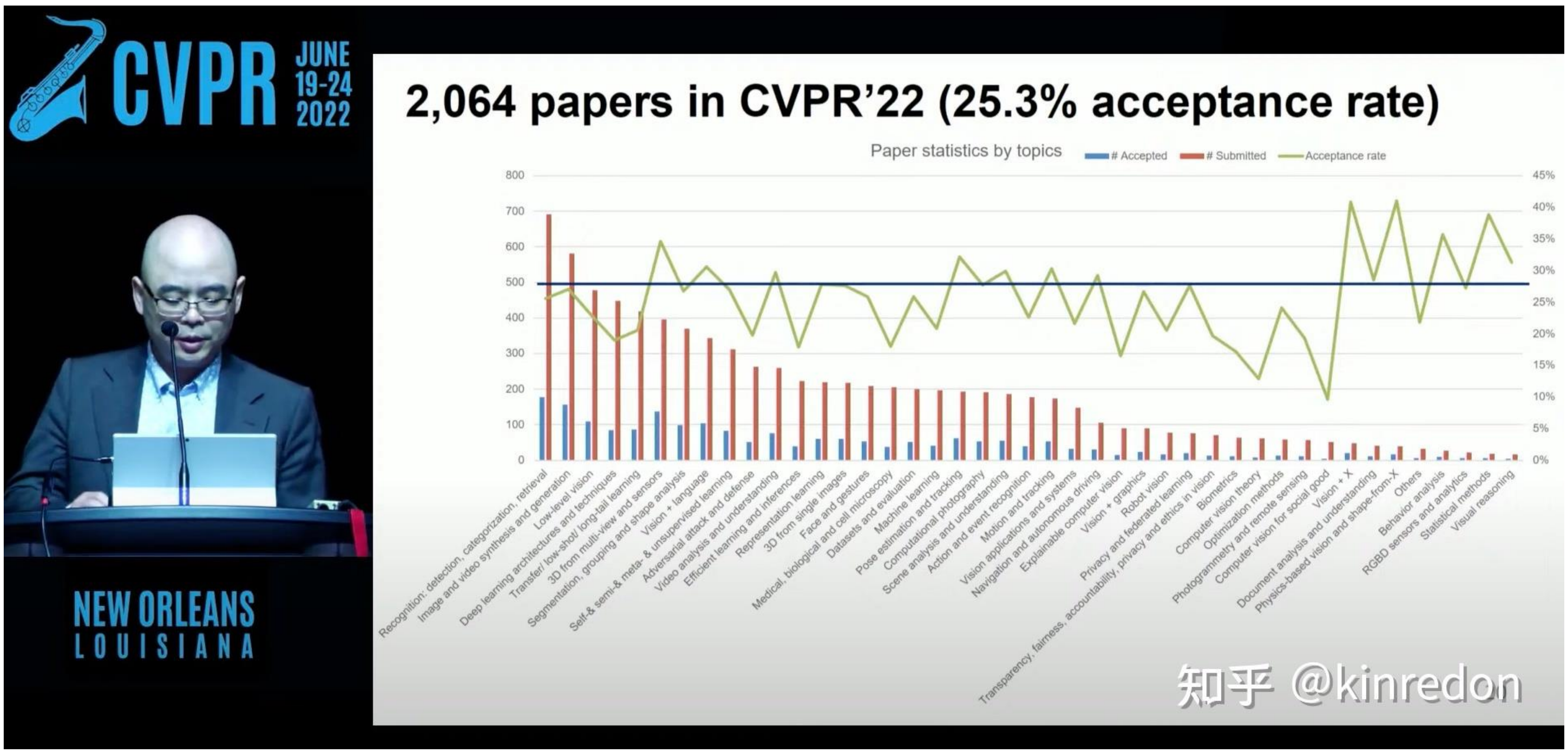
Domain Adaptation简介

柯水洲

2022.7.21

一 技术现状

Domain Adaptation 是迁移学习中的一个重要方向，每年在各大顶会上都有很多文章被接受，以 CVPR 2022 为例，其中 transfer/low-shot/long-tail learning 在众多 Topic 中排第四，有 400+ 投稿，接受的文章近百。



一 技术现状

Training Data

Testing Data

99.5% 57.5%

The results are from: http://proceedings.mlr.press/v37/ganin15.pdf

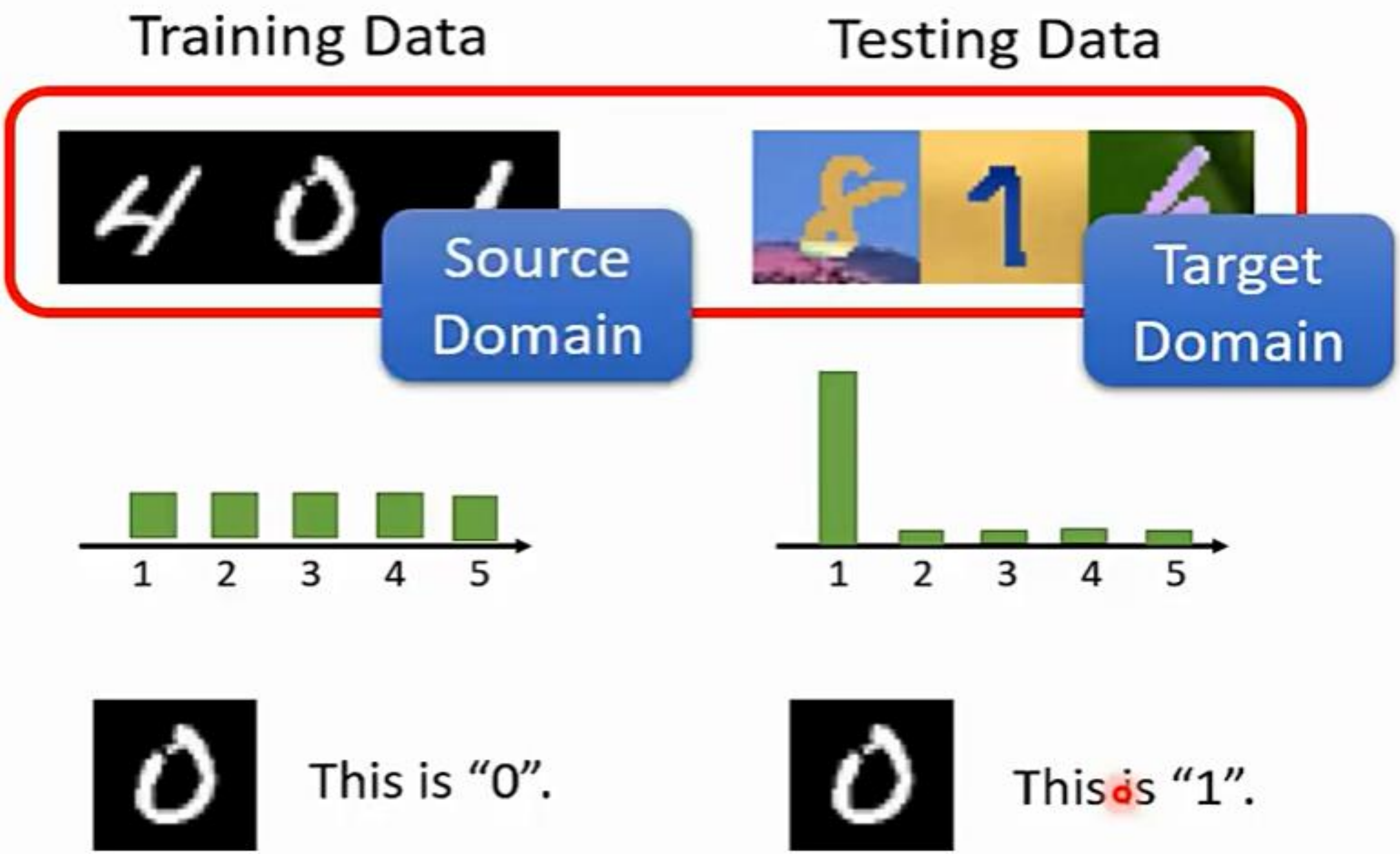
Domain shift: Training and testing data have different distributions. → **Domain adaptation**

一 技术现状

Domain Shift

$p(x,y)=p(x|y)p(y)=p(y|x)p(x)$

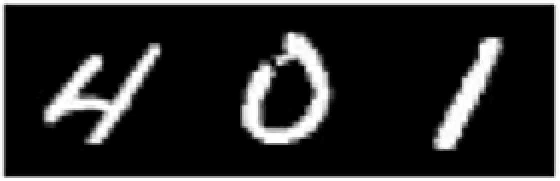
输入的边缘概率分布p（x）
输出标签的边缘概率分布p（y）
对应的条件概率分布p（x | y）



一 技术现状

Domain Adaptation

Source Domain
(with labeled data)



“4” “0” “1”

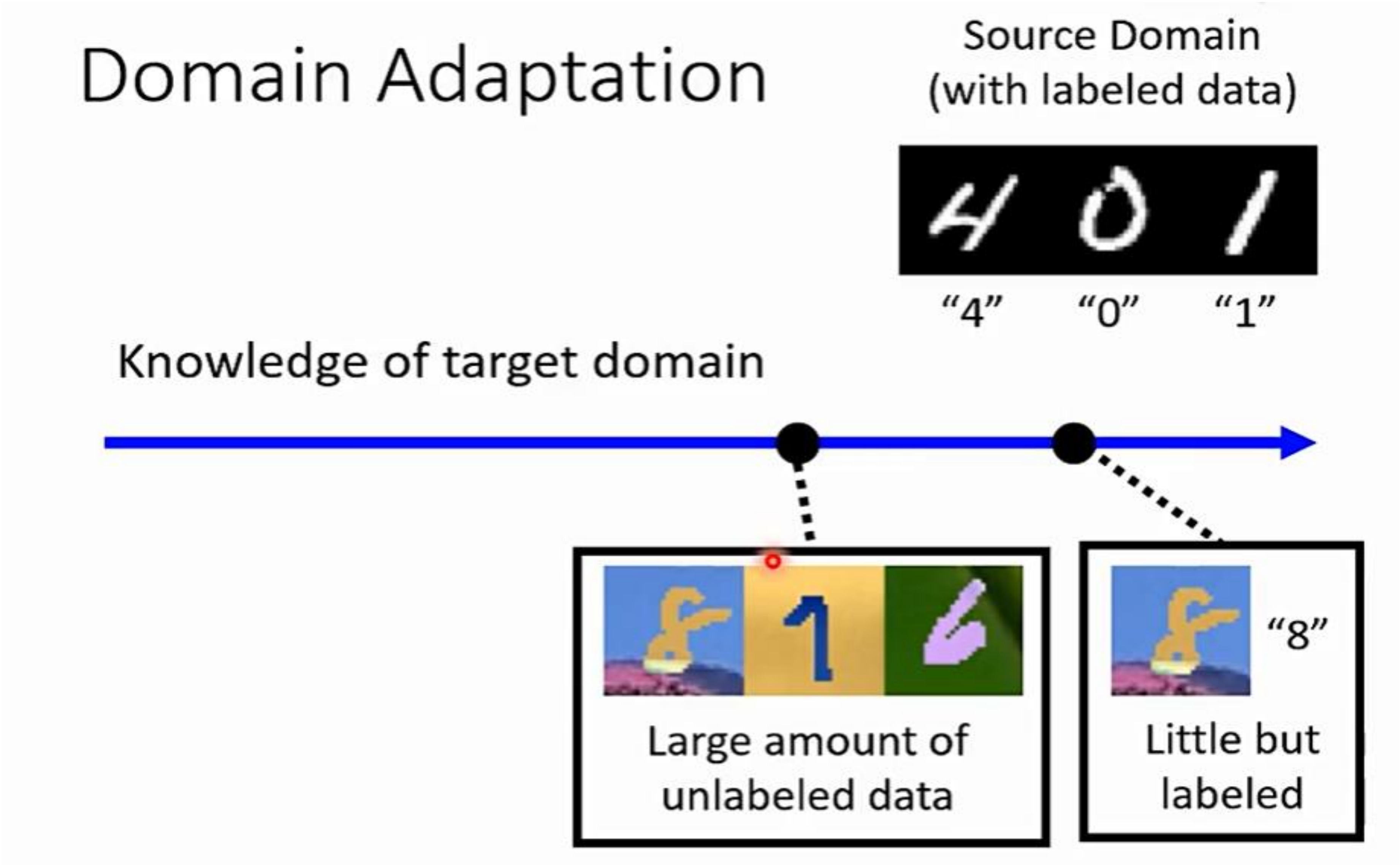
Knowledge of target domain

“8”

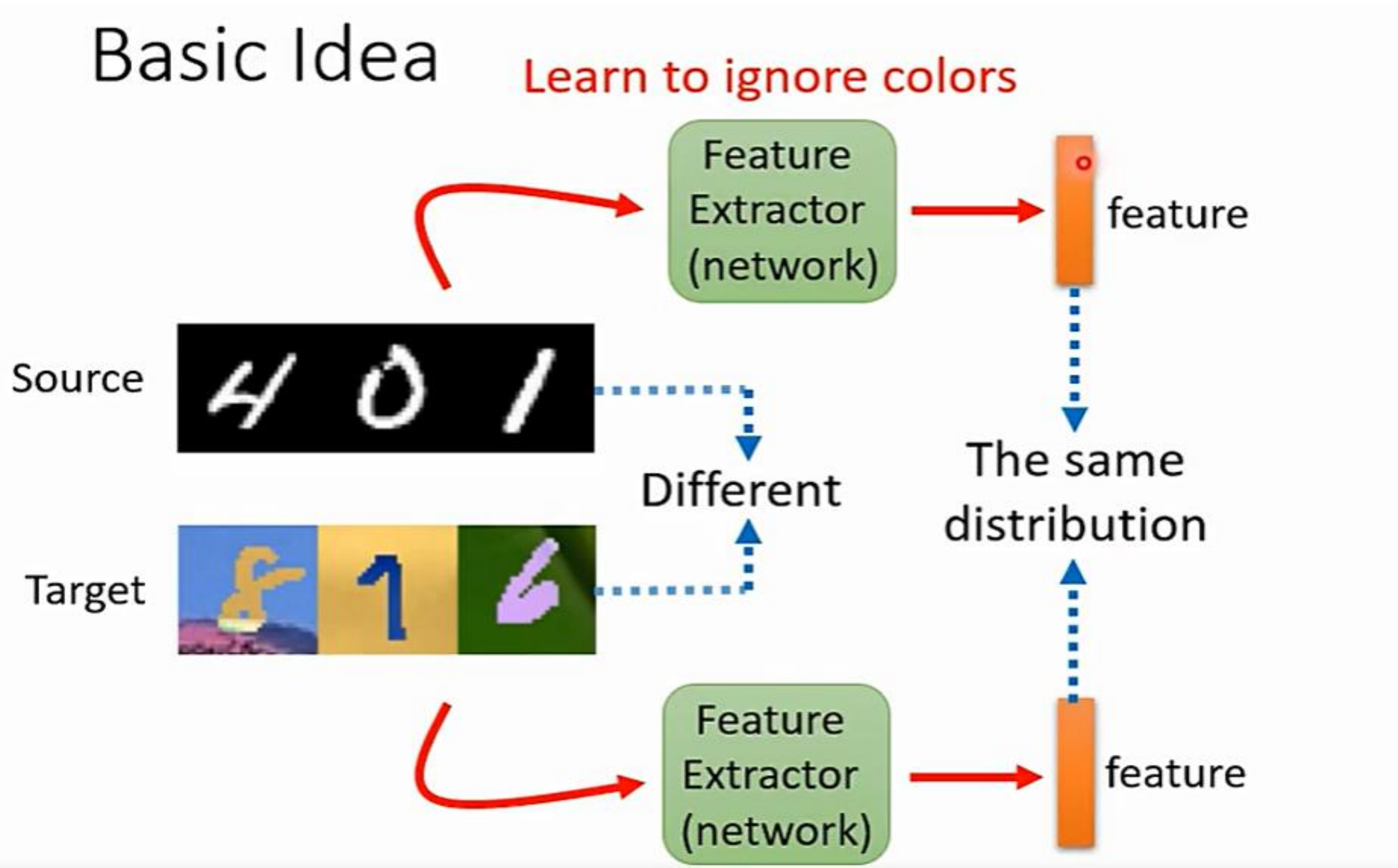
Little but labeled

- Idea: training a model by source data, then fine-tune the model by target data
- Challenge: only limited target data, so be careful about overfitting

一 技术现状

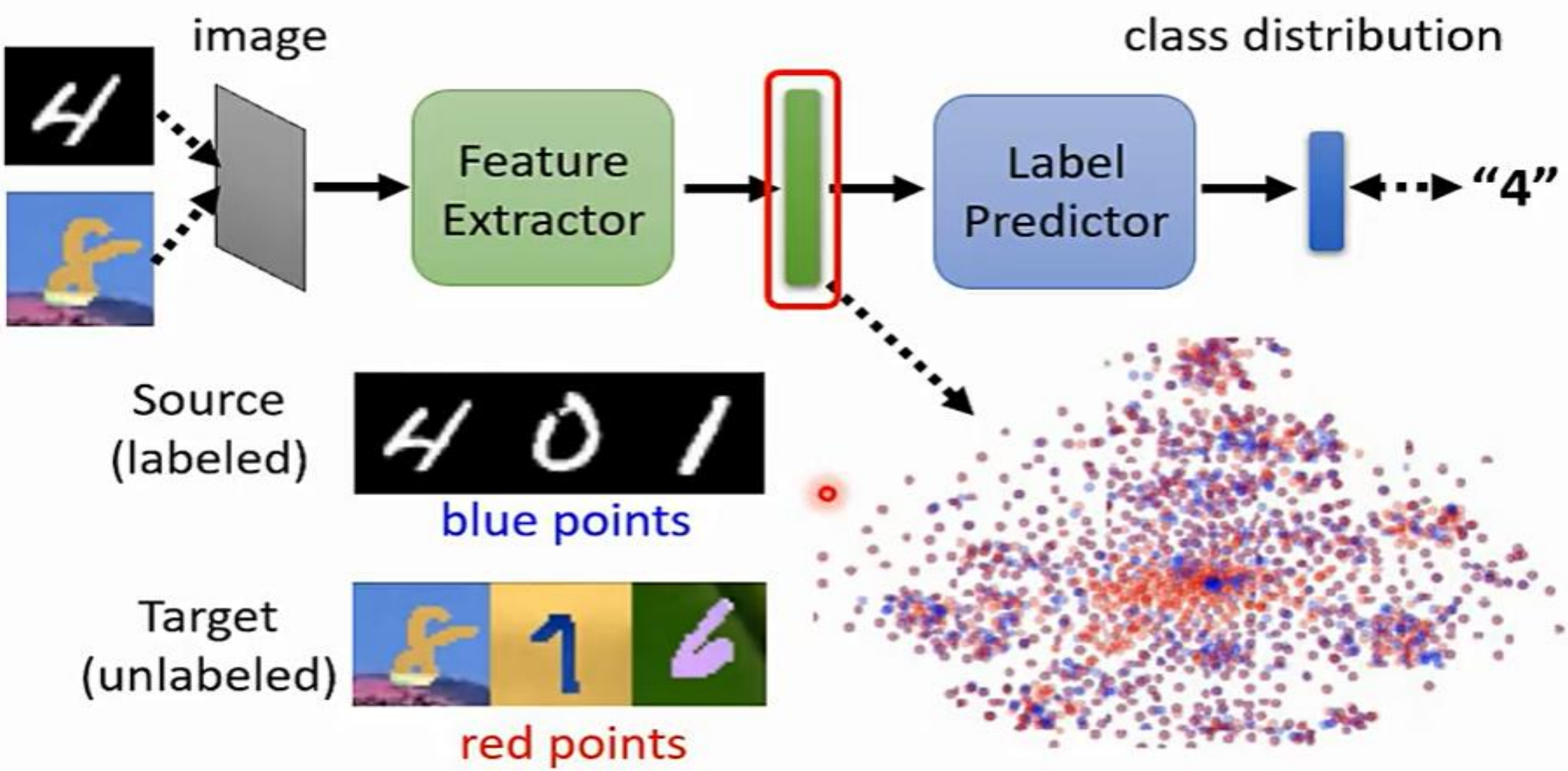

Domain Adaptation
Source Domain
(with labeled data)
"4" "0" "1"
Knowledge of target domain
Large amount of
unlabeled data
"8"
Little but
labeled

一 技术现状


Basic Idea
Learn to ignore colors
Feature
Extractor
(network)
feature
Source
Different
The same
distribution
Target
Feature
Extractor
(network)
feature

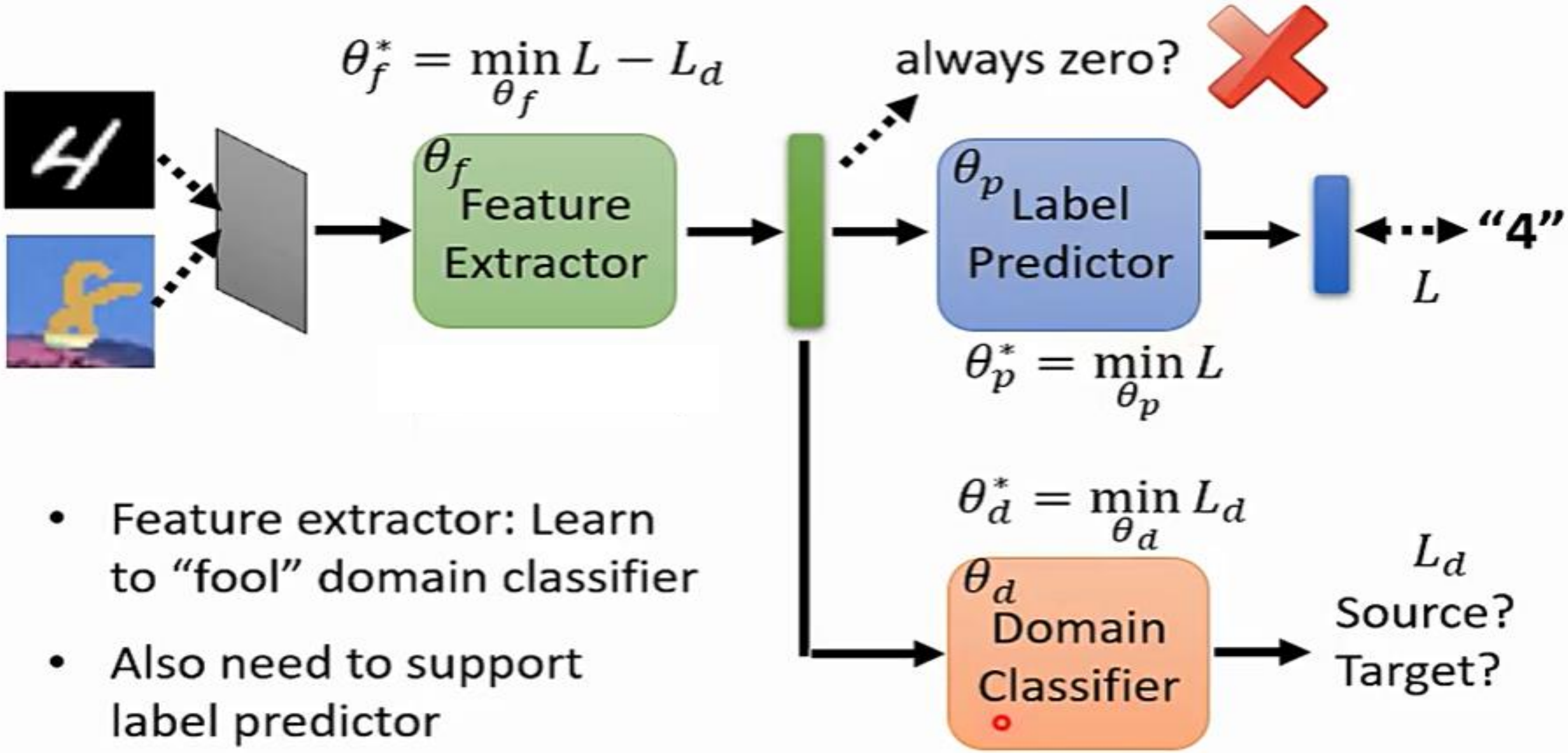
一 技术现状

Domain Adversarial Training



一 技术现状

Domain Adversarial Training



- Feature extractor: Learn to "fool" domain classifier
- Also need to support label predictor

一 技术现状

Domain Adversarial Training

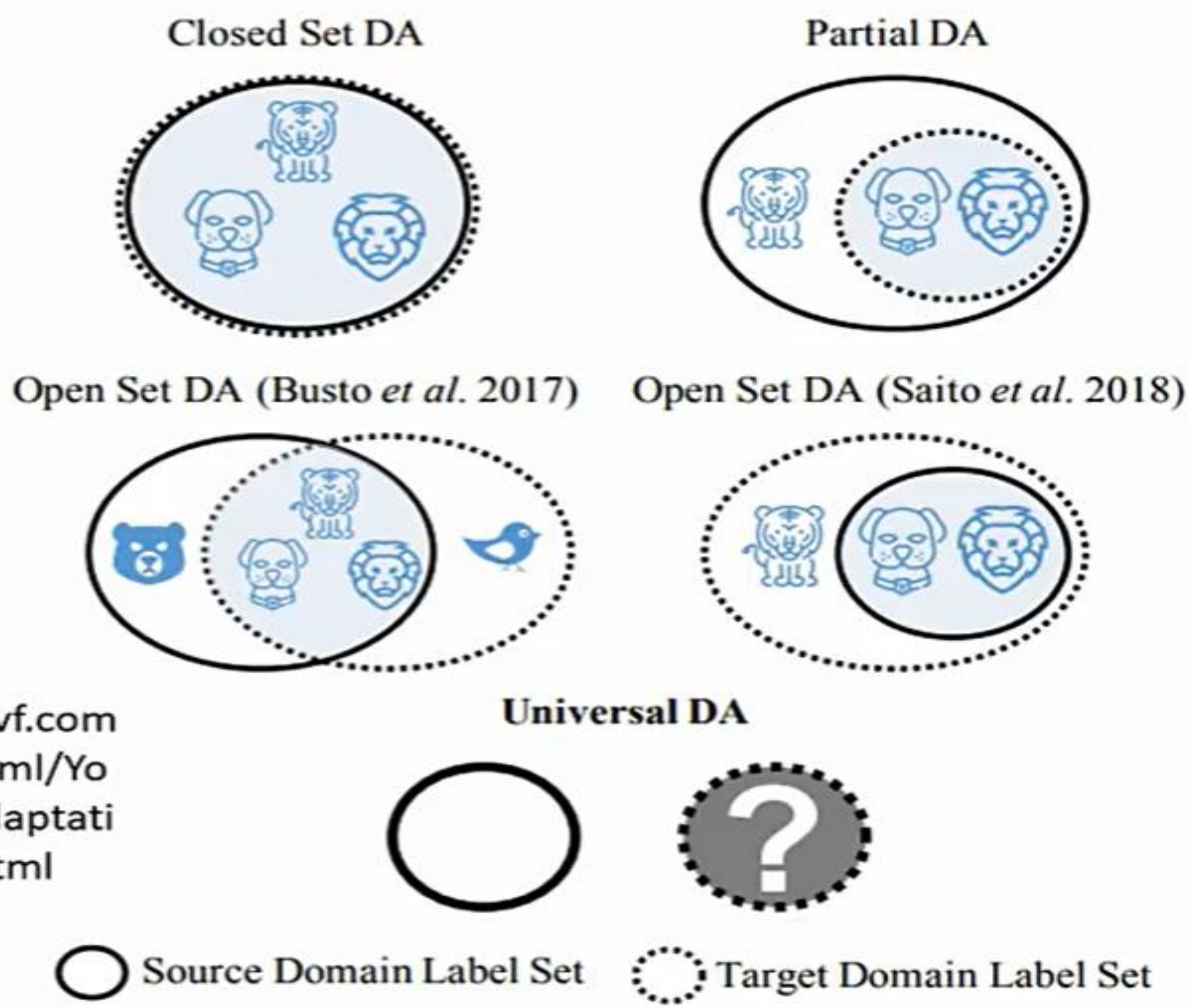
Yaroslav Ganin, Victor Lempitsky, Unsupervised Domain Adaptation by Backpropagation, ICML, 2015

Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, Domain-Adversarial Training of Neural Networks, JMLR, 2016

METHOD (SOURCE / TARGET)	MNIST / MNIST-M	SYN NUMBERS / SVHN	SVHN / MNIST	SYN SIGNS / GTSRB
SOURCE ONLY	.5749	.8665	.5919	.7400
PROPOSED APPROACH	**.8149** (57.9%)	**.9048** (66.1%)	**.7107** (29.3%)	**.8866** (56.7%)
TRAIN ON TARGET	.9891	.9244	.9951	.9987

一 技术现状

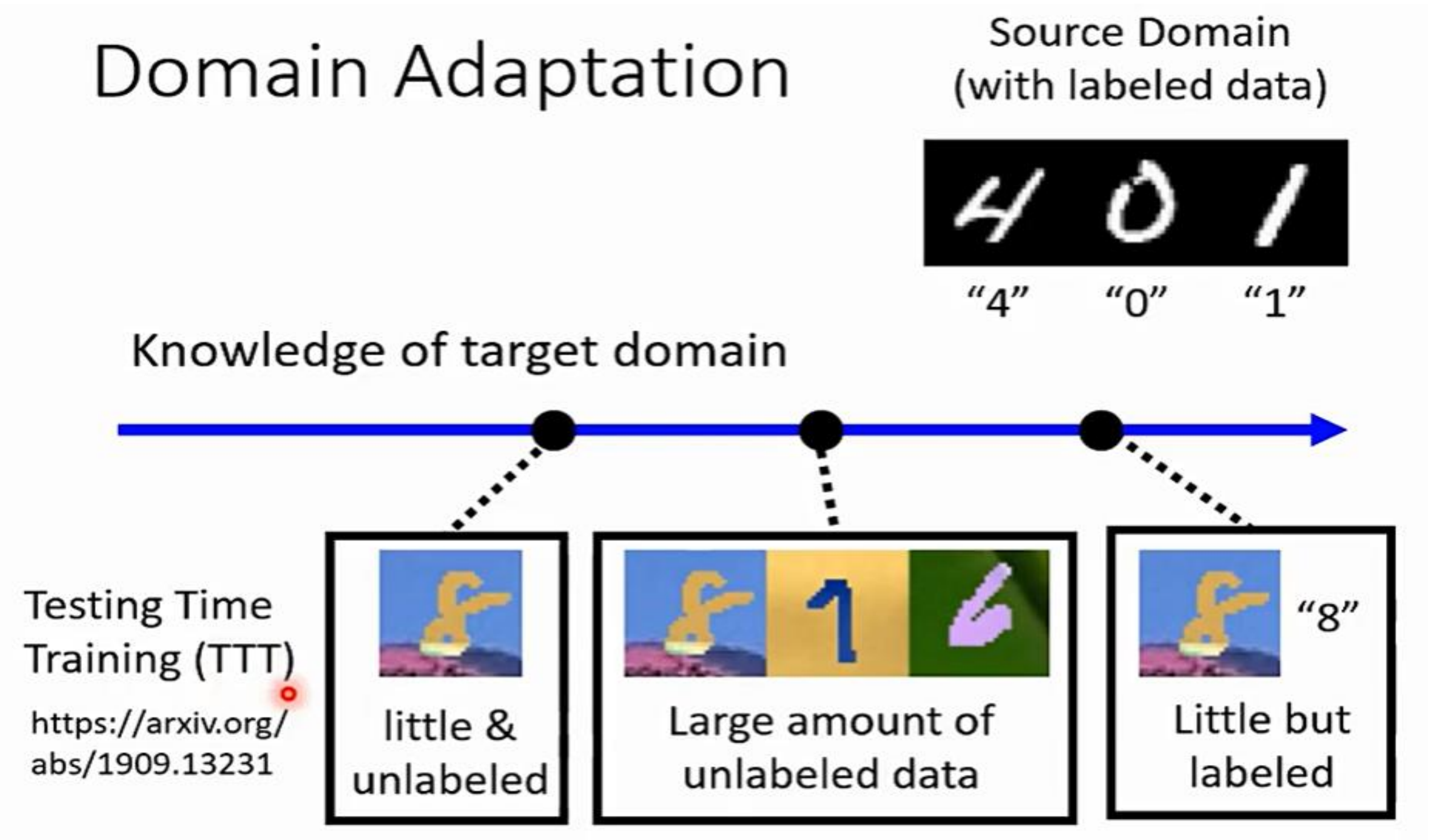
Outlook



Universal domain adaptation

https://openaccess.thecvf.com/content_CVPR_2019/html/You_Universal_Domain_Adaptation_CVPR_2019_paper.html

一 技术现状


Domain Adaptation
Source Domain
(with labeled data)
"4"
"0"
"1"
Knowledge of target domain
Testing Time
Training (TTT)
https://arxiv.org/
abs/1909.13231
little &
unlabeled
Large amount of
unlabeled data
"8"
Little but
labeled

二 论文应用

数据介绍

MICCAI 2018手术器械分割：达芬奇，肾切除术，9 种objects (8种器械+组织），11种语义关系（切割，烧灼，扎环等)

TORS数据集：达芬奇，口咽癌手术，5种objects，（4种器械+组织），5种语义关系（操作，抓握，烧灼等)

数据由自己标注：按照(object1，predicate，object2)的模式，如A monopolar curved scissor is cutting tissue.

L

GT: A monopolar curved scissor cutting tissue | **GT:** A prograsp forcep retracting tissue | **GT:** A clip applier is clipping a clip | **GT:** A suction is suctioning blood and a spatulated monopolar cautery is cauterizing tissue

(a) Source domain (b) Target domain

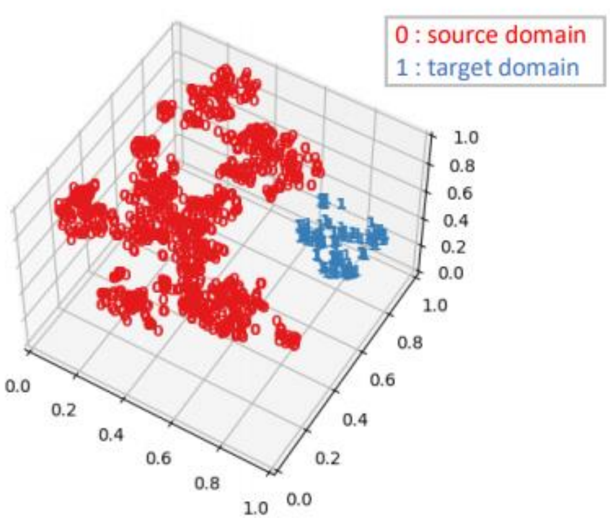
Xu, Mengya et al. Learning Domain Adaptation with Model Calibration for Surgical Report Generation in Robotic Surgery, ICRA 2021,6

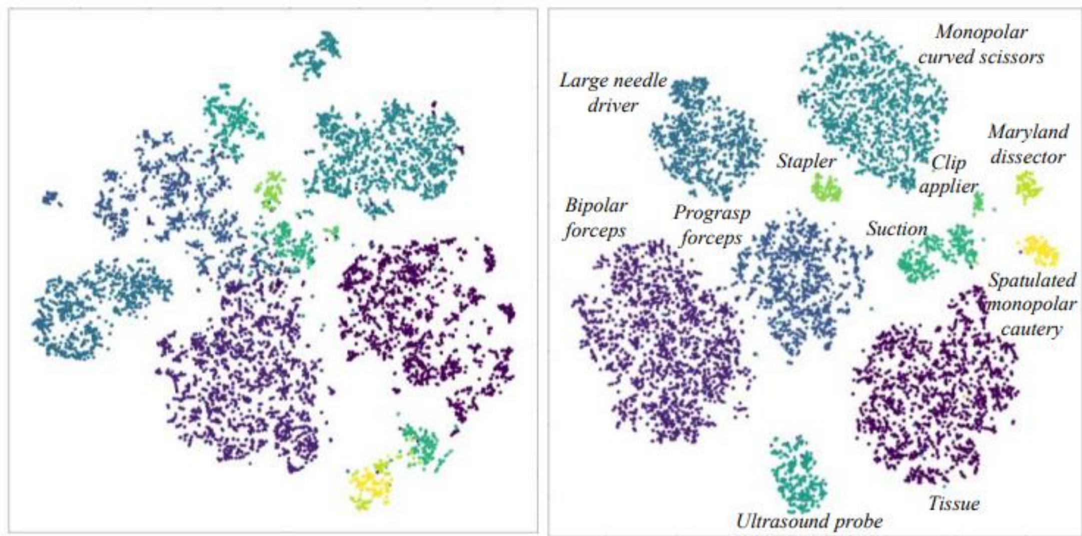
二 论文应用

文章内容

Label Smoothing：分类中防止过拟合的方法，用于模型校准

Domain Adaptation：源域和目标域往往属于同一类任务，但是分布不同。下图可视化展示不同域数据之间的分布情况





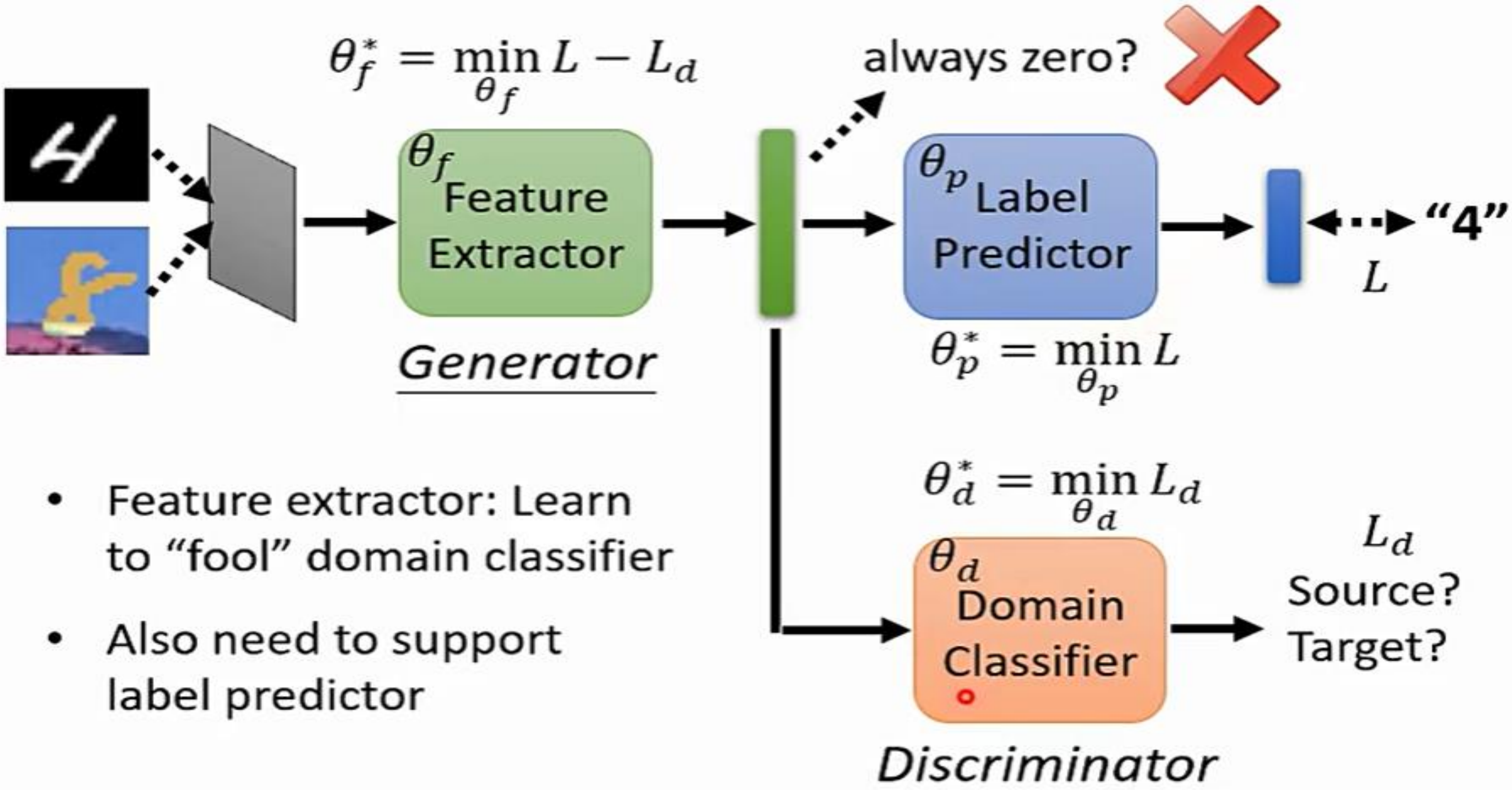
W/O LS　　W LS

Domain Shift：背景、器械、组织

提取特征的类别分布

二 论文应用

Domain Adversarial Training

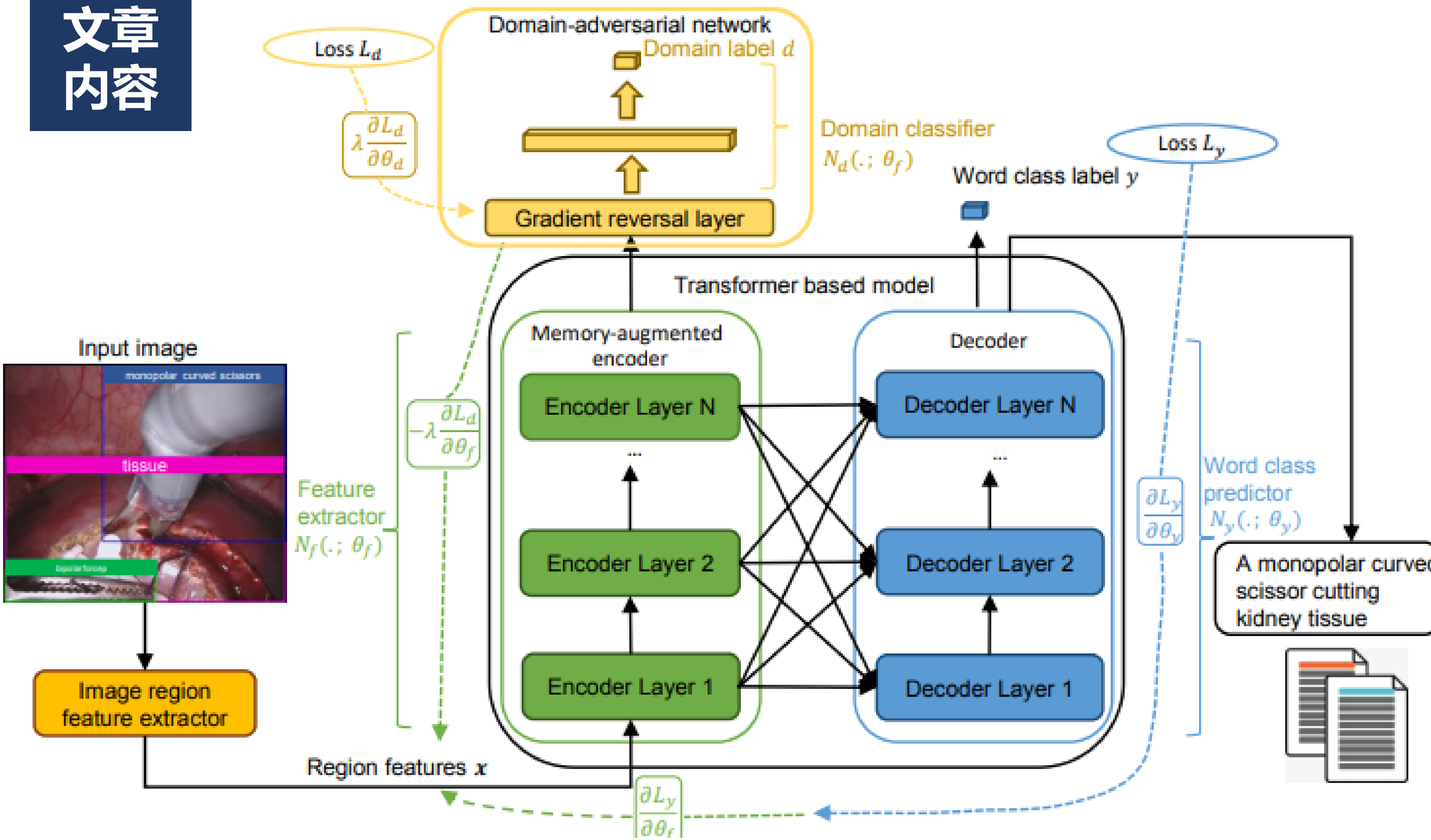


- Feature extractor: Learn to "fool" domain classifier
- Also need to support label predictor

Ganin Y. Unsupervised domain adaptation by backpropagation[C]//ICML. 2015.

二 论文应用

文章内容



Domain-adversarial network
Loss L_d
Domain label d
Domain classifier
$N_d(.;\theta_f)$
$\lambda\frac{\partial L_d}{\partial \theta_d}$
Gradient reversal layer
Loss L_y
Word class label y
Transformer based model
Memory-augmented encoder
Encoder Layer N
...
Encoder Layer 2
Encoder Layer 1
Decoder
Decoder Layer N
...
Decoder Layer 2
Decoder Layer 1
Input image
monopolar curved scissors
tissue
Feature extractor
$N_f(.;\theta_f)$
$-\lambda\frac{\partial L_d}{\partial \theta_f}$
Word class predictor
$N_y(.;\theta_y)$
$\frac{\partial L_y}{\partial \theta_y}$
A monopolar curved scissor cutting kidney tissue
Image region feature extractor
Region features x
$\frac{\partial L_y}{\partial \theta_f}$

二 论文应用

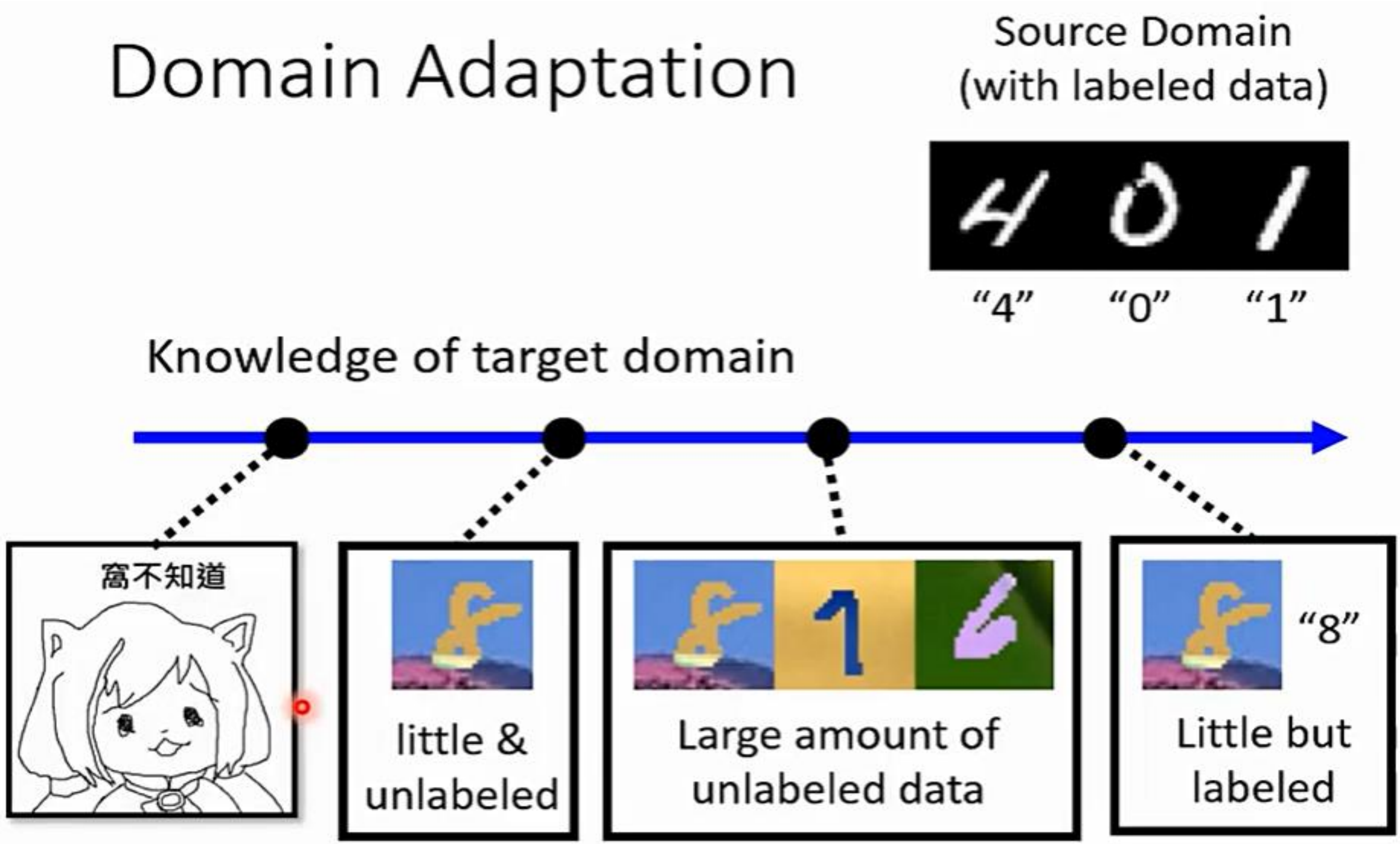
实验结果

实验方面：分别在SD和TD使用常用的无监督和半监督设置：Unsupervised DA、zero-shot、one-shot、few-shot 进行实验验证。

验证集不变，区别只在于训练集上：UDA，在SD训练模型的直接用于TD评估；zero-shot，使用TD中的覆盖了85%词汇的caption进行训练；One-shot，包括所有词语类别的最少图像样本

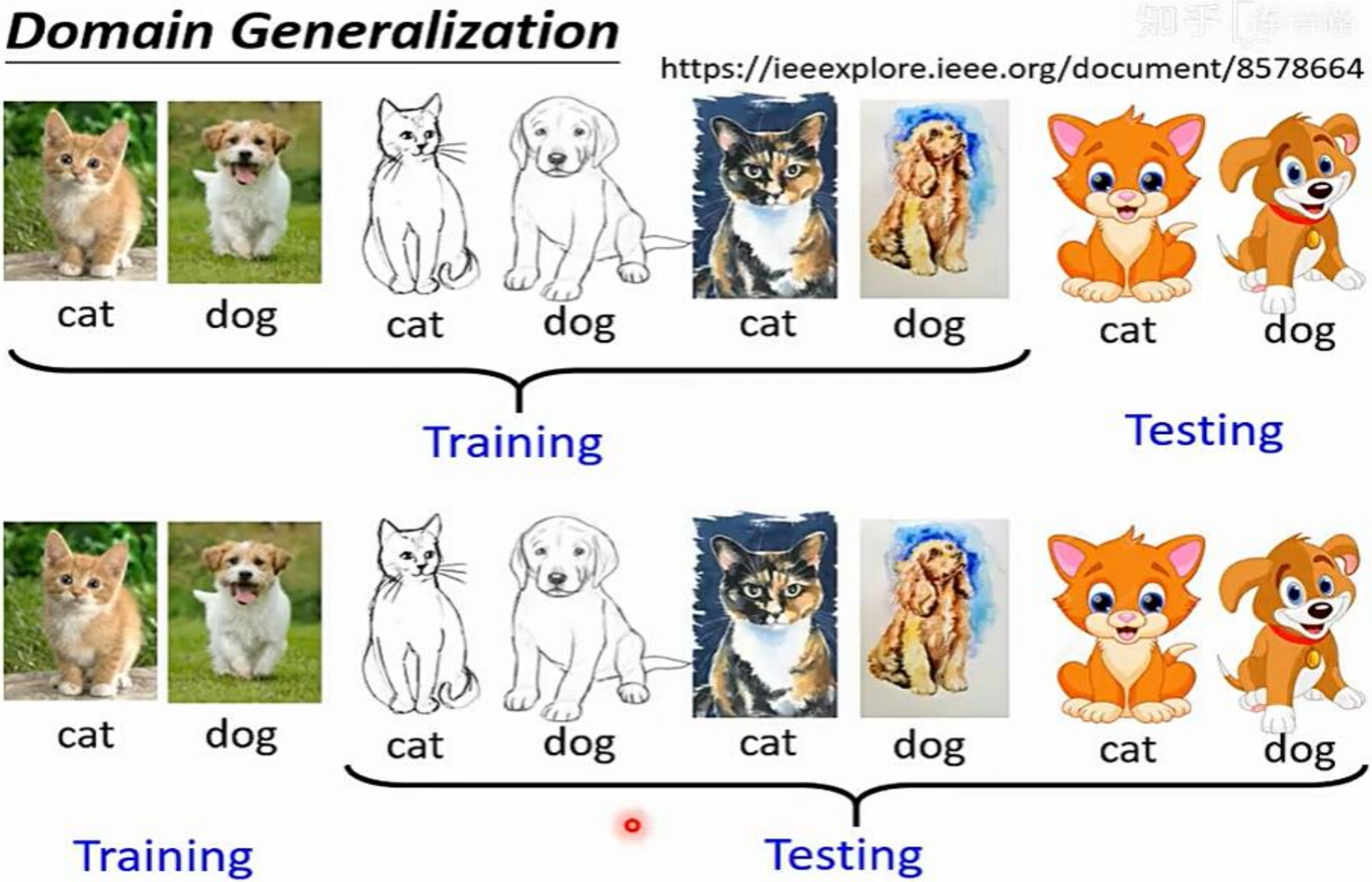
			BLEU-1↑	BLEU-2↑	BLEU-3↑	BLEU-4↑	METEOR↑	ROUGE↑	CIDEr↑
SD		M^2 Transformer [18]	0.5054	0.4543	0.4055	0.3646	0.4441	0.6355	1.7878
		Ours	0.5228	0.4730	0.4262	0.3861	0.4567	0.6495	2.2598
TD	UDA	M^2 Transformer [18]	0.2302	0.1059	0.0469	0.0267	0.1286	0.2956	0.1305
		Ours	0.2493	0.1150	0.0517	0.0289	0.1390	0.3129	0.1517
	Zero-shot	M^2 Transformer [18]	0.3204	0.2463	0.1923	0.1502	0.2371	0.4413	0.2874
		Ours	0.3118	0.2406	0.185	0.1409	0.2401	0.4336	0.3395
	One-shot	M^2 Transformer [18]	0.3746	0.3285	0.2939	0.2646	0.3449	0.5101	0.6367
		Ours	0.4042	0.372	0.3433	0.3161	0.4066	0.5385	0.8615
	Few-shot	M^2 Transformer [18]	0.4096	0.3803	0.3532	0.3265	0.4203	0.5489	0.9770
		Ours	0.4141	0.3888	0.3637	0.3375	0.4357	0.5538	0.9828

三 拓展总结

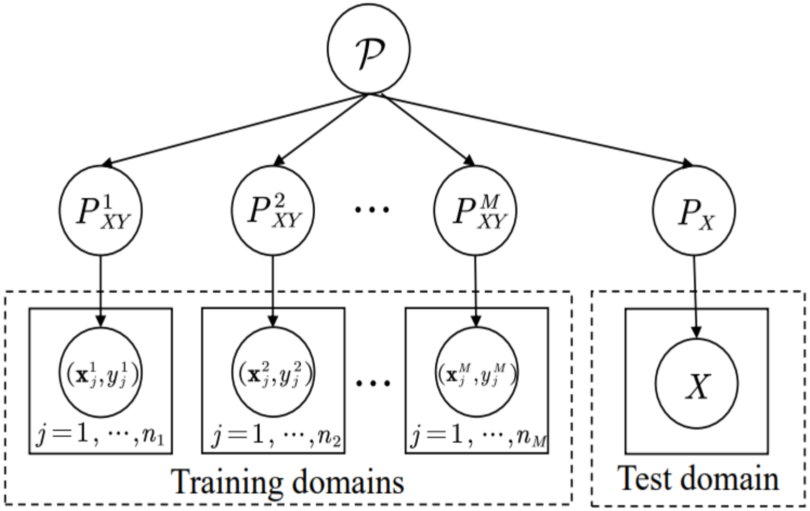


Domain Adaptation
Source Domain
(with labeled data)
"4" "0" "1"
Knowledge of target domain
窩不知道
little &
unlabeled
Large amount of
unlabeled data
"8"
Little but
labeled

三 拓展总结


Domain Generalization
https://ieeexplore.ieee.org/document/8578664
cat
dog
cat
dog
cat
dog
cat
dog
Training
Testing
cat
dog
cat
dog
cat
dog
cat
dog
Training
Testing

三 拓展总结



定义1：X、Y是输入和输出变量，P_{xy}是数据分布。

定义2：S表示training domains。每个domain之间的数据分布P不同。DG的目的就是从多个training domains中学习一函数h，再test domain中达到精度最高。

Definition 1 (Domain). *Let $\mathcal{X}$ denote a nonempty input space and $\mathcal{Y}$ an output space. A domain is composed of data that are sampled from a distribution. We denote it as $\mathcal{S} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n \sim P_{XY}$, where $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^d$, $y \in \mathcal{Y} \subset \mathbb{R}$ denotes the label, and P_{XY} denotes the joint distribution of the input sample and output label. X and Y denote the corresponding random variables.*

Definition 2 (Domain generalization). *As shown in Fig. 2, in domain generalization, we are given M training (source) domains $\mathcal{S}_{train} = \{\mathcal{S}^i \mid i = 1, \cdots, M\}$ where $\mathcal{S}^i = \{(\mathbf{x}_j^i, y_j^i)\}_{j=1}^{n_i}$ denotes the i-th domain. The joint distributions between each pair of domains are different: $P_{XY}^i \neq P_{XY}^j, 1 \leq i \neq j \leq M$. The goal of domain generalization is to learn a robust and generalizable predictive function $h : \mathcal{X} \to \mathcal{Y}$ from the M training domains to achieve a minimum prediction error on an unseen test domain $\mathcal{S}_{test}$ (i.e., $\mathcal{S}_{test}$ cannot be accessed in training and $P_{XY}^{test} \neq P_{XY}^i$ for $i \in \{1, \cdots, M\}$):*

$$\min_h \mathbb{E}_{(\mathbf{x},y)\in\mathcal{S}_{test}}[\ell(h(\mathbf{x}), y)], \tag{1}$$

where $\mathbb{E}$ is the expectation and $\ell(\cdot,\cdot)$ is the loss function.

Wang J, Lan C, Liu C, et al. Generalizing to unseen domains: A survey on domain generalization[J]. IEEE Transactions on Knowledge and Data Engineering, 2022.

三 拓展总结

多任务学习：针对一批数据，一次性学习多个任务（不针对新的domain）

迁移学习：再source任务训练的模型，用来增强target任务结果，更多使用Pretrain Finetune的策略（target数据不可见，且source和target任务一样的，分布不同）

领域适应：可以访问测试数据，实现测试任务上的精度最高（target不可见）

元学习：根据已有的数据和任务，学习一个函数用于新的数据（是DG中常用的学习策略）

终身学习：不断地接受新的数据、任务的同时，不遗忘以前学习的任务（对以前数据可见）

zero-shot学习：根据已有学习对未见过的数据进行分类（DG的target同类不同分布）

Learning paradigm	Training data	Test data	Condition	Test access
Multi-task learning	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{Y}^i \neq \mathcal{Y}^j, 1 \leq i \neq j \leq n$	✓
Transfer learning	$\mathcal{S}^{src}, \mathcal{S}^{tar}$	$\mathcal{S}^{tar}$	$\mathcal{Y}^{src} \neq \mathcal{Y}^{tar}$	✓
Domain adaptation	$\mathcal{S}^{src}, \mathcal{S}^{tar}$	$\mathcal{S}^{tar}$	$P(\mathcal{X}^{src}) \neq P(\mathcal{X}^{tar})$	✓
Meta-learning	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{S}^{n+1}$	$\mathcal{Y}^i \neq \mathcal{Y}^j, 1 \leq i \neq j \leq n+1$	✓
Lifelong learning	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{S}^i$ arrives sequentially	✓
Zero-shot learning	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{S}^{n+1}$	$\mathcal{Y}^{n+1} \neq \mathcal{Y}^i, 1 \leq i \leq n$	×
Domain generalization	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{S}^{n+1}$	$P(\mathcal{S}^i) \neq P(\mathcal{S}^j), 1 \leq i \neq j \leq n+1$	×

Wang J, Lan C, Liu C, et al. Generalizing to unseen domains: A survey on domain generalization[J]. IEEE Transactions on Knowledge and Data Engineering, 2022.

Q&A

合肥工業大學