57.8	36.2	58.5	69.3
YOLOv9-E [34]	500	57	189	20.53	55.6	72.8	60.6	40.2	61.0	71.4
Gold-YOLO-L [33]	300	75	152	9.21	53.3	70.9	-	33.8	58.9	69.9
YOLOv10-L* [32]	500	24	120	7.66	53.2	70.1	58.1	35.8	58.5	69.4
YOLOv10-X* [32]	500	30	160	10.74	54.4	71.3	59.3	37.0	59.8	70.9
YOLO11-L* [13]	500	25	87	6.31	52.9	69.4	57.7	35.2	58.7	68.8
YOLO11-X* [13]	500	57	195	10.52	54.1	70.8	58.9	37.0	59.2	69.7
DETR-based Real-time Object Detectors										
RT-DETR-HG-L [43]	72	32	107	8.77	53.0	71.7	57.3	34.6	57.4	71.2
RT-DETR-HG-X [43]	72	67	234	13.51	54.8	73.1	59.4	35.7	59.6	72.9
D-FINE-L [27]	72	31	91	8.07	54.0	71.6	58.4	36.5	58.0	71.9
DEIM-D-FINE-L	**50**	31	91	8.07	**54.7**	72.4	59.4	36.9	59.6	71.8
D-FINE-X [27]	72	62	202	12.89	55.8	73.7	60.2	37.3	60.5	73.4
DEIM-D-FINE-X	**50**	62	202	12.89	**56.5**	74.0	61.5	38.8	61.4	74.2

Dense O2O and MAL enable faster convergence as well as better performance

Epochs	Dense O2O	MAL	AP	AP_{50}	AP_{75}
RT-DETRv2-R50 [24]					
72			53.4	71.6	57.4
36	✓		53.6	71.9	58.2
	✓	✓	**53.9**	**71.7**	**58.6**
D-FINE-L [27]					
72			54.0	71.6	58.4
36	✓		54.2	72.1	58.9
	✓	✓	**54.6**	**72.2**	**59.5**