
LSTD: A Low-Shot Transfer Detector for Object Detection (2018)

Hao Chen^{1,2} Yali Wang¹ Guoyou Wang² Yu Qiao^{1,3}

¹Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China

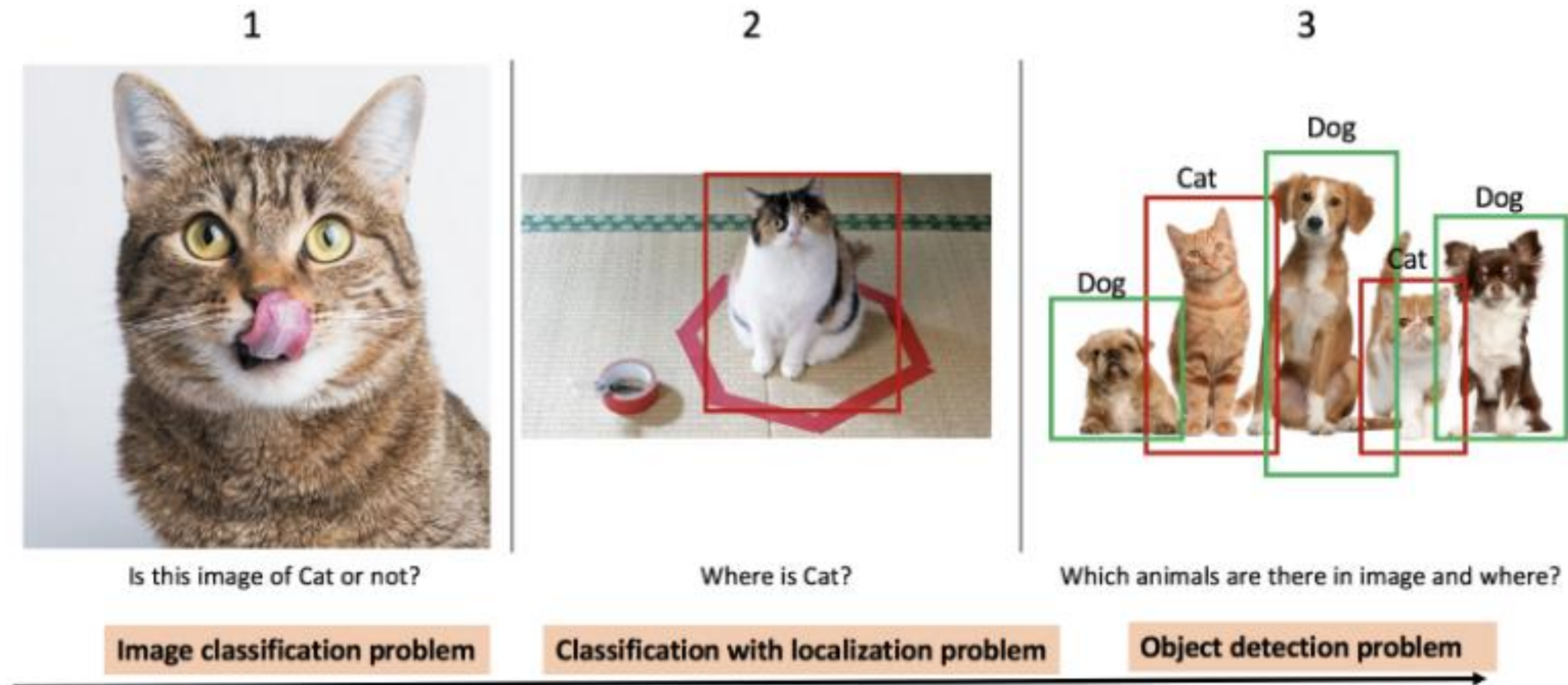
²Huazhong University of Science and Technology, China

³The Chinese University of Hong Kong, Hong Kong

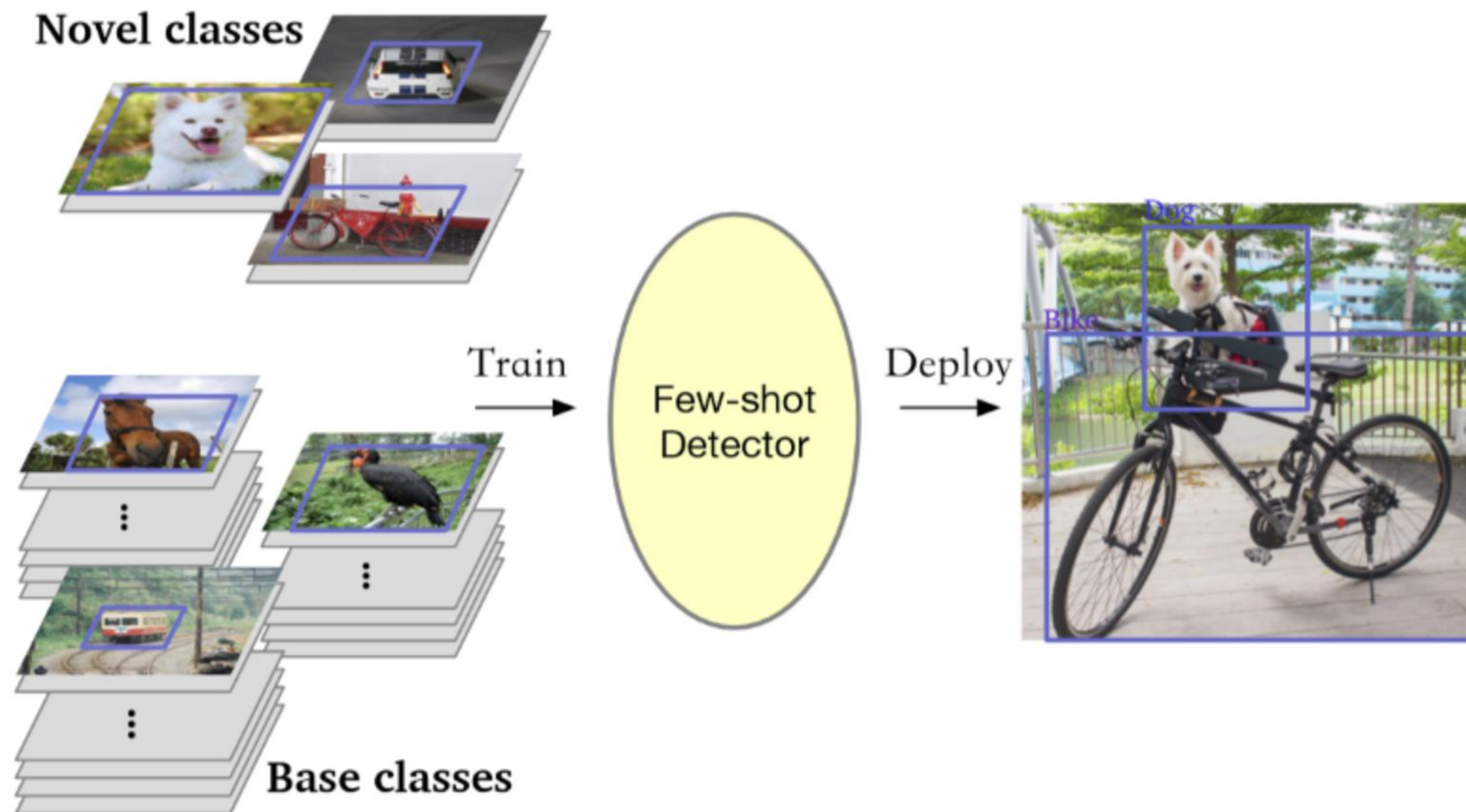
2023. 9. 20.

경영과학연구실 전재현

- Computer Vision Task



- Low-Shot(Few-Shot) Object detection



- Why Low-Shot Object Detection task is challenging?
 - 일반적으로 object detection task는 dataset에 의존적임
 - 하지만, 찾고자 하는 target object에 대해 annotated된 dataset이 부족한 경우가 있음
 - 적은 수의 data로 학습했을 때 overfitting이 일어날 가능성이 높음
- How to solve this problem?
 - Collecting extra detection images
 - Transfer learning(without extra data collection)

- Why Low-Shot Object Detection task is challenging?
 - 일반적으로 object detection task는 dataset에 의존적임
 - 하지만, 찾고자 하는 target object에 대해 annotated된 dataset이 부족한 경우가 있음
 - 적은 수의 data로 학습했을 때 overfitting이 일어날 가능성이 높음
- How to solve this problem?
 - Collecting extra detection images
 - **Transfer learning(without extra data collection)**

- **Problem Statement**

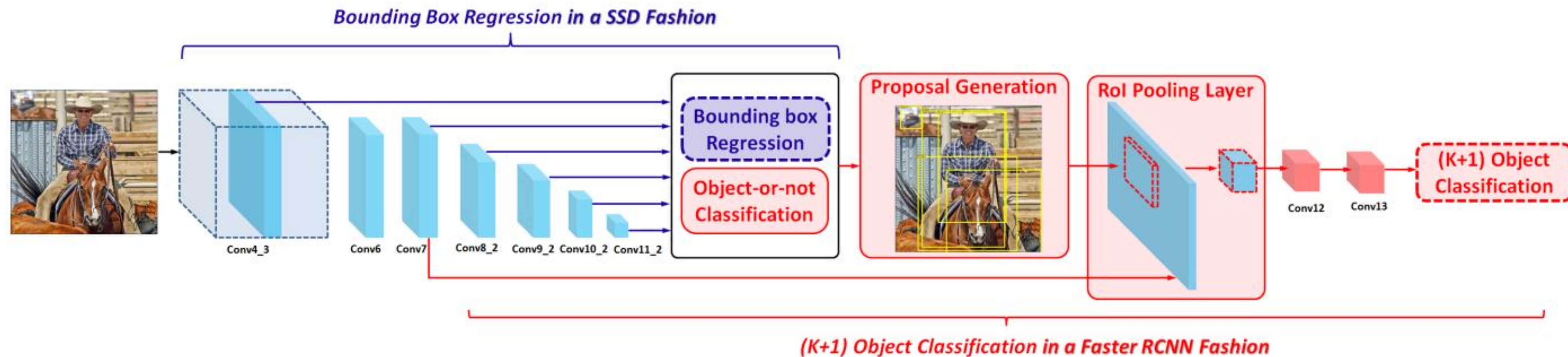
- Low-shot object detection task는 학습 data에 의존적이고, data 수가 적기 때문에 overfitting에 취약함

- **Key Idea**

- Object detection algorithm인 SSD와 Faster RCNN을 함께 사용하는 LSTD architecture 제안
- Low-shot object detection 분야에서 처음으로 transfer learning을 사용

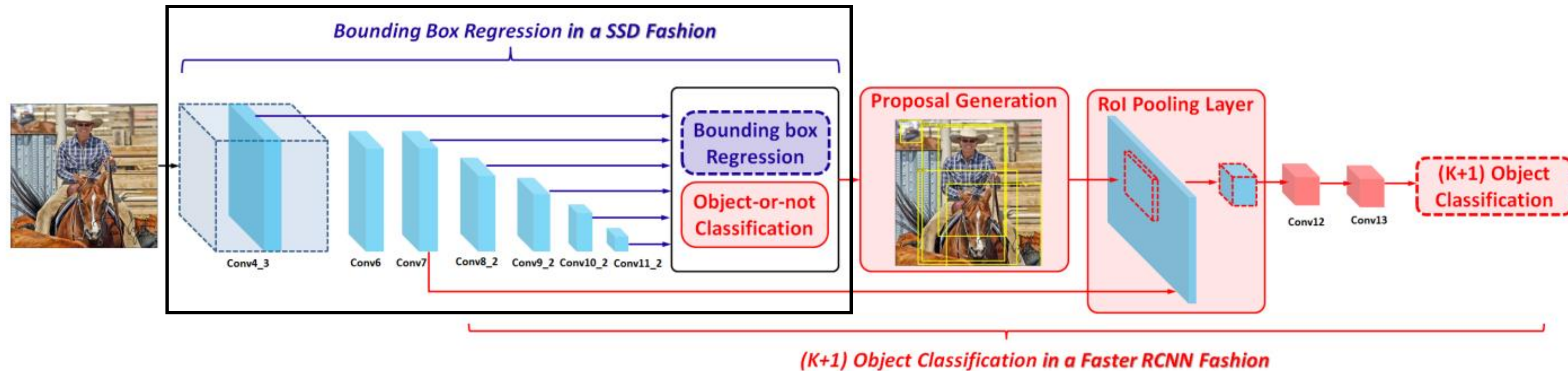
- Low-Shot Transfer Detector(LSTD)

- Target data set가 제한된 Low-Shot object detection을 효과적으로 하기 위한 아키텍처
- SSD와 Faster RCNN을 함께 사용



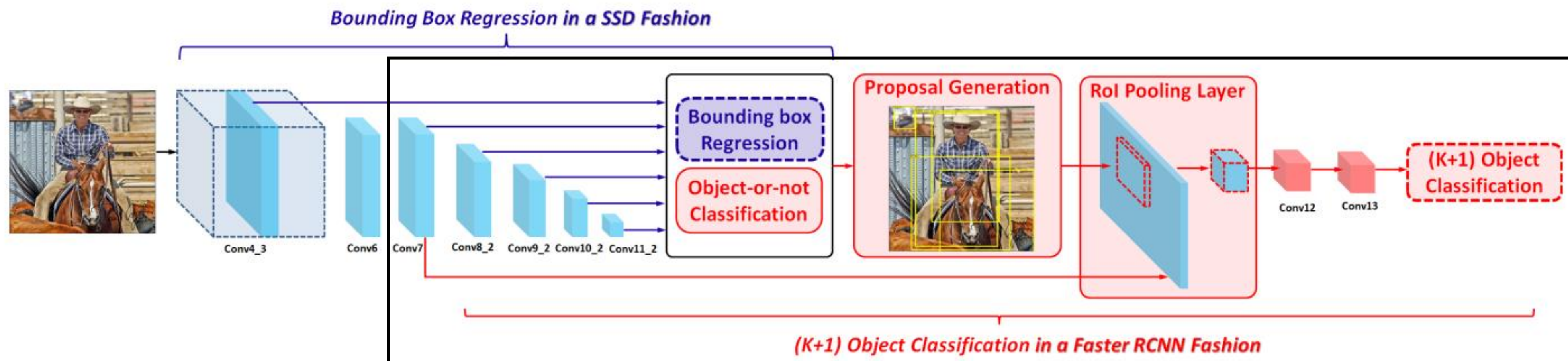
- Low-Shot Transfer Detector(LSTD) – Bounding Box Regression

- 여러 비율의 default bounding box 사용
- 다양한 크기의 객체가 존재할 때, localization에 적합한 방법
- 모든 객체의 카테고리간에 공유되기 때문에, data 수가 적은 low-shot object detection에서 size diversity에 robust할 수 있음



• Low-Shot Transfer Detector(LSTD) – (K+1) Object Classification

- 모든 default box에 대해 object-or-not으로 이진 분류를 진행
- Classification score를 활용하여 객체가 존재하는 proposal 생성
- RoI Pooling Layer를 통해 크기가 다른 bounding box들에 대한 (K+1) object classification 진행
- 객체와 배경의 이진분류를 통해 배경에 대한 학습이 가능해지고, 이는 적은 수의 target dataset으로도 배경과 객체를 분류할 수 있게 함



- Regularized Transfer Learning for LSTD

- 먼저, 대규모 source dataset을 이용하여 LSTD를 학습
- 그 후 pre-trained LSTD를 target domain에 맞게 finetuning
- Finetuning을 위해 BD(background depression)과 TK(transfer knowledge)를 도입

Algorithm 1 Regularized Transfer Learning of LSTD

1. Pre-training on Large-scale Source Domain

- Source-domain LSTD is trained with a large-scale detection benchmark.

2. Initialization for Small-scale Target Domain

- The last layer of $(K + 1)$ -object classifier is randomly initialized, due to object difference in source and target.
- All other parts of target-domain LSTD are initialized from source-domain LSTD.

3. Fine-tuning for Small-scale Target Domain

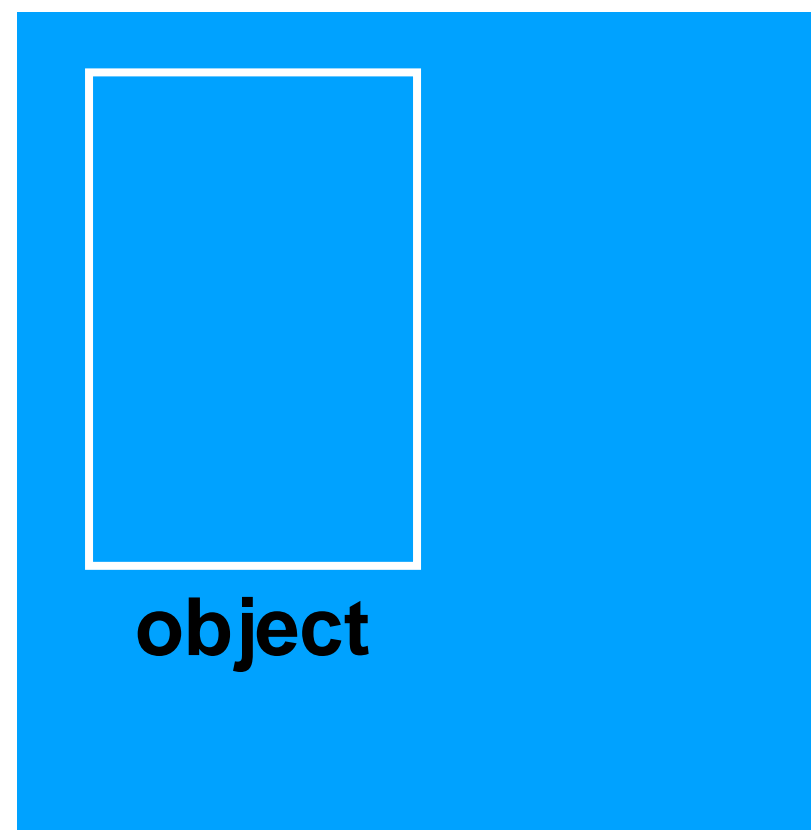
- We fine-tune target-domain LSTD with BD and TK regularizations (Eq. 1-6), based on the small training set.
-

$$\mathcal{L}_{total} = \mathcal{L}_{main} + \mathcal{L}_{reg}$$
$$\mathcal{L}_{reg} = \lambda_{BD}\mathcal{L}_{BD} + \lambda_{TK}\mathcal{L}_{TK}$$

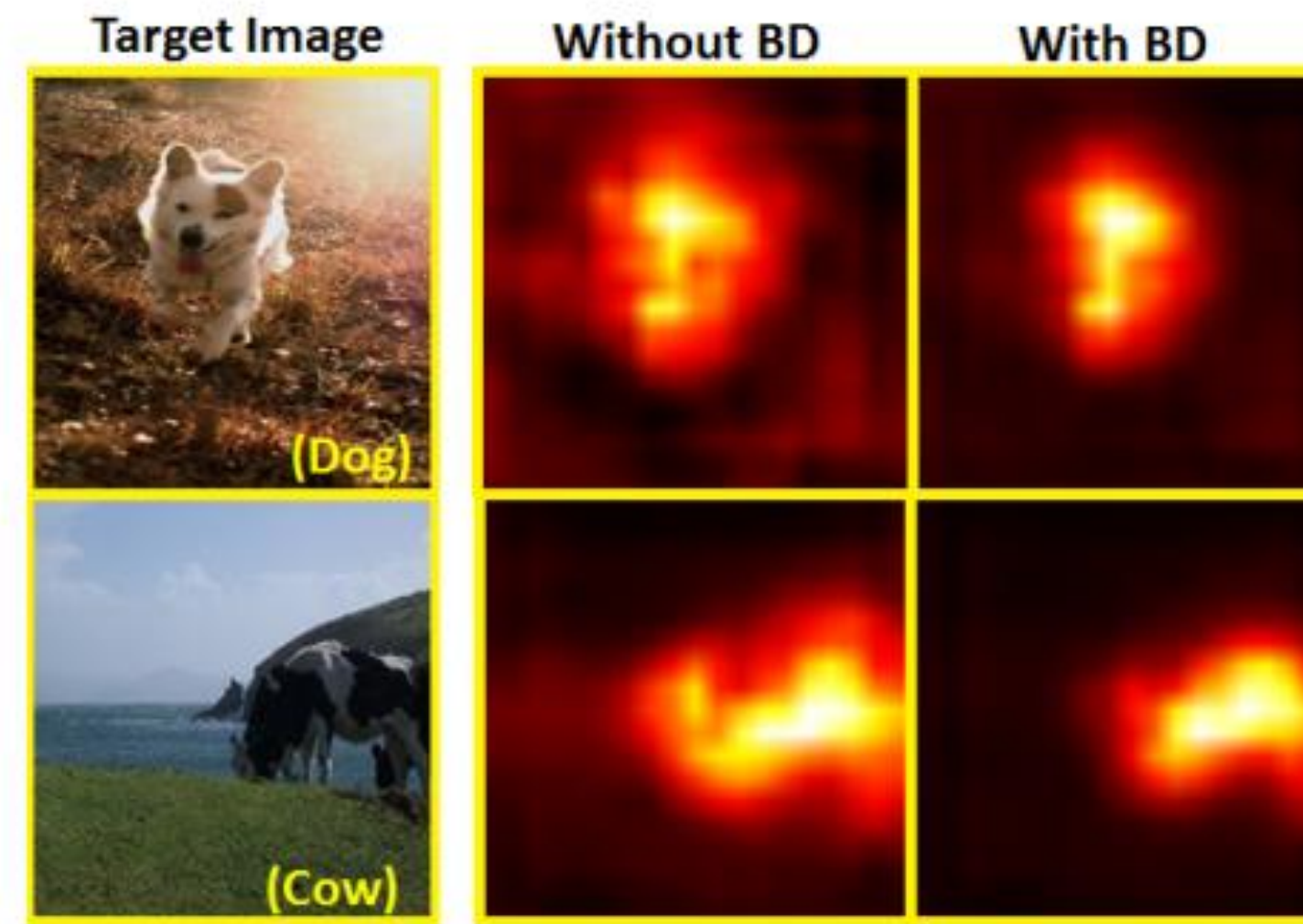
- Background-Depression(BD)

- Low-shot object detection에서 background가 영향을 줄 수 있음
- Target domain에 대해 학습을 진행할 때 객체가 존재하는 영역에 대해 masking 처리를 하여 새로운 dataset에서의 background에 대해 추가적으로 학습
- background가 object 영역에서 활성화되는 것을 방지하기 위해 \mathcal{L}_{BD} 를 도입

$$\mathcal{L}_{BD} = \|\mathbf{F}_{BD}\|_2$$



Target Image



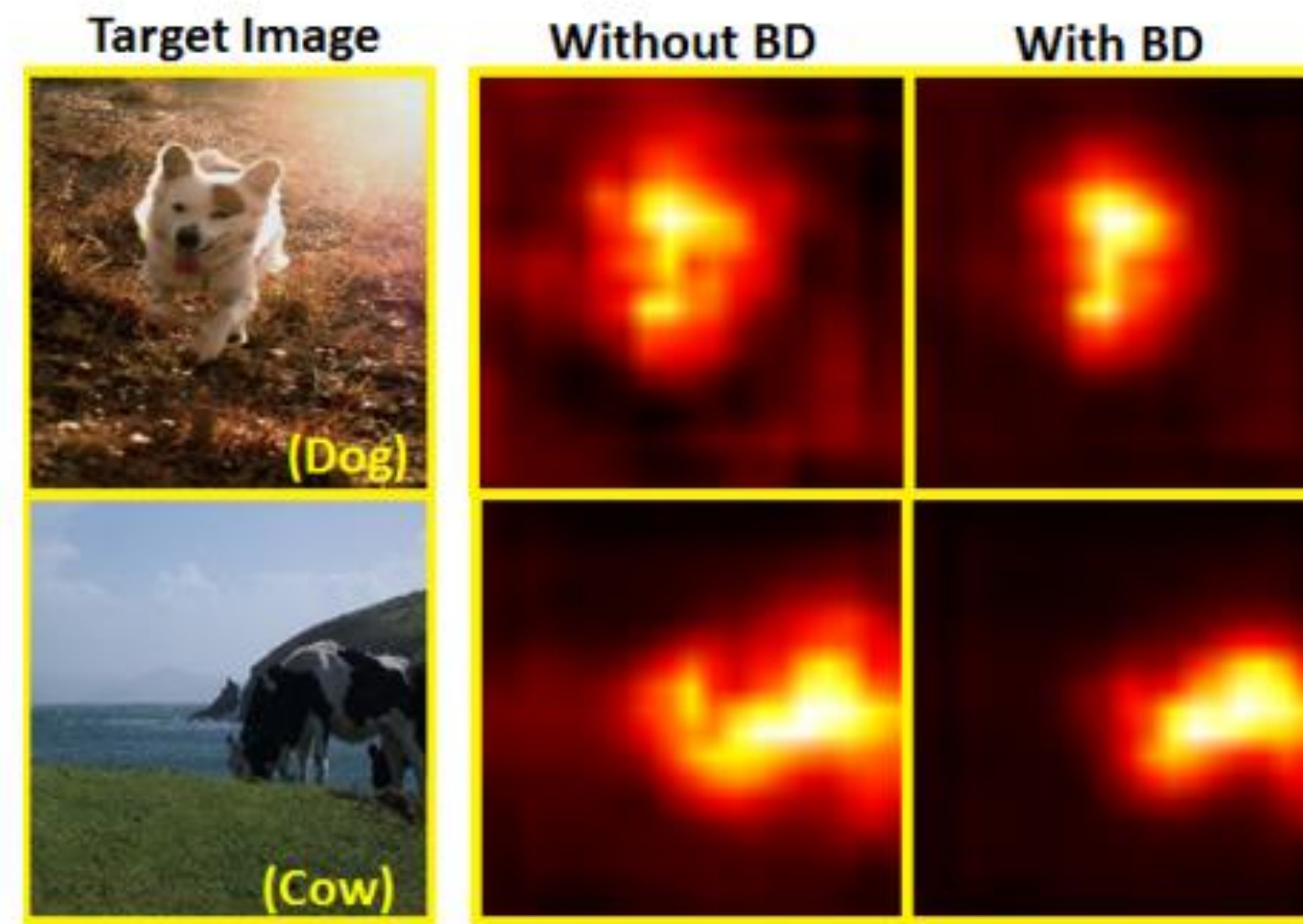
- Background-Depression(BD)

- Low-shot object detection에서 background가 영향을 줄 수 있음
- Target domain에 대해 학습을 진행할 때 객체가 존재하는 영역에 대해 masking 처리를 하여 새로운 dataset에서의 background에 대해 추가적으로 학습
- background가 object 영역에서 활성화되는 것을 방지하기 위해 \mathcal{L}_{BD} 를 도입

$$\mathcal{L}_{BD} = \|\mathbf{F}_{BD}\|_2$$



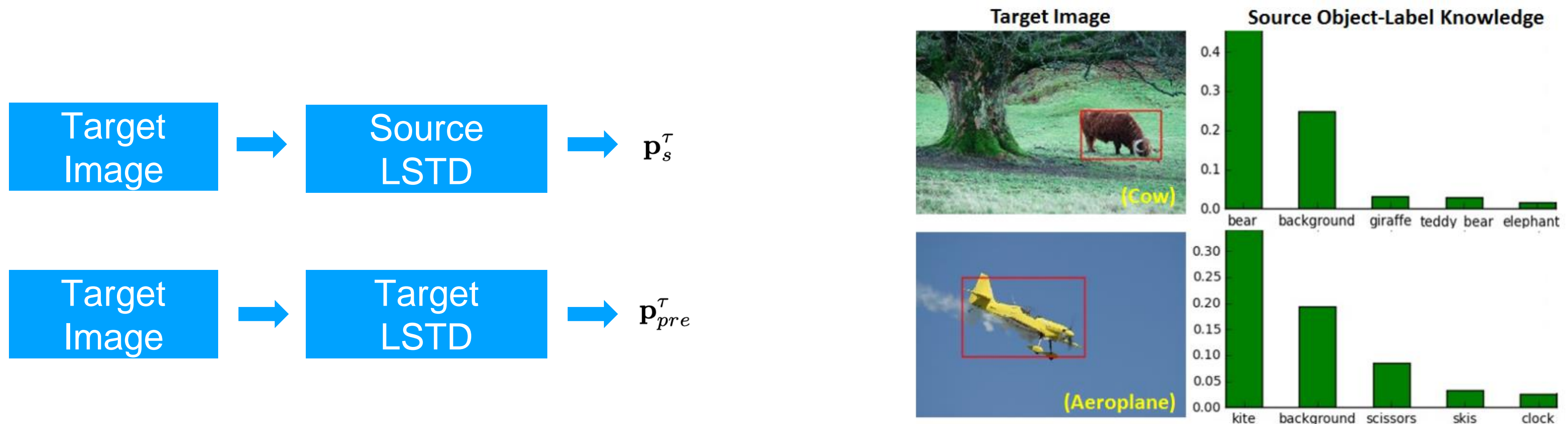
Target Image



- Transfer-Knowledge(TK)

- \mathbf{p}_s^τ : source domain에서 학습된 지식을 기반으로 한 target object에 대한 soften softmax 값
- \mathbf{p}_{pre}^τ : target domain에서 학습된 지식을 기반으로 한 target object에 대한 soften softmax 값
- 두 확률분포사이의 cross entropy loss를 최소화하여, target domain 네트워크가 source domain의 지식을 참고하도록 유도함

$$\mathcal{L}_{TK} = CrossEntropy(\mathbf{p}_s^\tau, \mathbf{p}_{pre}^\tau)$$



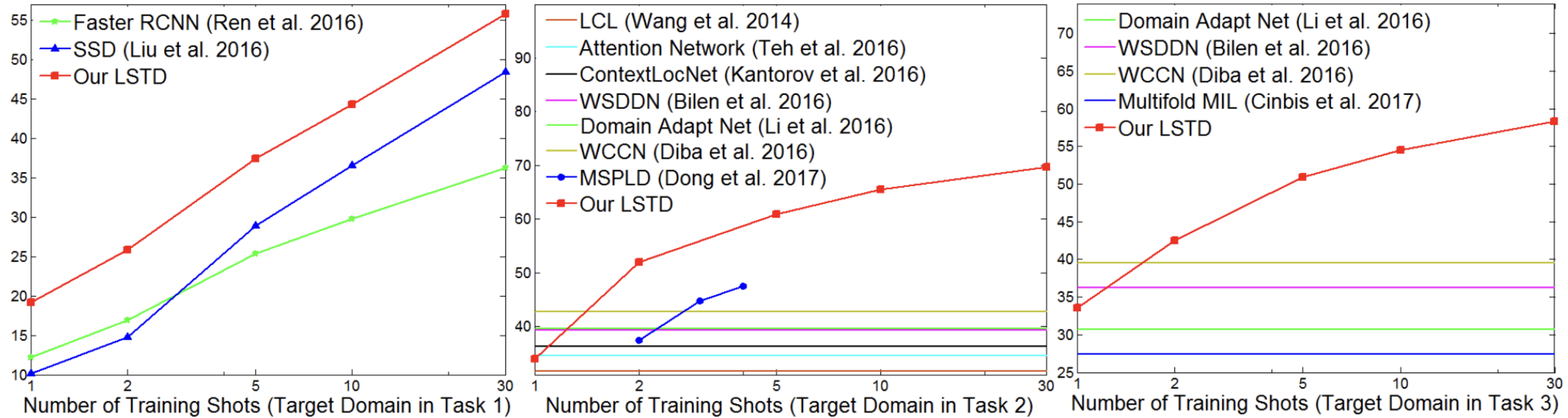
Experiment Results

- Faster RCNN vs. SSD vs. LSTD
 - Large Source dataset에 대해서는 큰 효과가 없음
 - 하지만 data가 적은 경우에 효과적인 구조임을 확인할 수 있음

| Deep Models | mAP | |
|-----------------------------|--------------|-----------------|
| | Large Source | Low-shot Target |
| Faster RCNN | 21.9 | 12.2 |
| SSD | 25.1 | 10.1 |
| Our LSTD _{conv5_3} | 24.7 | 15.9 |
| Our LSTD _{conv7} | 25.2 | 16.5 |

Experiment Results

- Comparison with the State-of-the-art



- Conclusion
 - Low-shot object detection을 효과적으로 수행하기 위해 transfer learning 기법을 해당 task에 처음으로 도입
 - Transfer difficulty를 줄이기 위해 LSTD라는 새로운 architecture를 제안함
 - 하지만 해당 task에서는 현재 meta-learning 기법이 많이 사용되고, transfer learning 기법을 사용하는 연구가 거의 없다는 점에서 transfer learning의 한계가 있는 것 같다는 생각이 듭