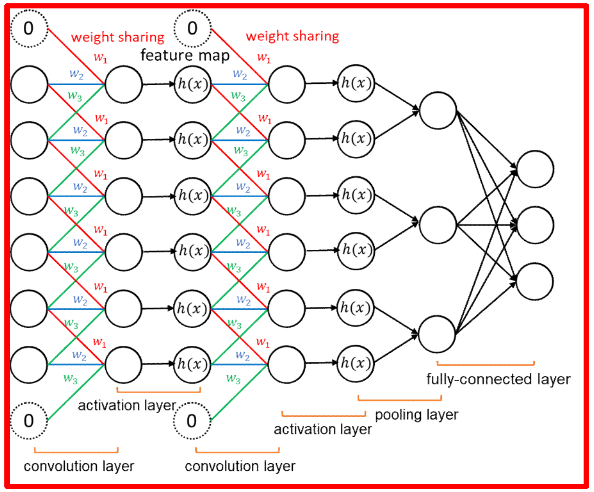
IIT-H and RIKEN-AIP Joint Workshop on Machine Learning and Applications March 15, 2019

Knowledge Transfer for Visual Recognition

The University of Tokyo
RIKEN AIP (Team leader of Medical Machine Intelligence)
Tatsuya Harada

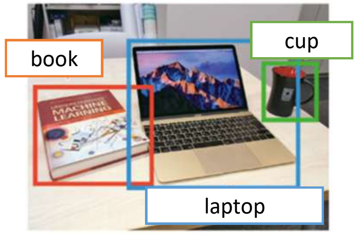
Deep Neural Networks for Visual Recognition

Deep Neural Networks

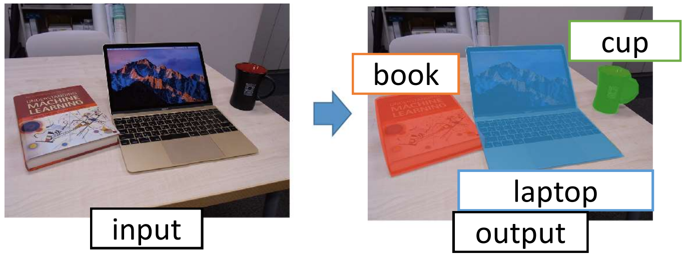


Applications

cellphone

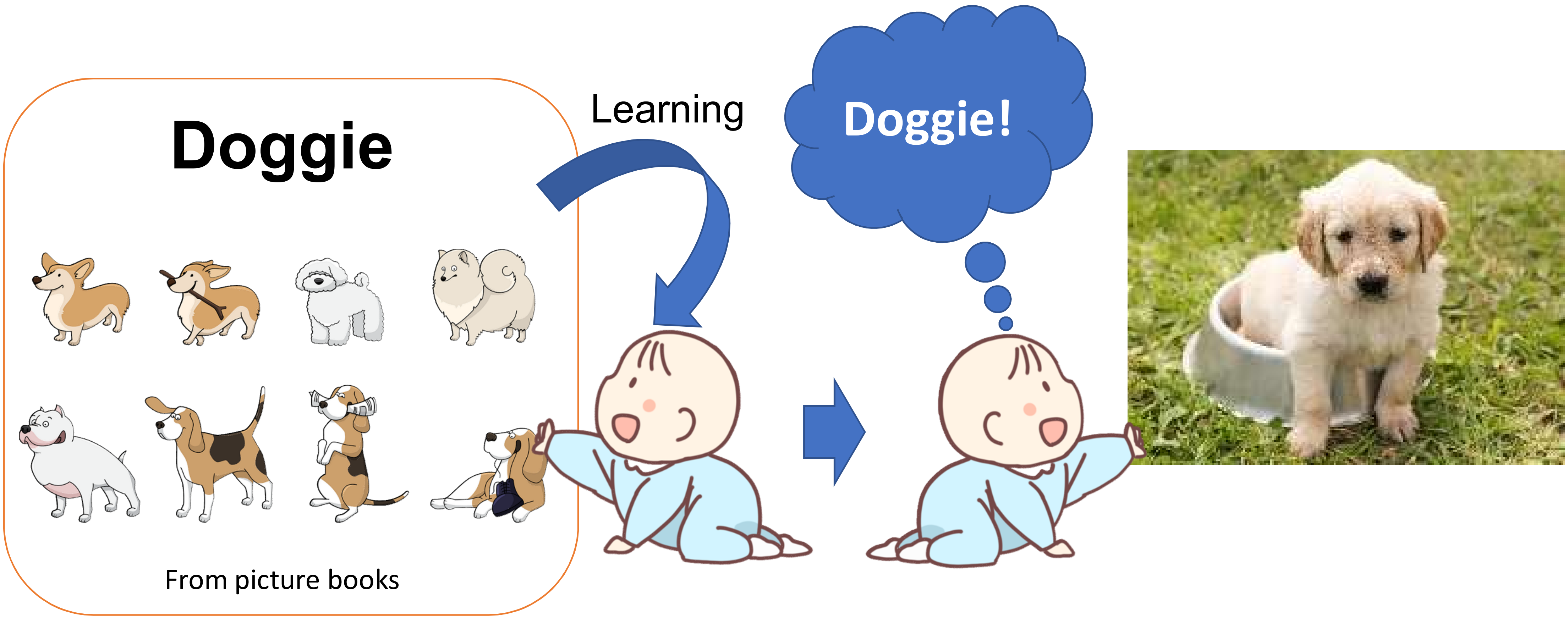


A yellow train on the tracks near a train station.



- Tasks in the visual recognition field
  - Object class recognition
  - Object detection
  - Image caption generation
  - Semantic and instance segmentation
  - Image generation
  - Style transfer
- DNNs becomes an indispensable module.
- A large amount of labeled data is needed to train DNNs.
- Reducing annotation cost is highly required.

Knowledge Transfer



Domain Adaptation

Image by Chiemsee2016 on Pixabay

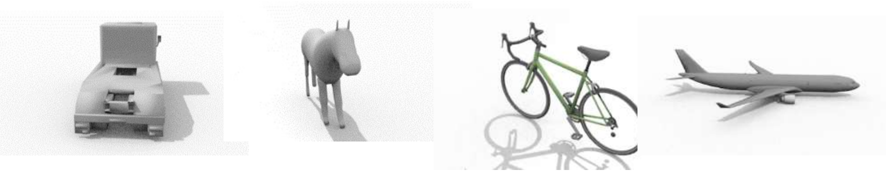
Image by GraphicMama-team on Pixabay

Domain Adaptation (DA)

- Problems
  - Supervised learning model needs many labeled examples
  - Cost to collect them in various domains
- Goal
  - Transfer knowledge from source (rich supervised data) to target (small supervised data) domain
  - Classifier that works well on target domain.
- Unsupervised Domain Adaptation (UDA)
  - Labeled examples are given only in the source domain.
  - There are no labeled examples in the target domain.

Source domain

Synthetic images, labeled



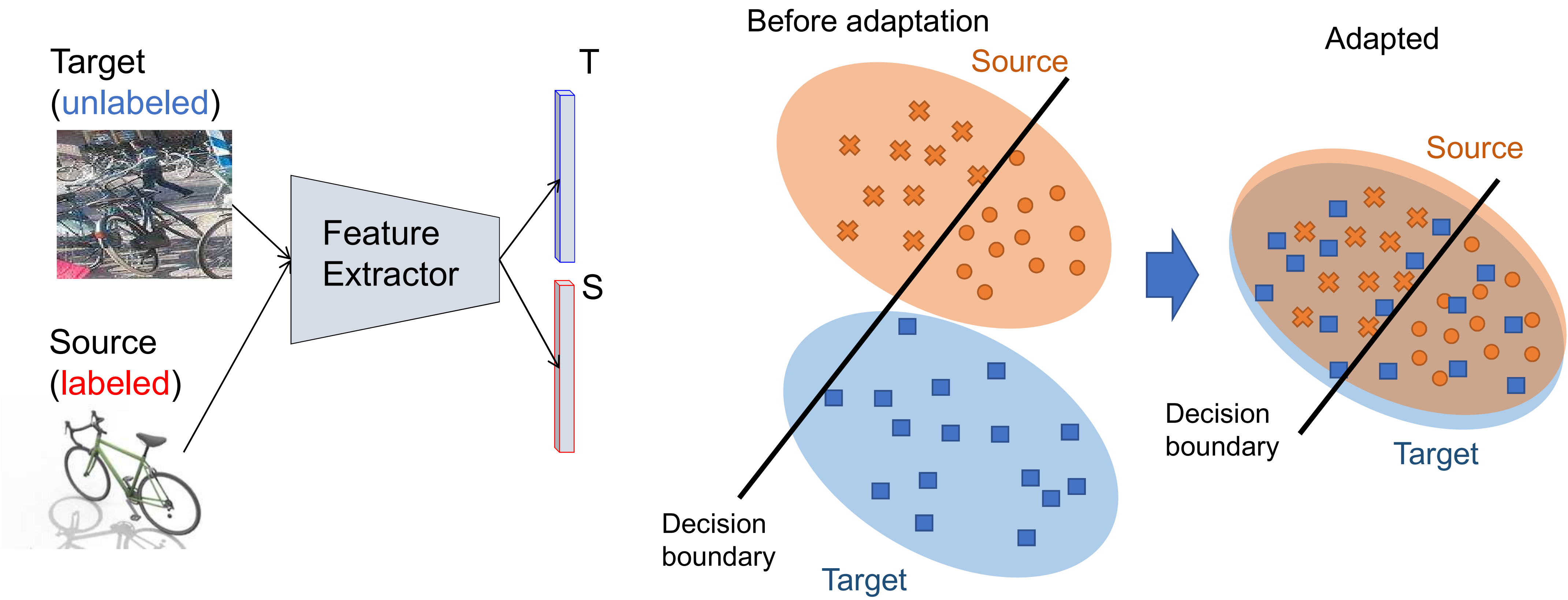
Target domain

Real images, unlabeled

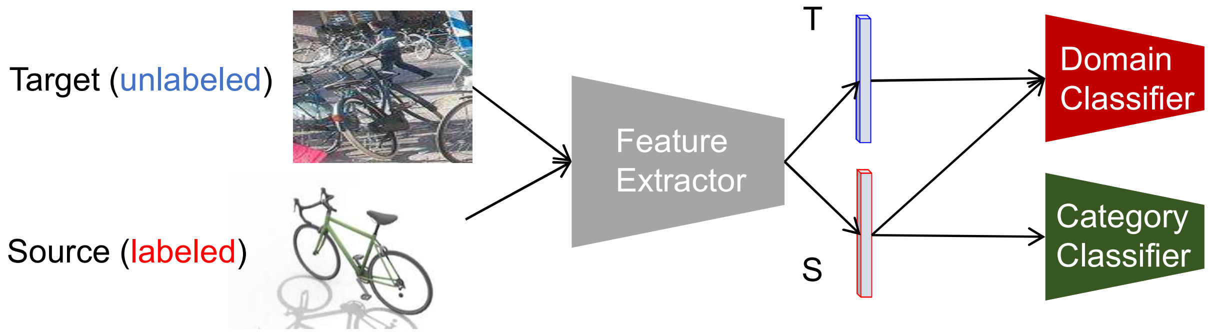
Distribution Matching for Unsupervised Domain Adaptation

◻Distribution matching based method

- Match distributions of source and target features
  - Domain Classifier (GAN) [Ganin et al., 2015]
  - Maximum Mean Discrepancy [Long et al., 2015]

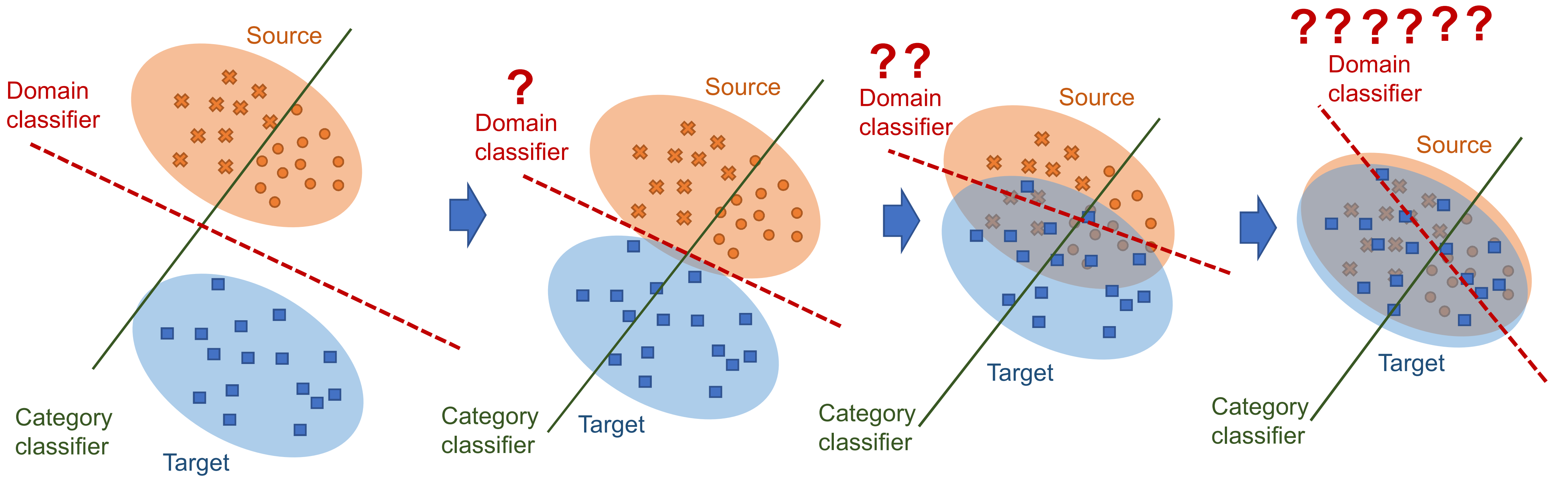


Adversarial Domain Adaptation



- Training the feature generator in a adversarial way works well!
- Category classifier, domain classifier, feature extractor
- Problems
  - Whole distribution matching
  - Ignorance of category information in source domain

Tzeng, Eric, et al. Adversarial discriminative domain adaptation. CVPR, 2017.



Unsupervised Domain Adaptation using Classifier Discrepancy

Kuniaki Saito[1], Kohei Watanabe[1], Yoshitaka Ushiku[1], Tatsuya Harada[1, 2]
1: The University of Tokyo, 2: RIKEN

CVPR 2018, oral presentation


K. Saito

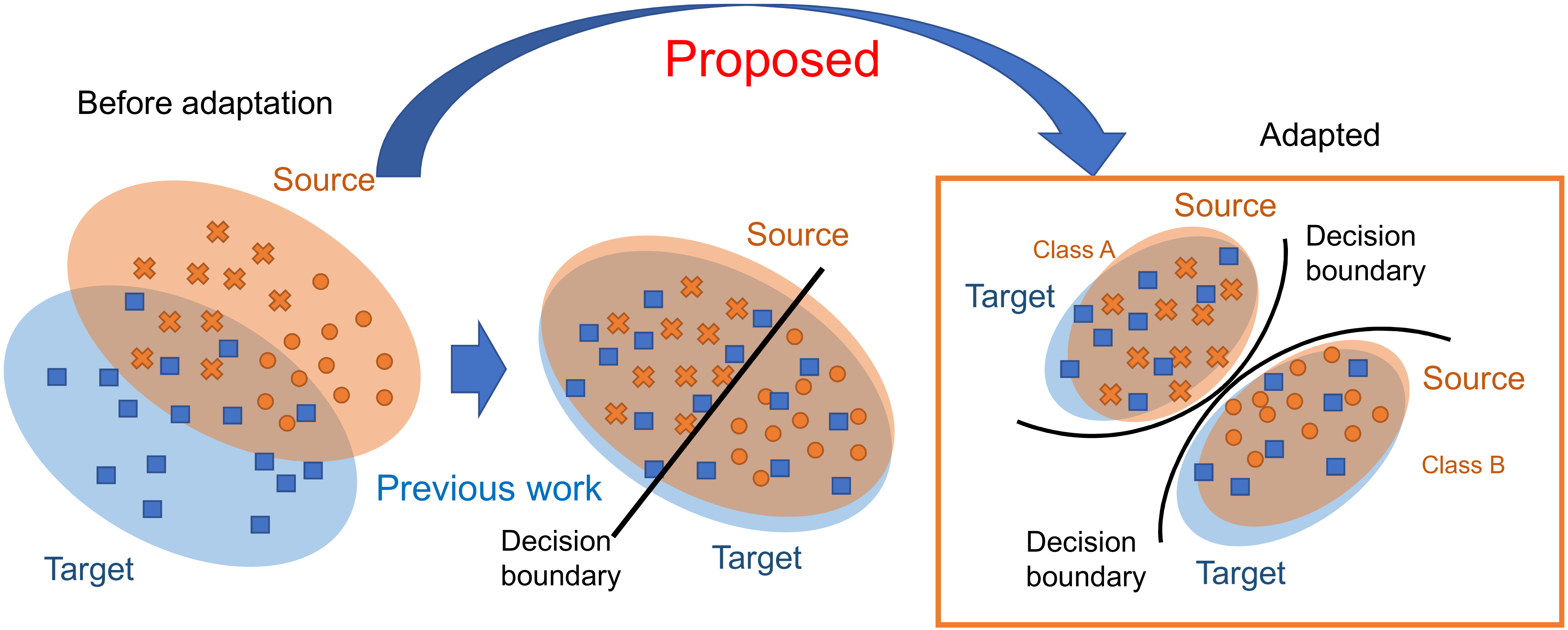

K. Watanabe


Y. Ushiku


T. Harada

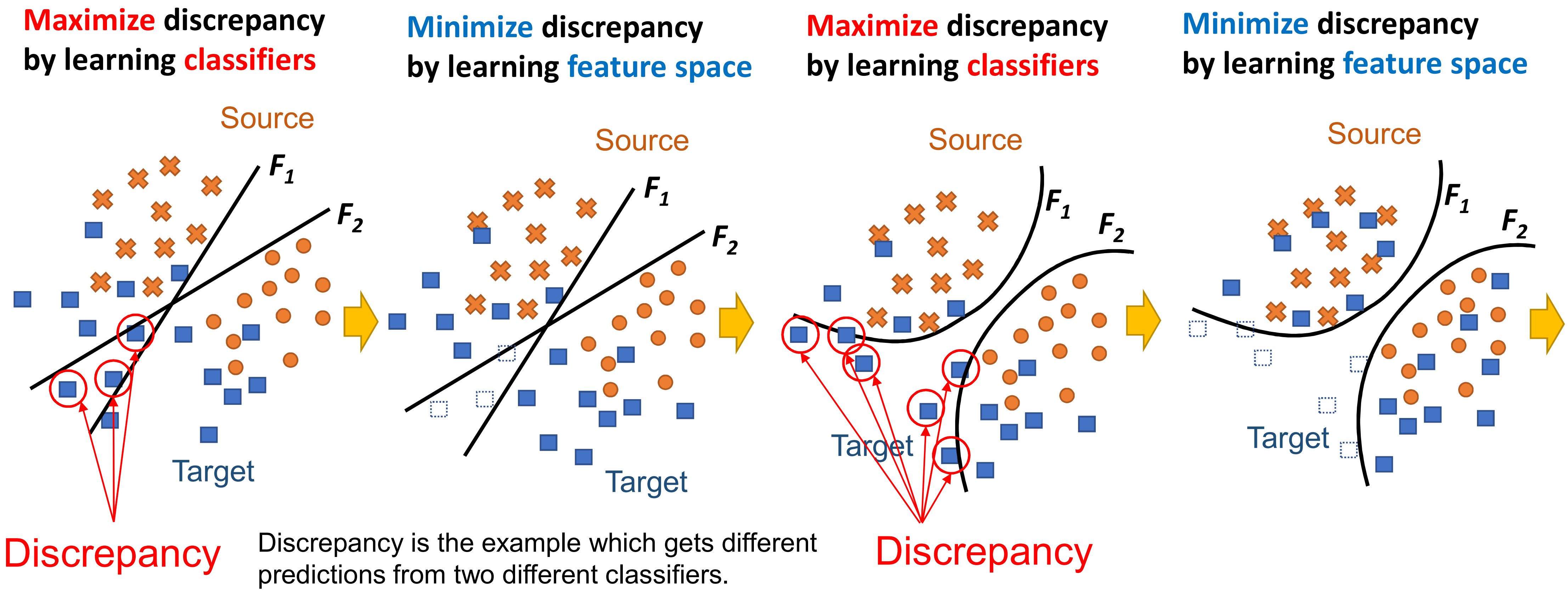
Proposed Approach

- Considering class specific distributions
- Using decision boundary to align distributions

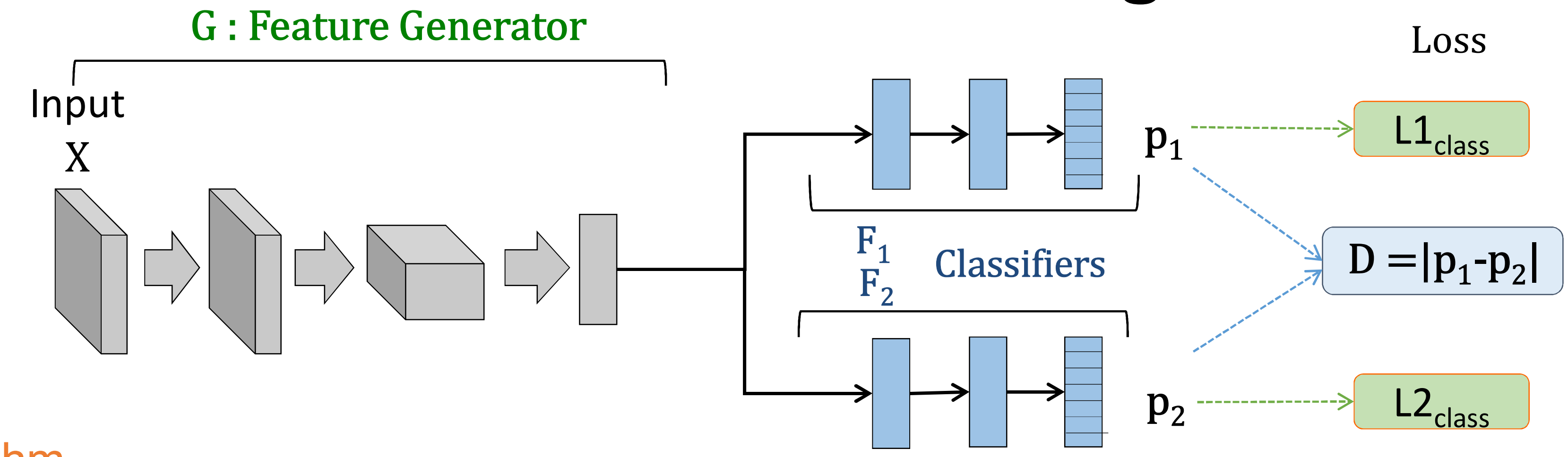


Key Idea

- ☐Maximizing discrepancy by learning two classifiers
- ☐Minimizing discrepancy by learning feature space



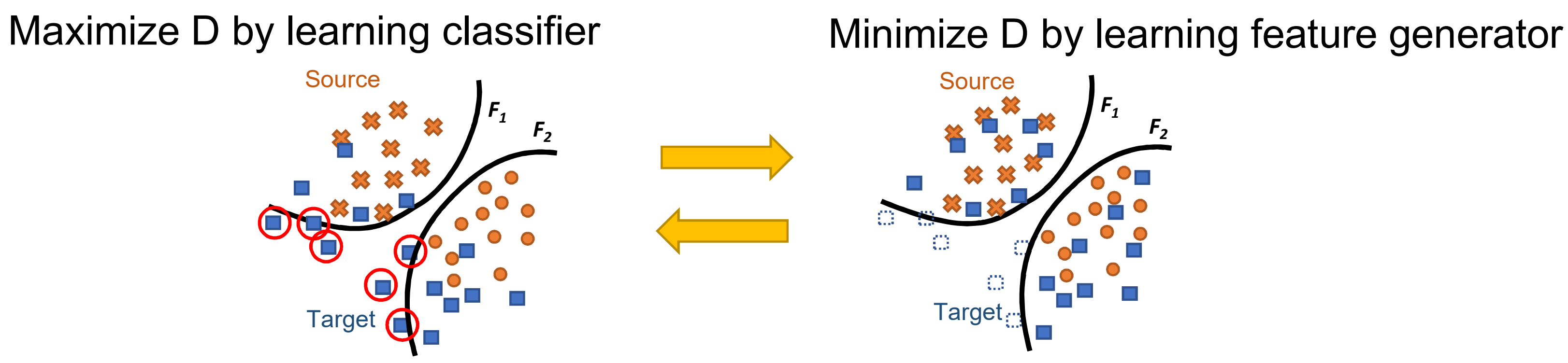
Network Architecture and Training



Algorithm

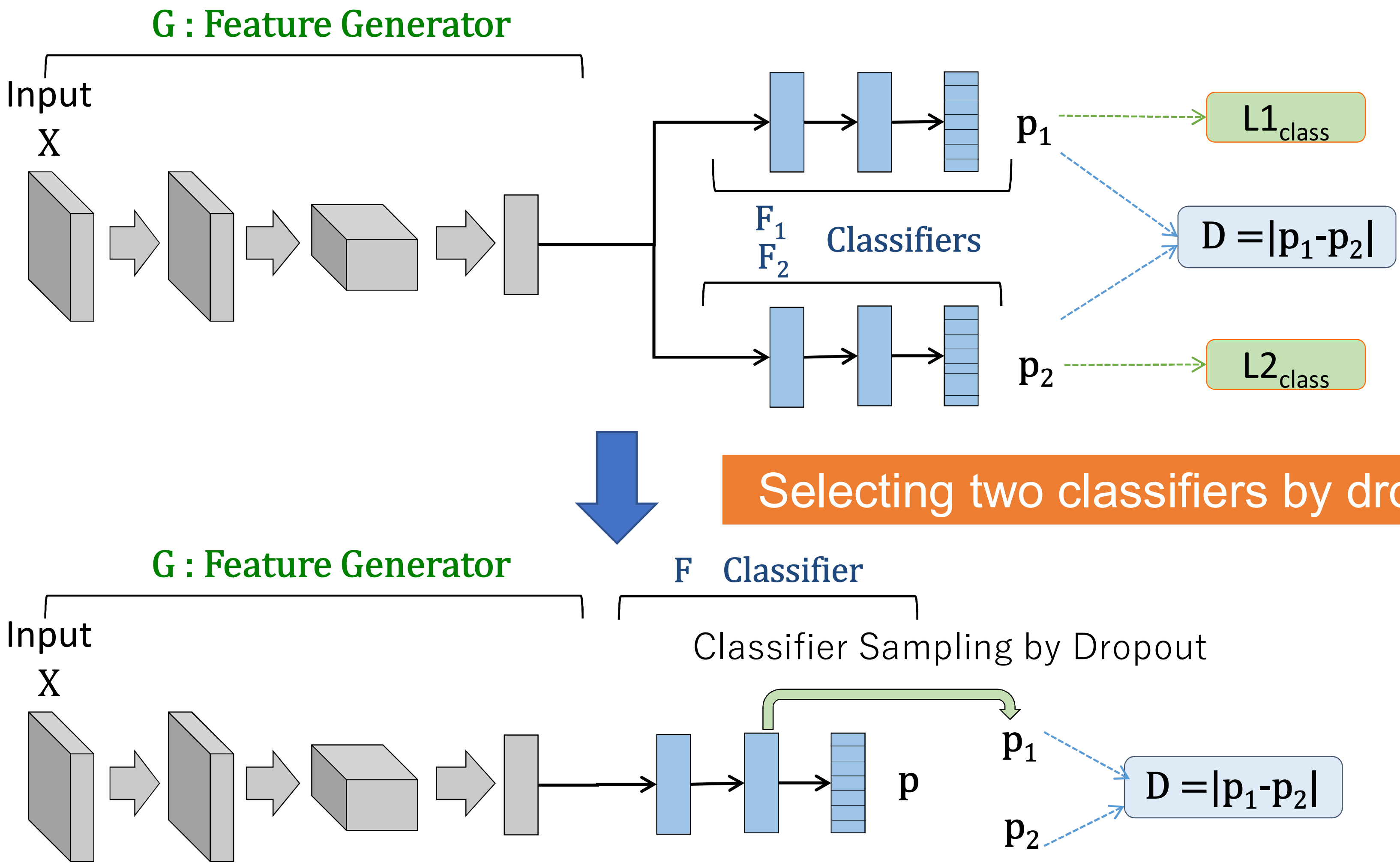
1. Fix generator G, and find classifiers F_1, F_2 that **maximize $D - (L_1 + L_2)$**
2. for $k = 1 : n$

   Fix classifiers F_1, F_2, and find feature generator G that **minimizes D**



Improving by Dropout

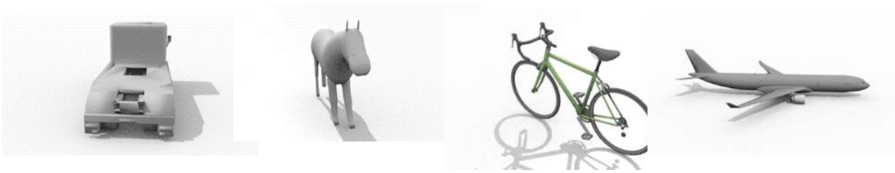
Adversarial Dropout Regularization
Kuniaki Saito, Yoshitaka Ushiku, Tatsuya Harada, Kate Saenko
ICLR 2018



Object Classification

- Synthetic images to Real images (12 Classes)
- Finetune pre-trained ResNet101 [He et al., CVPR 2016] (ImageNet)
- Source:images, Target:images

Source (Synthetic images)



Target (Real images)

Method	plane	bcycl	bus	car	hrs	knf	mcycl	prsn	plnt	sktbrd	trn	trck	mean
Source Only	55.1	53.3	61.9	59.1	80.6	17.9	79.7	31.2	81.0	26.5	73.5	8.5	52.4
MMD [Long et al., ICML 2015]	87.1	63.0	76.5	42.0	**90.3**	42.9	**85.9**	53.1	49.7	36.3	**85.8**	20.7	61.1
DANN [Ganin et al., ICML 2015]	81.9	**77.7**	82.8	44.3	81.2	29.5	65.1	28.6	51.9	**54.6**	82.8	7.8	57.4
Ours ($n = 4$)	**87.0**	60.9	**83.7**	**64.0**	88.9	**79.6**	84.7	**76.9**	**88.6**	40.3	83.0	**25.8**	**71.9**

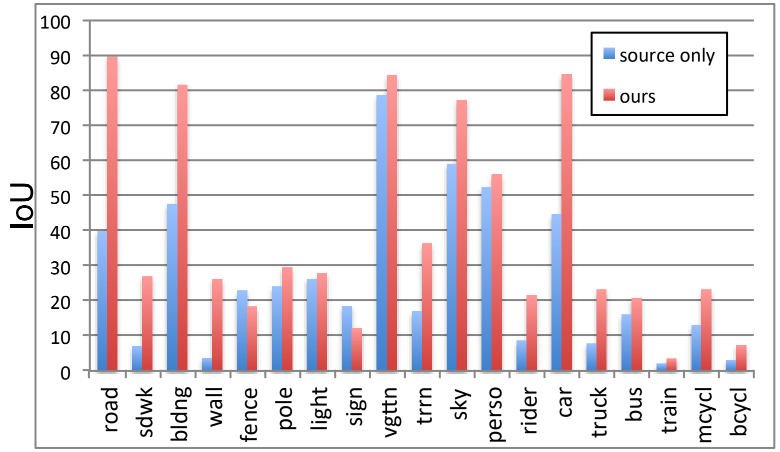
Semantic Segmentation

- Simulated Image (GTA5) to Real Image (CityScape)
- Finetuning of pre-trained VGG, Dilated Residual Network [Yu et al., 2017] (ImageNet)
  - Calculate discrepancy pixel-wise
- Evaluation by mean IoU (TP/(TP+FP+FN))

GTA 5 (Source)

CityScape(Target)

Network	Method	mIoU
VGG-16	Source Only	21.2
	FCN Wld [Hoffman et al., Arxiv 2017]	27.1
VGG-16	Source Only	22.3
	CrrclmDA (I) [Zhang el al., ICCV 2017]	23.1
VGG-16	Source Only	24.9
	Ours	28.8
DRN-105	Source Only	22.2
	Ours	**39.7**



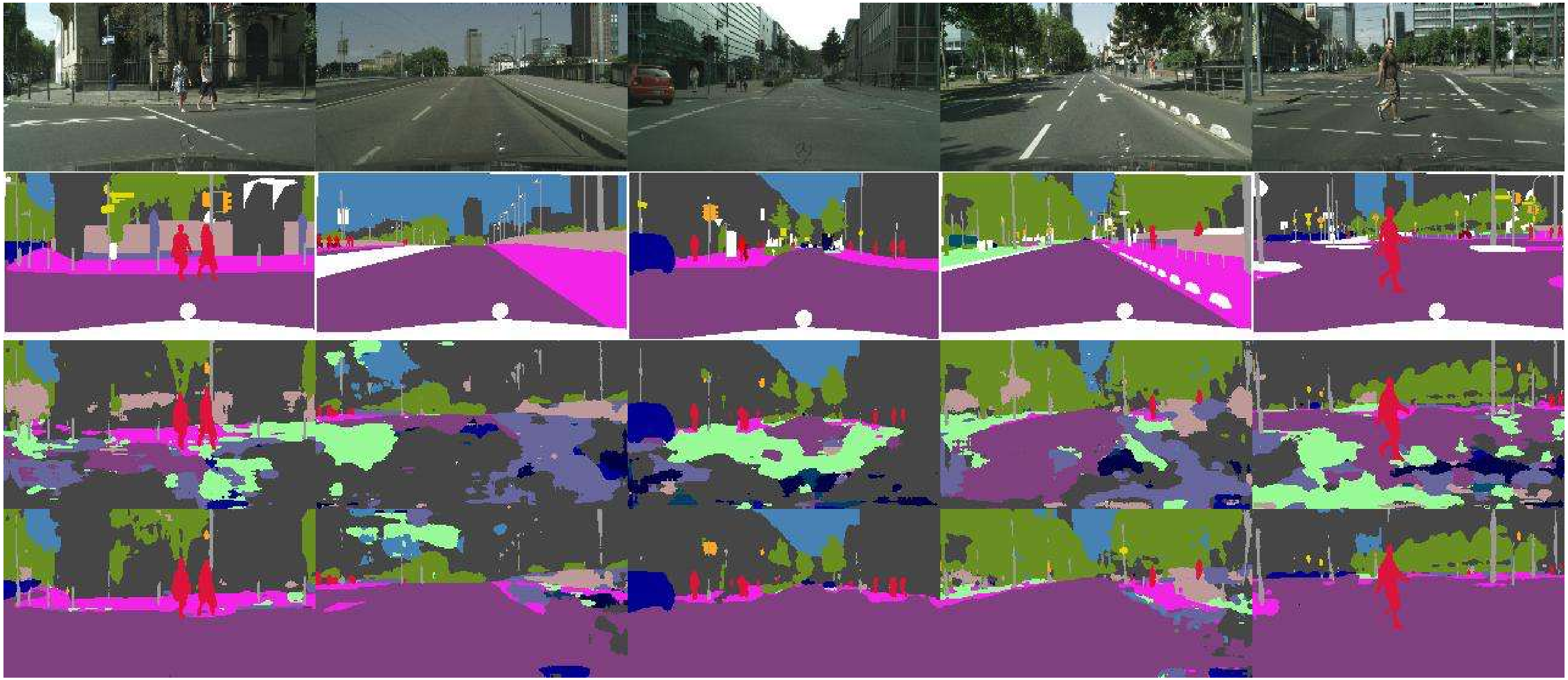
Qualitative Results

RGB

Ground truth

Source only

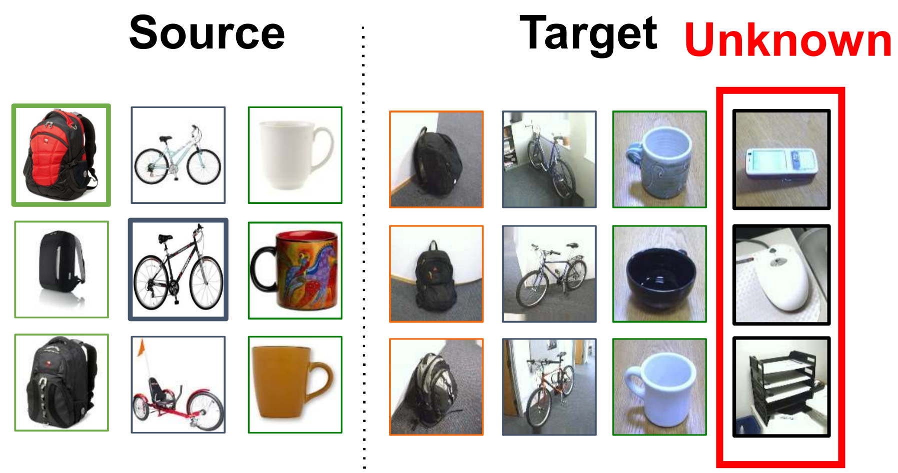
Adapted (ours)



Another Topics of Unsupervised Domain Adaptation

□Open-set Domain Adaptation

- Kuniaki Saito, Shohei Yamamoto, Yoshitaka Ushiku, Tatsuya Harada. Open Set Domain Adaptation by Backpropagation. ECCV, 2018.



□Adaptive Object Detection

- Kuniaki Saito, Yoshitaka Ushiku, Tatsuya Harada, Kate Sanenko. Strong-Weak Distribution Alignment for Adaptive Object Detection. CVPR, 2019.