dongguk UNIVERSITY

META-RCNN: META LEARNING FOR FEW-SHOT OBJECT DETECTION

2021.09.8

Ershang Tian

Table of Contents

동국대학교
dongguk university

- BACKGROUND
- METHOD
- EXPERIMENTS
- CONCLUSION

BACKGROUND

Meta-RCNN

Meta-RCNN learns an object detector in an episodic learning paradigm on the (meta) training data.

This learning scheme helps acquire a prior which enables Meta-RCNN to do few-shot detection on novel tasks.

Built on top of the Faster RCNN model, in Meta-RCNN, both the Region Proposal Network (RPN) and the object classification branch are meta-learned.

The meta-trained RPN learns to provide class-specific proposals, while the object classifier learns to do few-shot classification. The novel loss objectives and learning strategy of Meta-RCNN can be trained in an end-to-end manner.

BACKGROUND

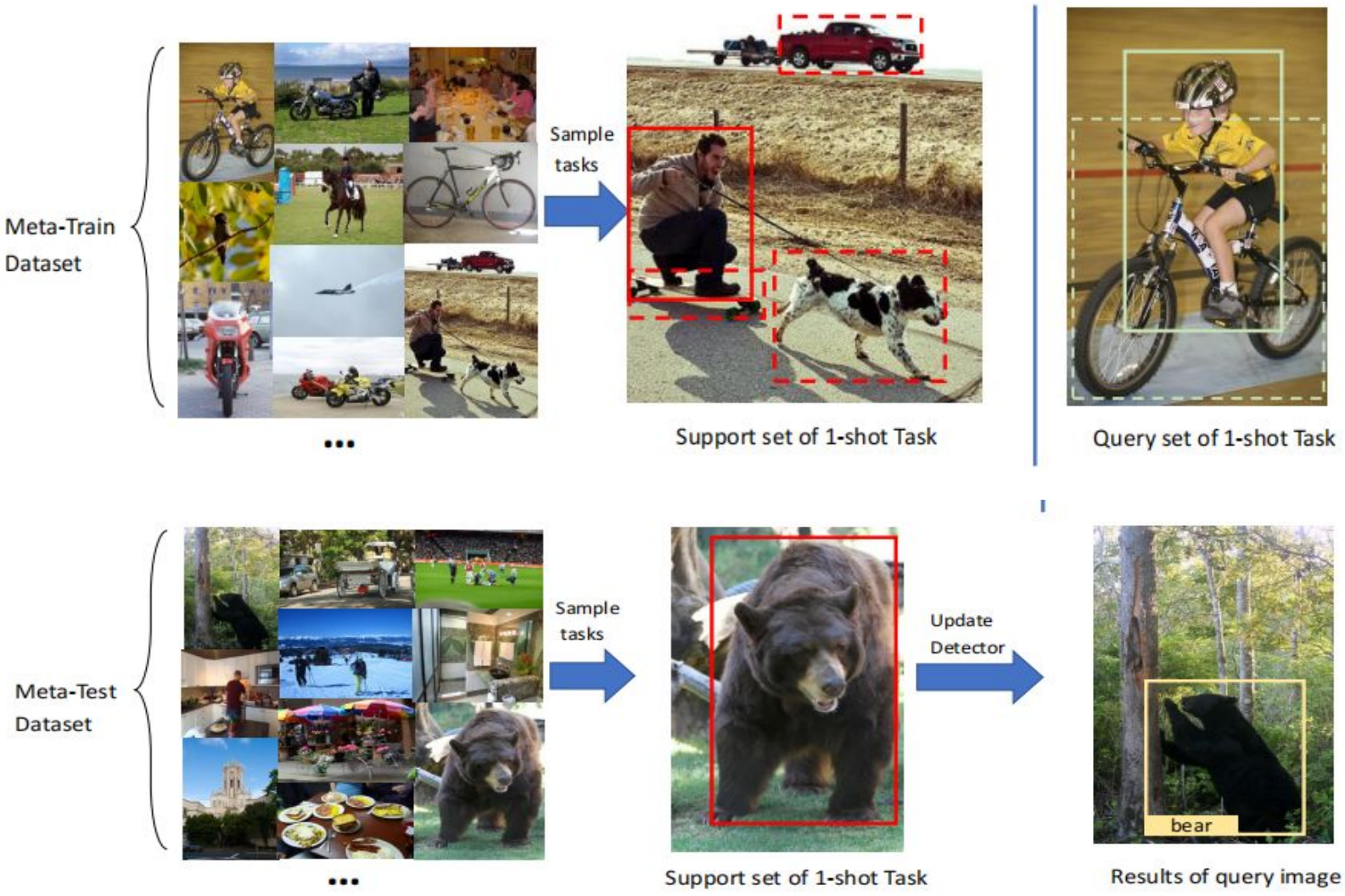
Meta-RCNN

In meta learning, a set of tasks in a few-shot setting is simulated from a large corpus of annotated data, and the model is optimized to perform well over these few shot tasks.

i) Detection algorithms not only require classifying objects but also need to correctly localize objects in cluttered backgrounds by using a Region Proposal Network (RPN) and bounding box (bbox) regressors.

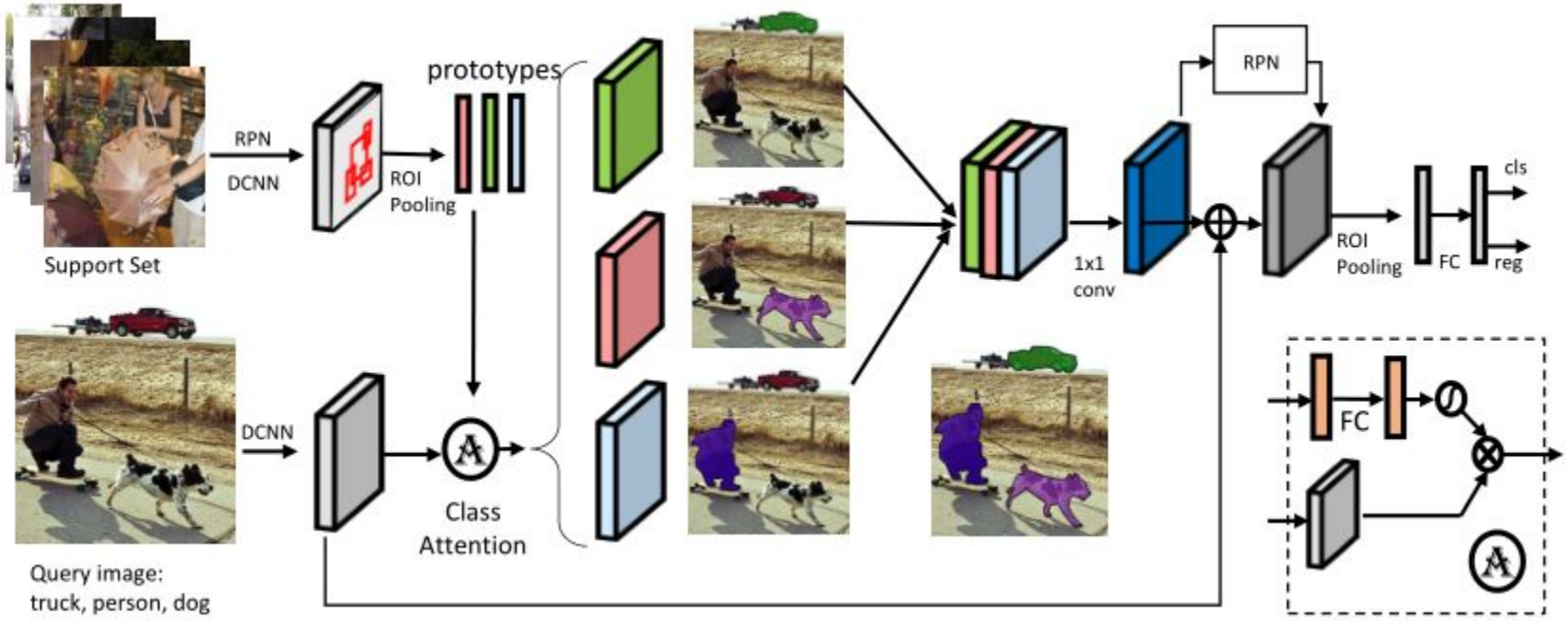
ii) For a given task with one (or few) annotated image(s), the annotated image may contain objects from several classes.

BACKGROUND



- BACKGROUND
- METHOD
- EXPERIMENTS
- CONCLUSION

METHOD



The Meta-RCNN workflow

- BACKGROUND
- METHOD
- **EXPERIMENTS**
- CONCLUSION

1. DATASETS AND IMPLEMENTATION DETAILS

DATASET	Train	#Img	#cls	Test	#Img	#cls
VOC-FSOD	VOC2007trainval	~ 4.9k	10	VOC2007test	~ 2.2k	10
IMAGENET-FSOD	ImageNet-LOC	~ 53k	100	ImageNet-LOC	~ 117k	214

Table 1: Two few-shot object detection benchmark testbeds for performance evaluation

Construct two benchmark testbeds to facilitate the performance evaluation for few-shot object detection in meta-learning settings.

The first is on Pascal VOC2007, and the second is on the animal subset of ImageNet-LOC dataset.

Experiments

2. RESULTS ON VOC-FSOD BENCHMARK

Method	5way-1shot	5way-3shot	5way-5shot
vanilla FRCN (Ren et al., 2015)	14.78% ± 1.02%	20.34% ± 1.26%	26.89% ± 1.23%
LSTD (Chen et al., 2018)	17.66% ± 1.65%	22.37% ± 0.81%	29.00% ± 1.28%
FRCN-PN	12.71% ± 0.70%	13.91% ± 0.70%	14.33% ± 0.61%
Meta-RCNN (ours)	**19.22% ± 1.01%**	**24.45% ± 1.20%**	**31.11% ± 0.88%**

Table 2: mAP Performance Evaluation on the VOC-FSOD BENCHMARK

vanilla FRCN : the vanilla Faster RCNN which is the most popular object detection algorithm with competitive performance on many benchmarks.

LSTD is a few-shot detection algorithm based on Faster RCNN. LSTD uses categorical regularization items which transfers knowledge of L dataset to S dataset.

FRCN-PN is a modified version of Faster RCNN using meta-learning, which replaces final FC classification layer with non-parametric prototypical network (PN), which is considered as a variant of RepMet.

Experiments

2. RESULTS ON VOC-FSOD BENCHMARK

Model	Backbone	5way-1shot	5way-3shot	5way-5shot
vanilla FRCN (Ren et al., 2015)	VGG16	24.9%	26.5%	28.4%
FRCN-PN	VGG16	24.7%	24.9%	26.1%
Meta-RCNN (ours)	VGG16	**26.1%**	**27.9%**	**33.7%**

Table 3: Recall evaluation of Meta-RCNN on VOC-FSOD BENCHMARK test set.

Performance of RPN: Here, we present the performance of RPN to validate our concerns of the negative impact of irrelevant categories.

We use regular FRCN and FRCN-PN as our baseline.

Experiments

3. RESULTS ON IMAGENET-FSOD BENCHMARK

Model	Backbone	50way-1shot	50way-5shot
vanilla FRCN (Ren et al., 2015)	VGG16	16.5%	34.3%
LSTD (Chen et al., 2018)	VGG16	19.2%	37.4%
RepMet (Schwartz et al., 2019)	DCN+FPN	24.1%	39.6%
Meta-RCNN (ours)	VGG16	24.6%	40.1%
Meta-RCNN (ours)	ResNet50	**25.1%**	**40.3%**

Table 4: mAP performance evaluation on IMAGENET-FSOD BENCHMARK.

On the **ImageNet-FSOD** benchmark, we adapt the weights of detector pretrained on **MSCOCO** trainval set, and then optimize Meta-RCNN based on this starting point.

- BACKGROUND
- METHOD
- EXPERIMENTS
- CONCLUSION

CONCLUSION

Meta learning based detection algorithm Meta RCNN, which is robust to few-shot learning, and the proposed training strategies make it more suit able in detection scenario.

Specifically it adapts the Faster RCNN method and enables meta-learning of the object classifier, the RPN and the bounding box regressor.

The RPN is meta-trained through a novel class-specific attention module.

THANKS

Q&A

Thank you!