JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

DiGeo: Discriminative Geometry-Aware Learning for Generalized Few-Shot Object Detection

Jiawei Ma, Yulei Niu, Jincheng Xu, Shiyuan Huang, Guangxing Han, Shih-Fu Chang

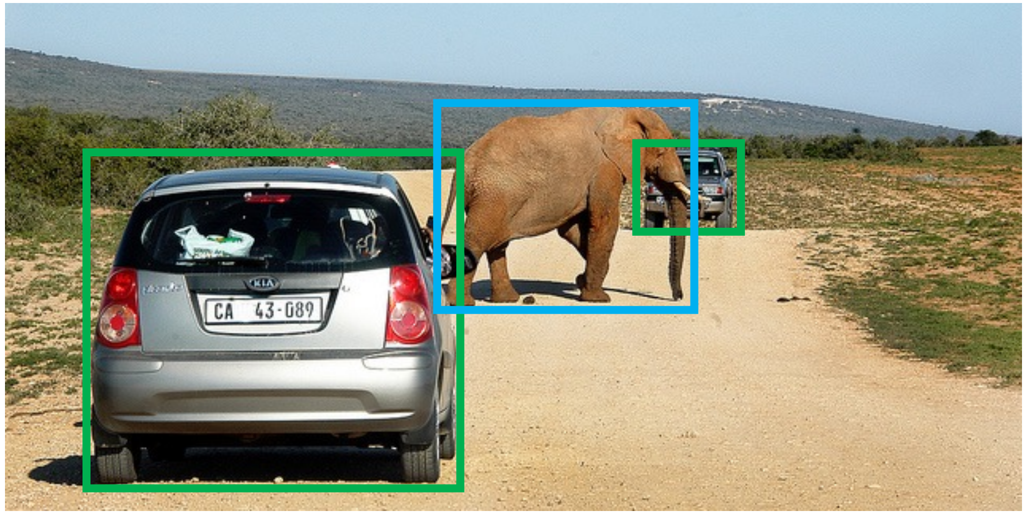
Digital Video | Multimedia Laboratory

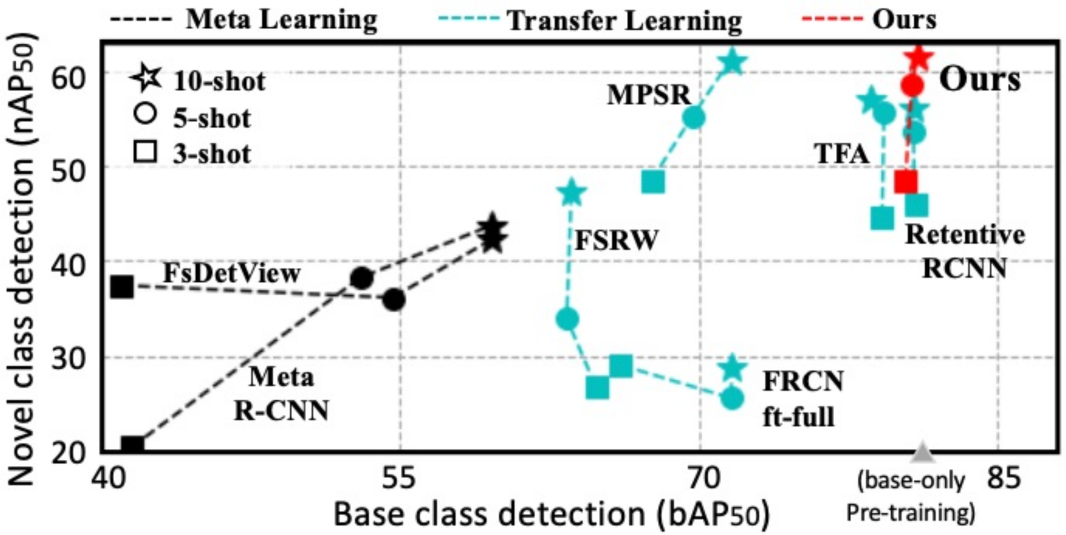
Poster ID: TUE-AM-305

Background

Generalized Few-shot Object Detection

The detection prediction for base classes (Car) and novel classes (elephant) should be naturally considered.



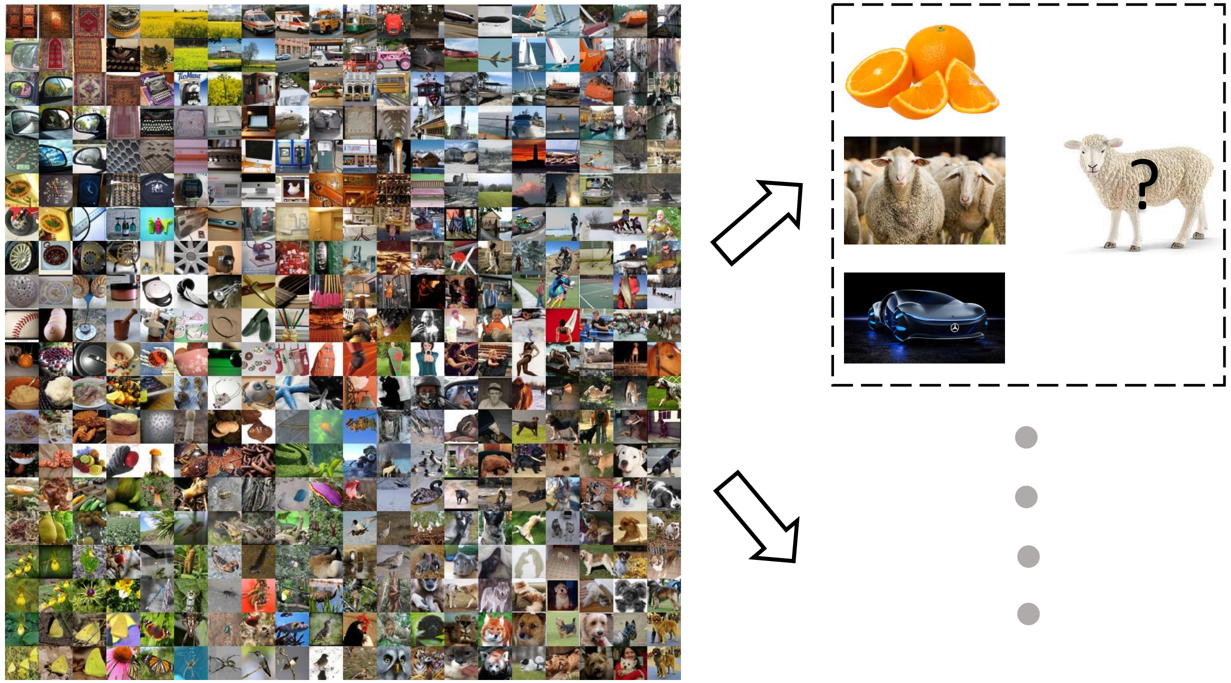


Learning features with ***discriminative geometry*** is important to detect all classes.

Background

Few-shot Learning

Given a base dataset to pre-train a feature extractor/detector, the learned model is adapted to new unseen classes with very limited training data.



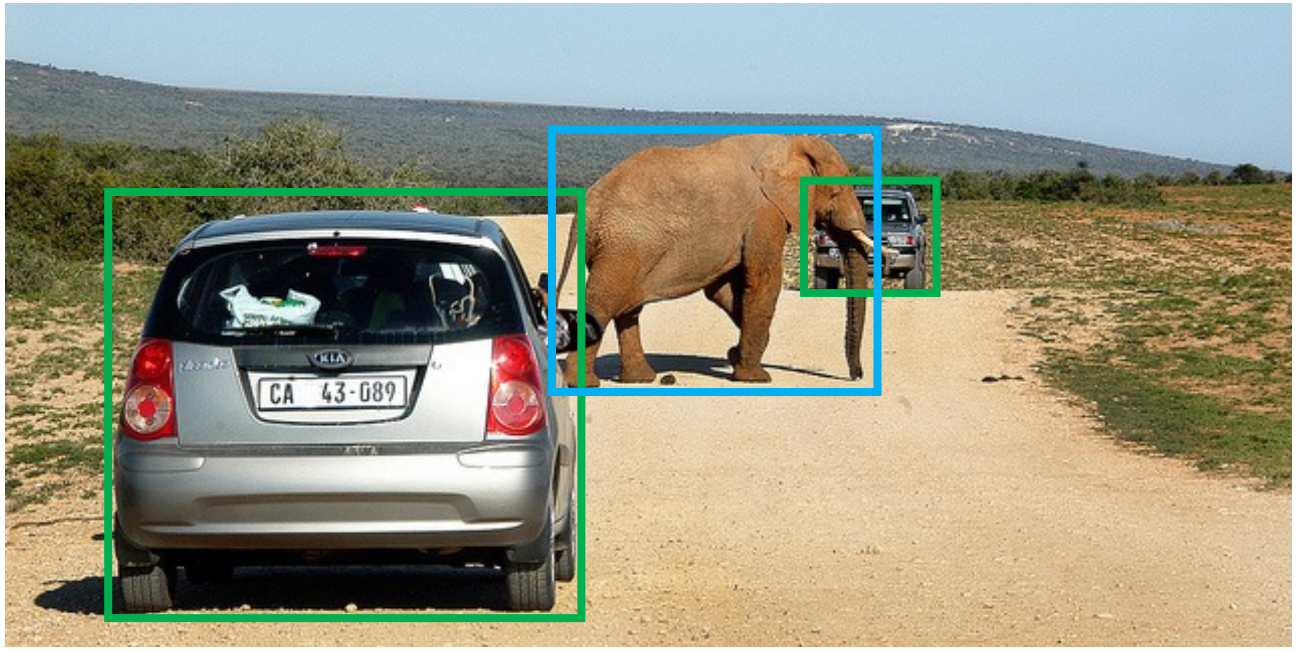
Base set

Few-shot task

Background

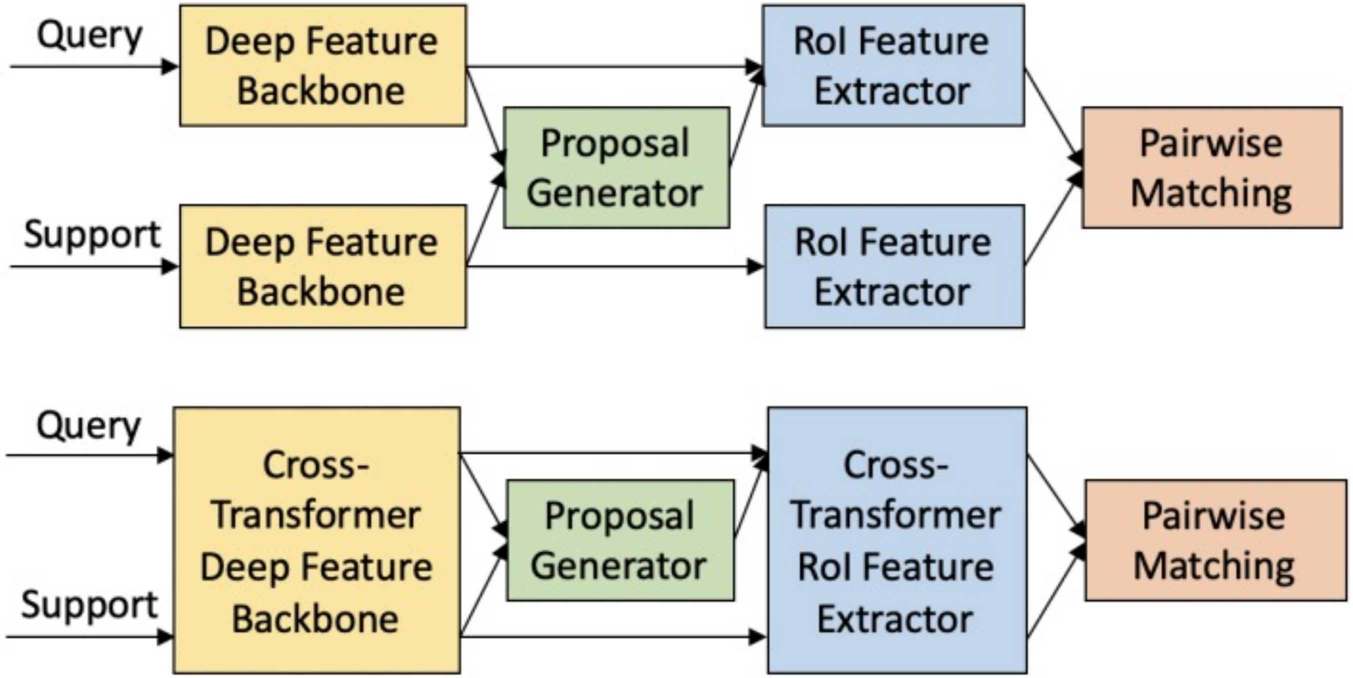
Generalized Few-shot Object Detection

For object detection, as an image may contain instances of both base (car) classes and novel (elephant) classes, the detection prediction on all classes should be naturally considered.



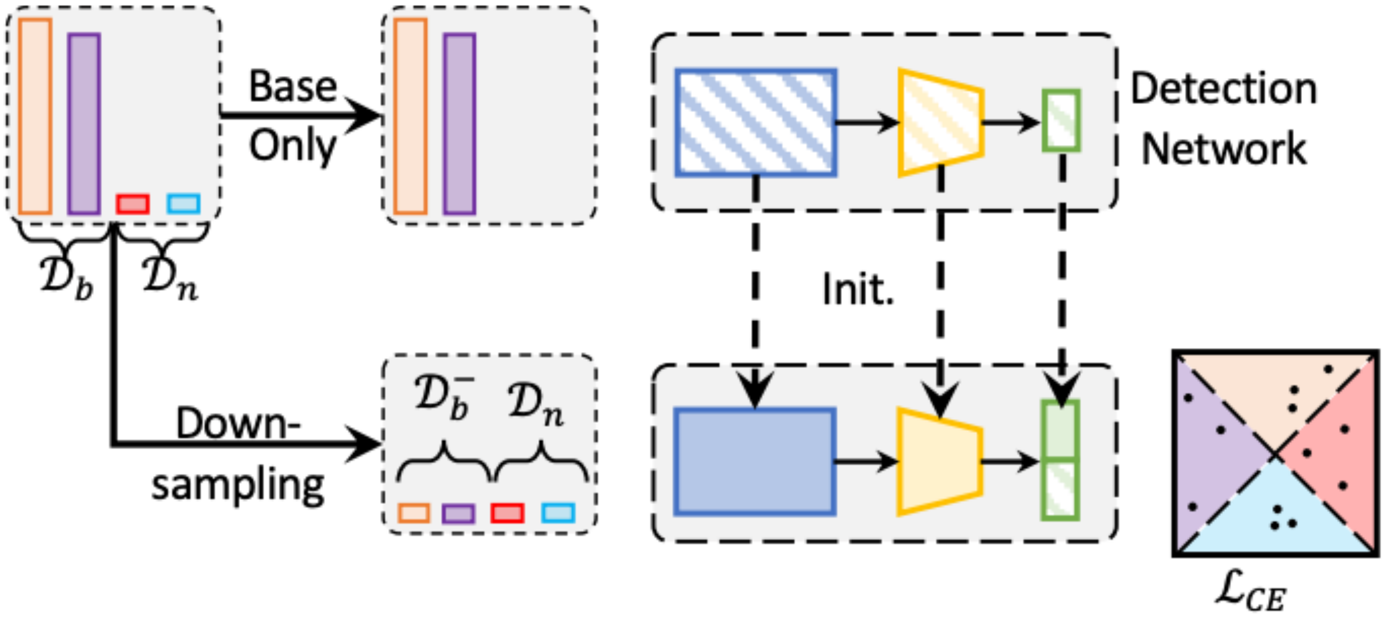
Rethinking Conventional Training Framework

(Pairwise Matching) Conventional meta-learning-based approaches focus on alignment between the test image and the limited training data. However, such training strategy is not efficient and significantly hurt the based detection.



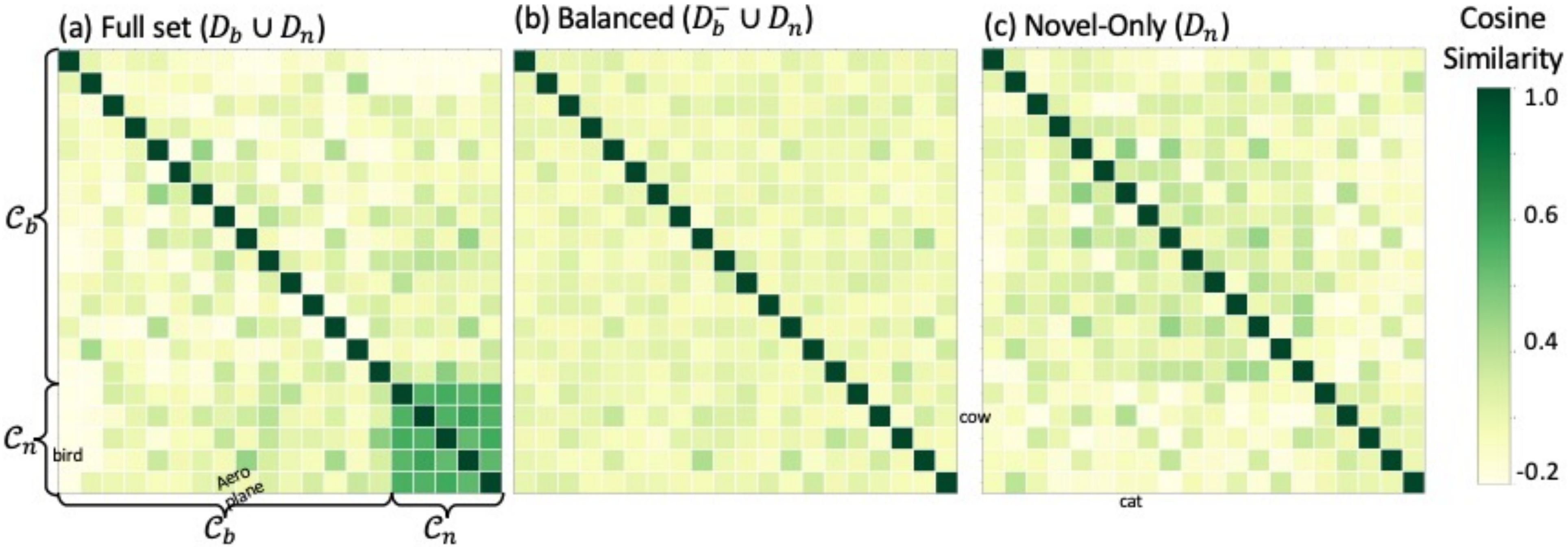
Rethinking Conventional Training Framework

(Two-Stage Finetuning) To adapt to new dataset, the Transfer Learning approaches finetune the pre-trained base detector on a balanced dataset, but suffer from base forgetting. Setting different classifiers for base and novel classes still hurt the adaptation efficiency.



Motivation

Though suffering from base forgetting, building a balanced dataset among all classes (base + novel) is essential to learn a well-separated classifier for object detection.



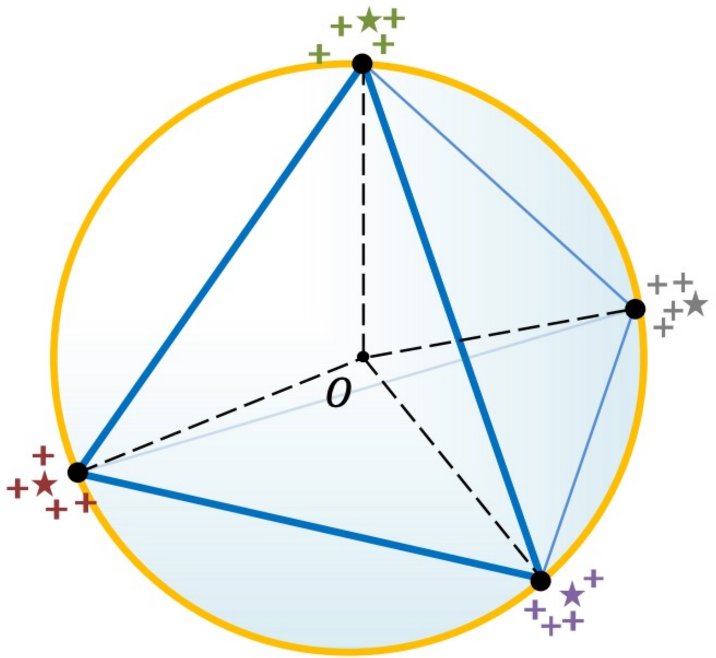
Motivation

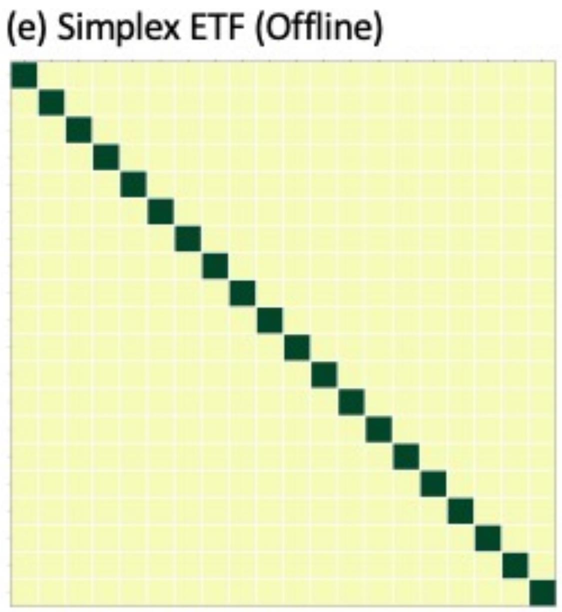
The instance-center similarity on the few-shot classes is not even properly converged

Approach

Discriminative Geometry-Aware Learning

Inter-class separation: Use Simplex ETF as a fixed classifier for all classes during training.

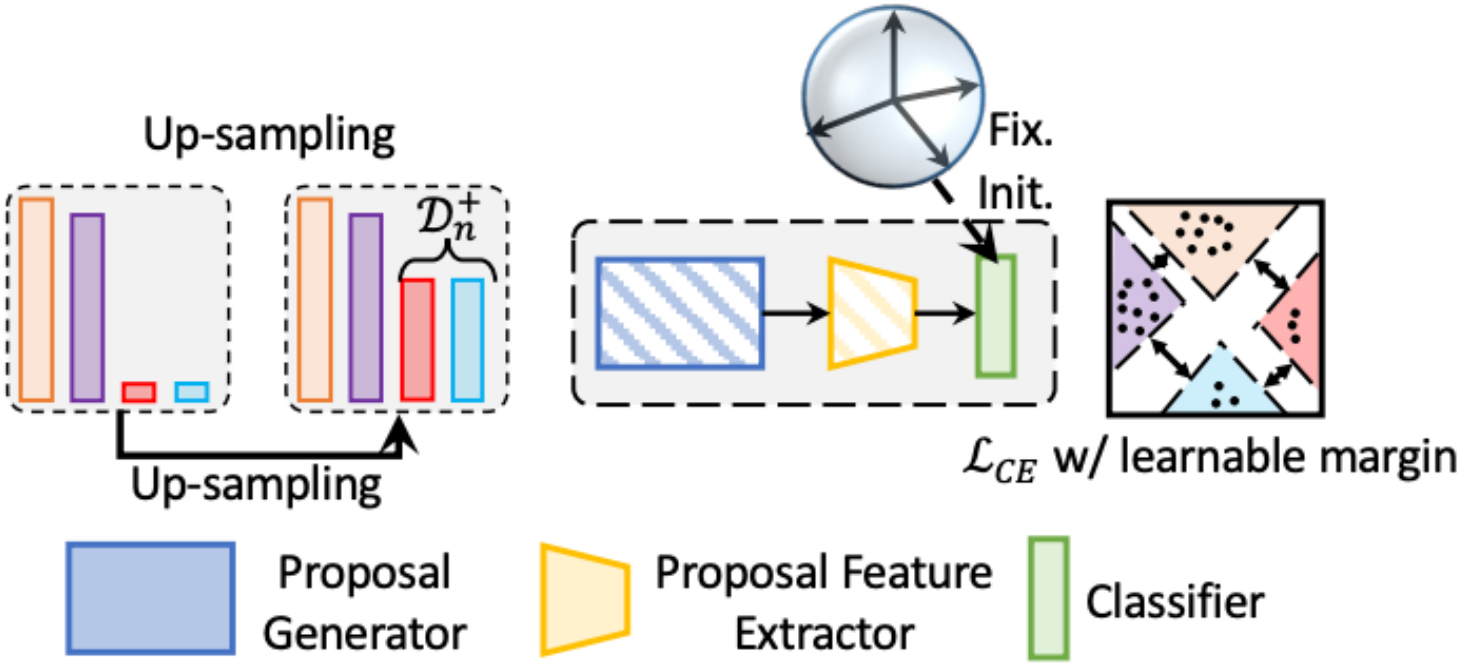


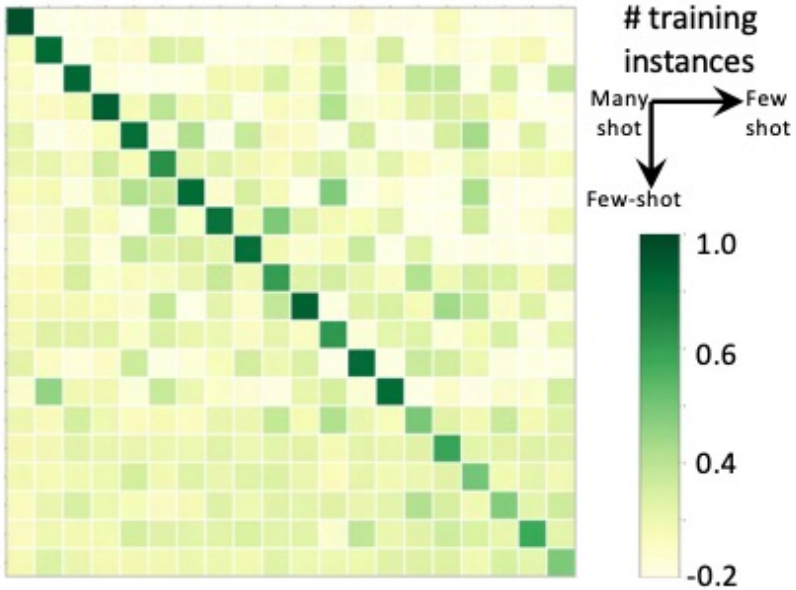


Discriminative Geometry-Aware Learning

Inter-class separation: Use Simplex ETF as a fixed classifier for all classes during training.

Intra-class compactness: Apply class-specific margins to push features close to the class centers.





Experimental Study

Our one-stage approach (self-distillation applied) can improve the detection precision on all classes.

Approach	split 1					split 2					split 3					Avg
	1	2	3	5	10	1	2	3	5	10	1	2	3	5	10	
Meta-Learning Approaches																
Meta RCNN [59]*	17.5	30.5	36.2	49.3	55.6	19.4	33.2	34.8	44.4	53.9	20.3	3.0	41.2	48.0	55.1	38.0
FSRW [39]	53.5	50.2	55.3	56.0	59.5	55.1	54.2	55.2	57.5	58.9	54.2	53.5	54.7	58.6	57.6	55.6
FsDetView [58]*	36.4	40.3	40.1	50.0	55.3	36.3	43.7	41.6	45.8	54.1	37.0	39.5	40.7	50.7	54.8	44.4
Transfer-Learning Approaches																
TFA w/ fc [55]	69.3	66.9	70.3	73.4	73.2	64.7	66.3	67.7	68.3	68.7	67.8	68.9	70.8	72.3	72.2	69.5
TFA w/ cos [55]	69.7	68.2	70.5	73.4	72.8	65.5	65.0	67.7	68.0	68.6	67.9	68.6	71.0	72.5	72.4	69.5
FRCN-ft-full [59]*	55.4	57.1	56.8	60.1	60.9	50.1	53.7	53.6	55.9	55.5	58.5	59.1	58.7	61.8	60.8	57.2
MPSR [57]	56.8	60.4	62.8	66.1	69.0	53.1	57.6	62.8	64.2	66.3	55.2	59.8	62.7	66.9	67.7	62.1
Retentive R-CNN [6]†	**71.3**	**72.3**	72.1	74.0	74.6	66.8	**68.4**	70.2	70.7	71.5	**69.0**	**70.9**	72.3	73.9	74.1	71.5
DiGeo (Ours)	69.7	70.6	**72.4**	**75.4**	**76.1**	**67.5**	**68.4**	**71.4**	**71.6**	**73.6**	68.6	**70.9**	**72.9**	**74.4**	**75.0**	**71.9**

*: results reported by Retentive R-CNN [6] and TFA [55]. †: Model ensembling.

Experimental Study

Our approach significantly improve the adaptation efficiency when more training data is available under the few-shot setup.

Approach	split 1					split 2					split 3					Avg.
	1	2	3	5	10	1	2	3	5	10	1	2	3	5	10	
FRCN-ft-full [59]*	15.2	20.3	29	25.5	28.7	13.4	20.6	28.6	32.4	38.8	19.6	20.8	28.7	42.2	42.1	27.1
TFA w/ fc [55]	36.8	29.1	43.6	55.7	57	18.2	29	33.4	35.5	39.0	27.7	33.6	42.5	48.7	50.2	38.7
TFA w/ cos [55]	39.8	36.1	44.7	55.7	56	23.5	26.9	34.1	35.1	39.1	30.8	34.8	42.8	49.5	49.8	39.9
MPSR [57]	**42.8**	43.6	48.4	55.3	61.2	**29.8**	28.1	41.6	**43.2**	47.0	**35.9**	40.0	43.7	48.9	51.3	**44.0**
Meta RCNN [59]*	16.8	20.1	20.3	38.2	43.7	7.7	12.0	14.9	21.9	31.1	9.2	13.9	26.2	29.2	36.2	22.8
FSRW [39]	14.8	15.5	26.7	33.9	47.2	15.7	15.3	22.7	30.1	39.2	19.2	21.7	25.7	40.6	41.3	27.3
FsDetView [58]*	25.4	20.4	37.4	36.1	42.3	22.9	21.7	22.6	25.6	29.2	32.4	19.0	29.8	33.2	39.8	29.2
Retentive R-CNN [6]	42.4	**45.8**	45.9	53.7	56.1	21.7	27.8	35.2	37.0	40.3	30.2	37.6	43.0	49.7	50.1	41.1
Prior	35.8	37.9	45.7	56.4	61.0	22.7	28.4	39.8	41.4	48.8	30.8	36.4	45.4	52.4	53.8	42.4
DiGeo	37.9	39.4	**48.5**	**58.6**	**61.5**	26.6	**28.9**	**41.9**	42.1	**49.1**	30.4	**40.1**	**46.9**	**52.7**	**54.7**	**44.0**

*: results reported by Retentive R-CNN [6] and TFA [55].†: Model ensembling. Full tables can be found in Supp.

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

DiGeo: Discriminative Geometry-Aware Learning for Generalized Few-Shot Object Detection

Contact: {jiawei.m, yulei.niu}@columbia.edu

GitHub: https://github.com/Phoenix-V/DiGeo