

# 迁移学习概述

2020.4

# 什么是迁移学习

- 是指利用数据、任务或模型之间的相似性，将在旧领域学习过的模型，应用于新领域的一种学习过程。
- **核心问题：**找到新问题和原问题的相似性，实现知识的迁移。

冬末春初，  
北京的天气渐渐暖了起来，

纽约的天气\_\_\_\_\_  
东京的天气\_\_\_\_\_  
布宜诺斯艾利斯的天气\_\_\_\_\_

• 举一反三

• 照猫画虎

• 东施效颦

## 负迁移 (Negative Transfer)

如果这个相似性找的不合理（两个领域之间不存在相似性，或者基本不相似），就会大大损害迁移学习的效果，称为负迁移。

比如：拿骑自行车的经验来学习开汽车



热点新闻

19-12-3 08:00 from 微博云剪



【不会打牌的段子手不是好棋手！柯洁拿下斗地主冠军】11月30日，围棋世界冠军柯洁成功拿下了腾讯斗地主锦标赛全民星赛冠军。柯洁9局结束以8300分的成绩获得冠军。首次参加斗地主比赛就取得了这么好的成绩。柯洁表示很激动，这是第一次拿到斗地主的冠军。对自己的冠军成绩很满意，只是对前两局的牌面不太满意，整体发挥没有什么致命失误。[辽视说天下的微博视频](#)



游戏频道 网络游戏 · 64.5K views

Share



# 为什么要研究迁移学习

## 大数据与少标注之间的矛盾

数据的标注是一个耗时且昂贵的操作，这给机器学习和深度学习的模型训练和更新带来了挑战。特定的领域可能因为没有足够的标定数据用来学习，使得这些领域一直不能很好的发展。

## 大数据与弱计算之间的矛盾

计算能力=\$\$\$. 如何降低成本，让普通人也能利用这些数据和模型就成了亟待解决的问题

## 普适化模型与个性化需求之间的矛盾

当前，我们对于每一个通用的任务都构建了一个通用的模型，来解决绝大多数的公共问题。但是具体到每个个体、每个需求，都存在其唯一性和特异性，一个普适化的通用模型根本无法满足。  
例如：导航模型，可以定位及导航所有的路线。但是不同的人有不同的需求，有的人喜欢走高速，有的人喜欢走偏僻小路，这就是个性化需求。

## 特定应用的需求

一个新的推荐系统，没有足够的用户数据，如何进行精准的推荐？  
一个崭新的图片标注系统，没有足够的标签，如何进行精准的服务？  
(冷启动问题)

# 迁移学习的分类

|        | 源域 & 目标域 | 源任务 & 目标任务 | 数据标签 |         | 任务方法  |
|--------|----------|------------|------|---------|---|
|        |          |            | 源数据  | 目标数据    |   |
| 传统机器学习 | 相同       | 相同         | 有标签  | 有标签     | 验证  |
|        |          |            |      | 无标签     | 测试  |
|        |          |            |      | 无标签     | 聚类  |
| 迁移学习   | 相同/相关    | 相关         | 有标签  | 有标签     | 多任务学习<br>fine-tune<br>自适应<br>TrAdaBoost           |
|        | 相关       | 相同/相关      | 有标签  | 无标签     | Domain-adversarial training<br>Zero-shot Learning |
|        | 相同/相关    | 相关         | 无标签  | 无标签/有标签 | 自我学习  |

## 域：

进行学习的主体，主要由数据和生成这些数据的概率分布这两部分构成。

## 源域：

有知识或有大量数据标注的领域，是被要迁移的对象

## 目标域：

最终要赋予知识、赋予标注的对象。知识从源域传递到目标域，就完成了迁移。

## 任务：

学习的目标。比如图片分类，文本翻译，等等。

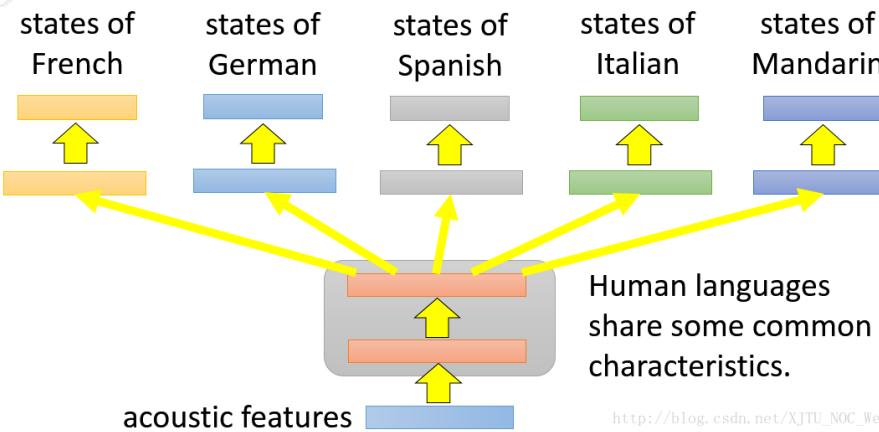
# 从迁移学习角度重看多任务学习与Fine-tune

## 多任务学习

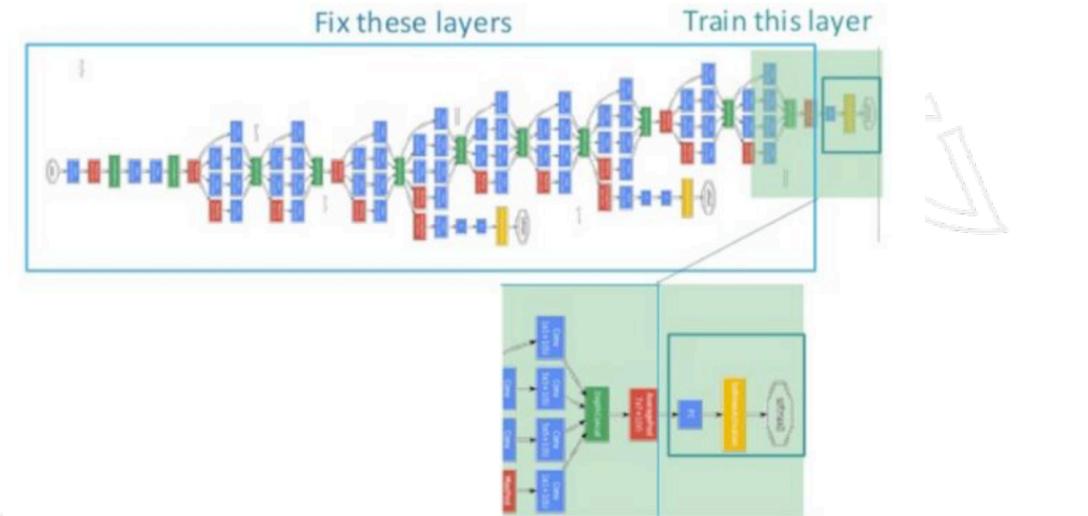
- 利用**多个学习任务**中所包含的有用信息来帮助提升**每个任务**的学习效果。在完成多个任务训练之后，产生的效果可能比训练单独的任务的效果好。

**例1：**从中文提取特定特征后，同时进行英文、日文、德语翻译训练

**例2：**从不同语言语音中提取发音特性进行中文、法语、德语识别



## Fine-tune



- 资源与时间成本小：**不需要针对新任务从头开始训练网络，而是利用别人已经训练好的网络，针对自己的任务再进行调整
- 实现简单：**只关注自己的任务即可
- 模型泛化性好：**预训练好的模型通常都是在大数据集上进行的。默认目标域样本数量远少于源域样本数量 (one-shot learning)

**Conservative training:** 用旧模型的参数初始化新模型，目标是使同样样本在两个模型的输出结果中  $L2$  Loss 最小

**Layer transfer:** 保留部分旧模型的参数，只用目标域数据调整一部分参数

# Fine-tune应用举例

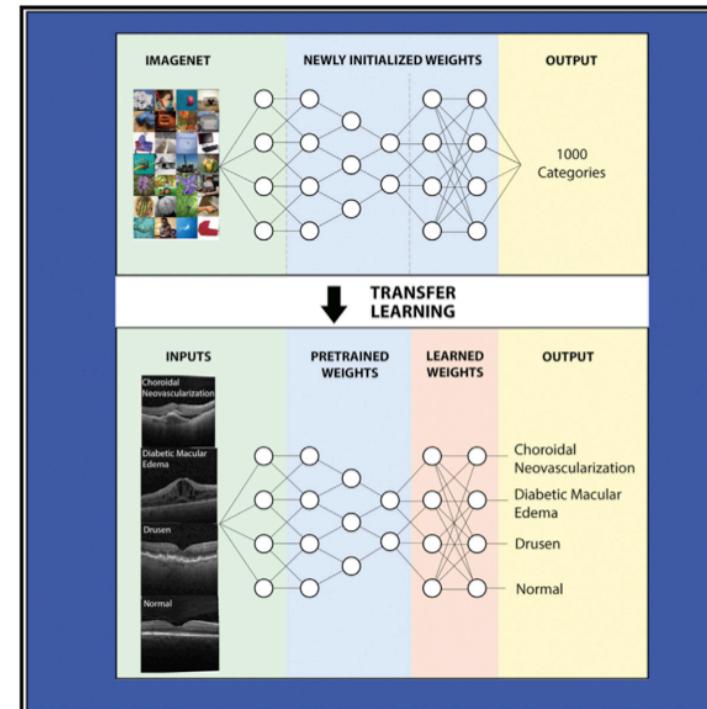
- 2018年《细胞》封面文章
- 广州妇女儿童医疗中心与加州大学圣迭戈分校合作完成
- 利用**ImageNet数据+InceptionV3+Fine-tune**的方法，实现了低资源视网膜OCT图像与肺部X光图像的人工智能识别
- 诊断结果**AUC超过99%**，超过部分人类专家
- OCT任务**训练数据10w+**（ImageNet图像1400w+），**测试集1000**，全部有标注（诊断信息）
- 没有进行非迁移学习模型的对比



Cell

## Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning

### Graphical Abstract



Resource

### Authors

Daniel S. Kermany, Michael Goldbaum, Wenjia Cai, ..., M. Anthony Lewis, Huimin Xia, Kang Zhang

### Correspondence

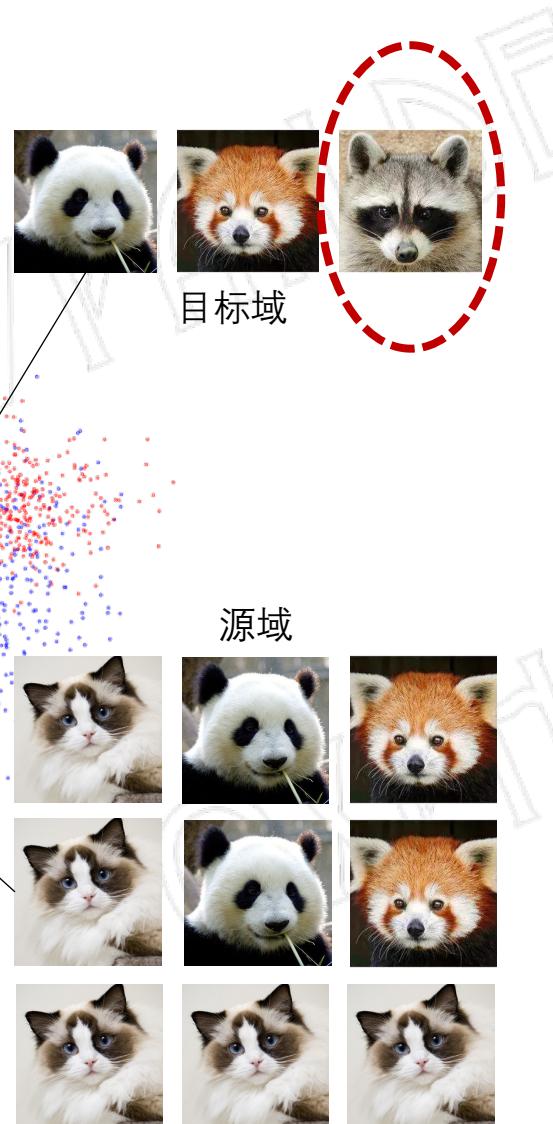
kang.zhang@gmail.com

### In Brief

Image-based deep learning classifies macular degeneration and diabetic retinopathy using retinal optical coherence tomography images and has potential for generalized applications in biomedical image interpretation and medical decision making.

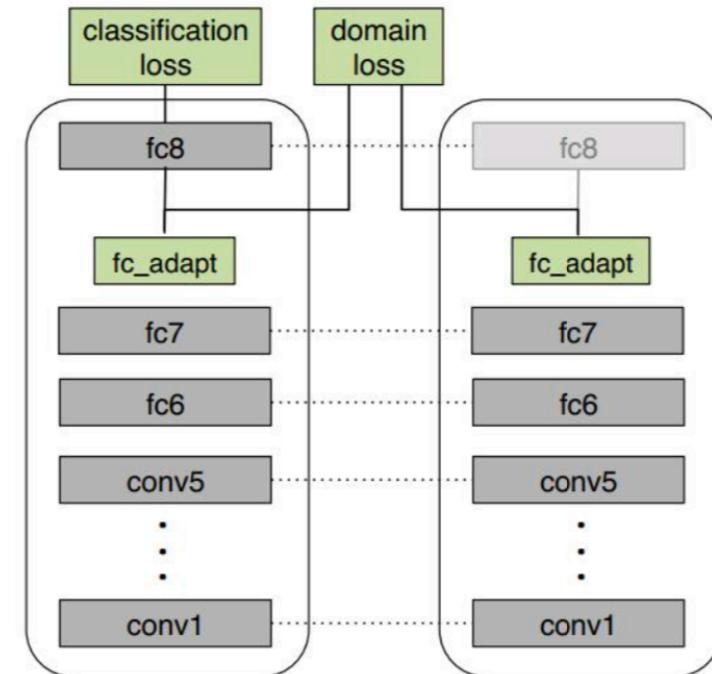
# 数据分布差异与自适应方法

在迁移学习中，通常假定源域和目标域的概率分布是不同且未知的，Fine-tune没有解决这一问题



- Maximum mean discrepancy适配层：MMD是计算两堆数据在高维空间中的均值距离的核函数。MMD层通过计算源域与目标域的距离，并被引入训练时的网络损失来减小域适应的问题。
- Deep Domain Confusion：深度网络的单层自适应。在AlexNet网络的倒数第二层加入了MMD层，
- Deep Adaptation Networks：多核MMD+多适配层

$$\ell = \ell_c(\mathcal{D}_s, \mathbf{y}_s) + \lambda MMD^2(\mathcal{D}_s, \mathcal{D}_t)$$



Deep Domain Confusion

# TrAdaBoost: 更新源域标签和目标域标签 (样本) 权重, 协同训练得到目标模型

## 算法基本思路：

在迁移时, 为了最大限度地和目标域相似, 我们可以人为地提高源域中目标域类别的样本权重。

- 合并源域和目标域的训练数据集
- 通过Boosting思路建立一种权重调整机制, 增加有效数据权重, 降低无效数据权重。

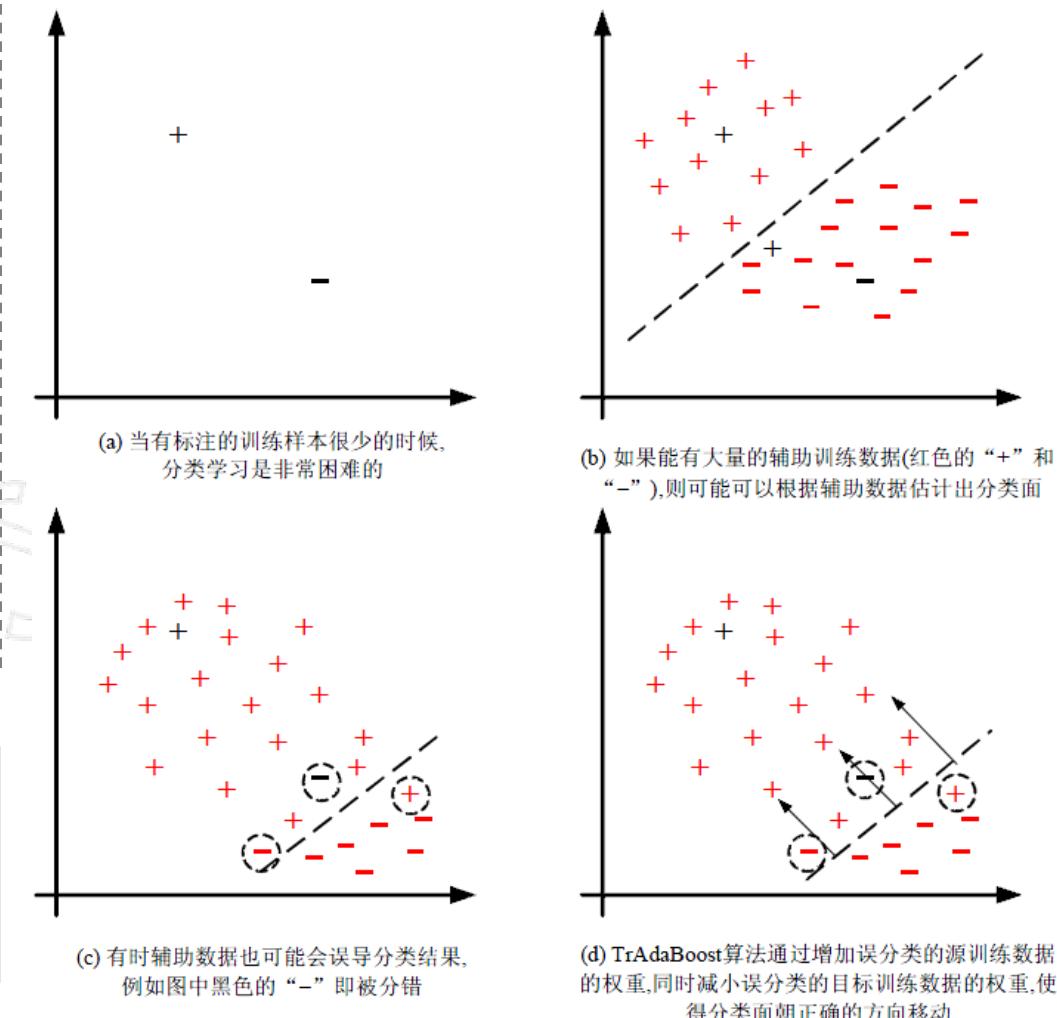
**有效数据**: 目标域训练数据中, 预测值与标签相差大的数据 (Boosting)

源域训练数据中, 与目标域数据匹配, 且预测值与标签接近的数据。

## 理想结果：

目标样本预测值与标签尽量匹配, 同时筛选出与目标域样本最匹配的 (权重大的) 源域样本作为训练辅助数据。

当源域与目标域同分布数据占比当低于0.1时, 算法效果明显, 当占比超过 0.1 时, TrAdaBoost 退化为 SVM 的效果。我们认为大于0.1时, 仅仅依靠目前数据就足够完成样本训练, 这种情况下, 辅助数据的贡献可以忽略。另外, 源域与目标域差别比较大时, 该方法效果较差

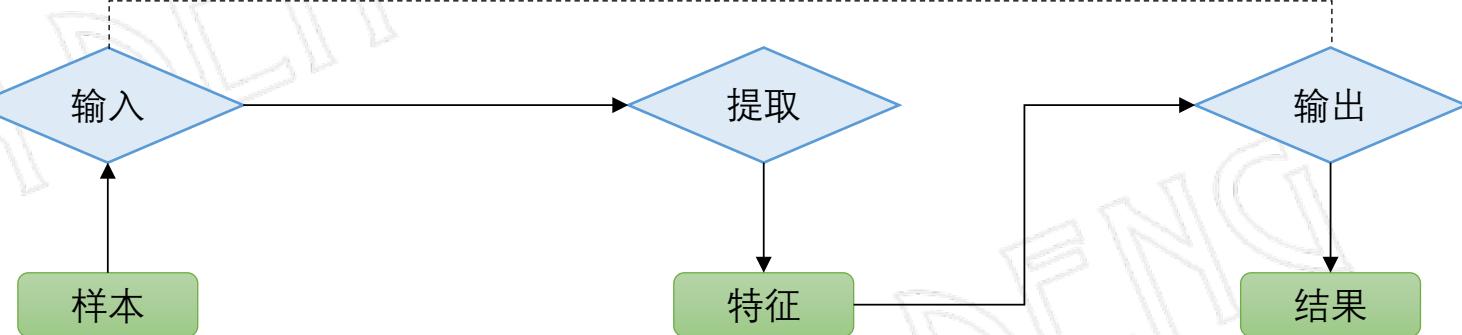


# 从迁移对象看迁移学习分类



## 基于参数的迁移：

发现源域和目标域之间的共享参数或先验关系。  
多任务学习/fine-tune



## 基于样本的迁移：

通过调整源域的标签和目标域标签的权重，协同训练得到目标模型。

## 基于特征的迁移：

把源域和目标域的特征变换到一个空间里，寻找“好”特征

自适应

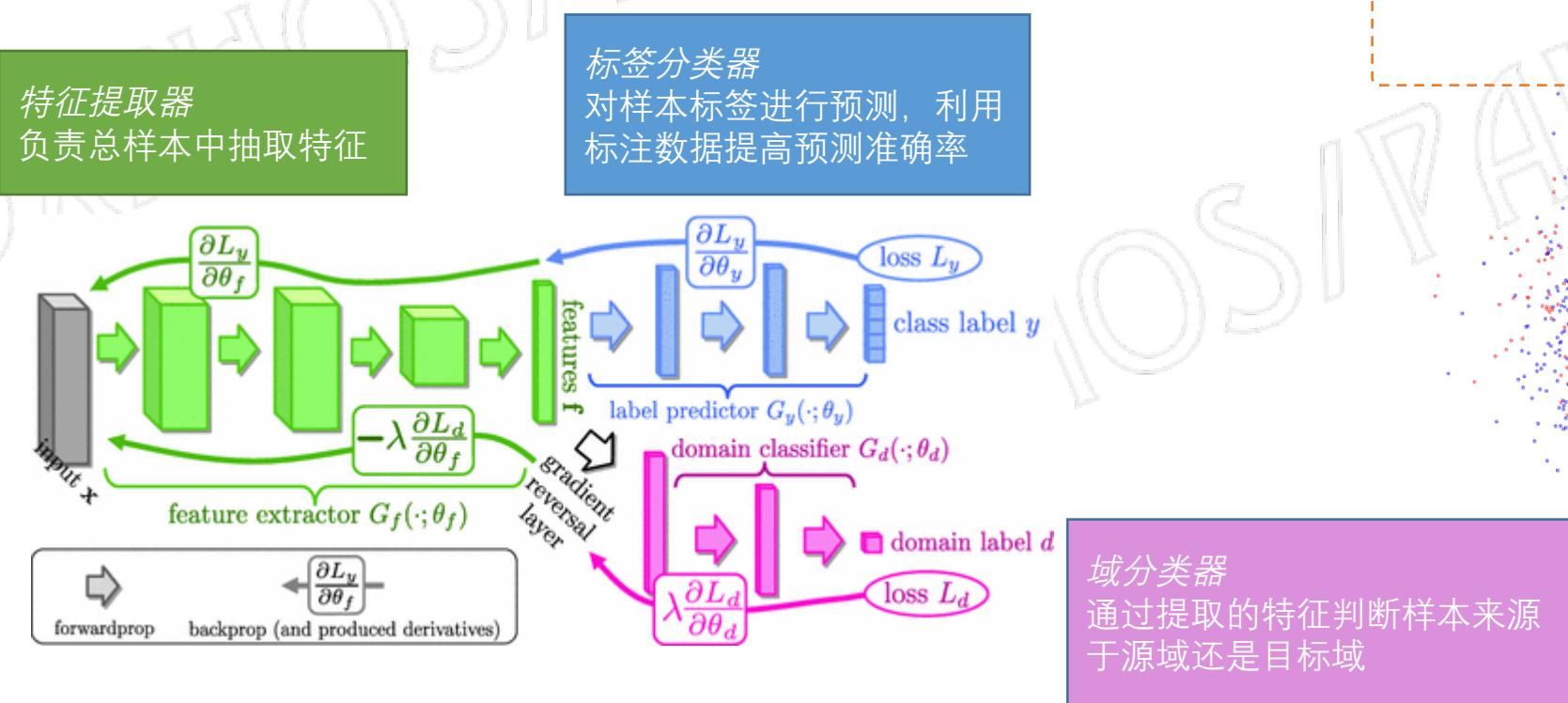
TrAdaBoost

# 迁移学习的分类

|        | 源域 & 目标域 | 源任务 & 目标任务 | 数据标签 |         | 任务方法  |         |
|--------|----------|------------|------|---------|---|---------|
|        |          |            | 源数据  | 目标数据    |   |         |
| 传统机器学习 | 相同       | 相同         | 有标签  | 有标签     | 验证  |         |
|        |          |            |      | 无标签     | 测试  |         |
|        |          |            |      | 无标签     | 聚类  |         |
| 迁移学习   | 相同/相关    | 相关         | 有标签  | 有标签     | 多任务学习<br>fine-tune                                | 基于参数的迁移 |
|        |          |            |      |         | 自适应<br>TrAdaBoost                                 | 基于样本的迁移 |
|        | 相关       | 相同/相关      | 有标签  | 无标签     | Domain-adversarial training<br>Zero-shot Learning | 基于特征的迁移 |
|        | 相同/相关    | 相关         | 无标签  | 无标签/有标签 | 自我学习  |         |

# Domain-adversarial training : 提取源域与目标域共同特征

- Domain-adversarial training与GAN的思路有相似之处，其中的域分类器的作用是**消除域特异性特征**，提取源域与目标域的通用特征。 (discriminator)
- 通过域分类器与标签分类器，找到非域特异的样本特征，并尽可能准确地进行预测 (**最大化标签分类器准确度，最小化域分类器准确度**)



# Domain-adversarial training : 提取源域与目标域共同特征

*Domain-adversarial training* 无法处理标签在源域中不存在的问题

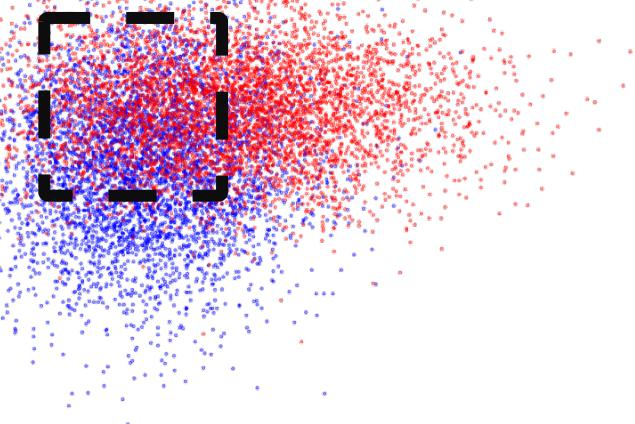
源域 :  $(D_s, L_s)$  ; 目标域 :  $(D_t, )$



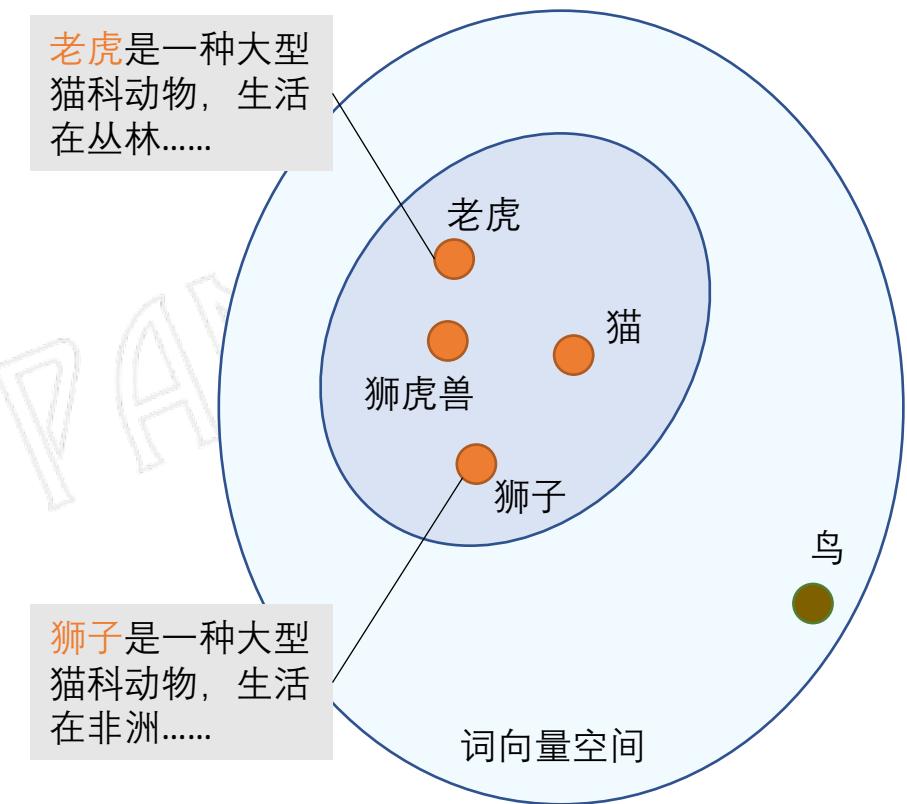
目标域



源域

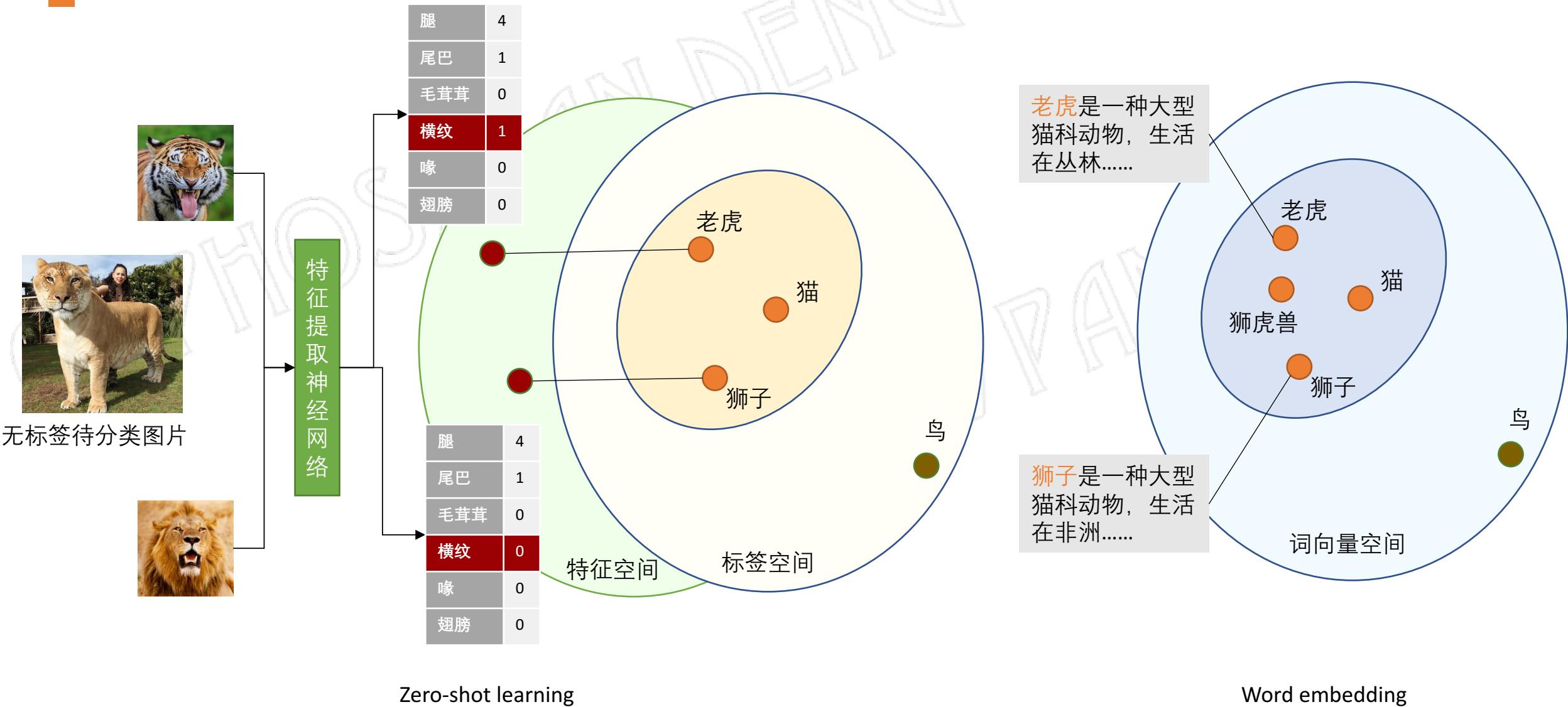


Zero-shot learning : 特征embedding, 建立特征->标签的映射函数

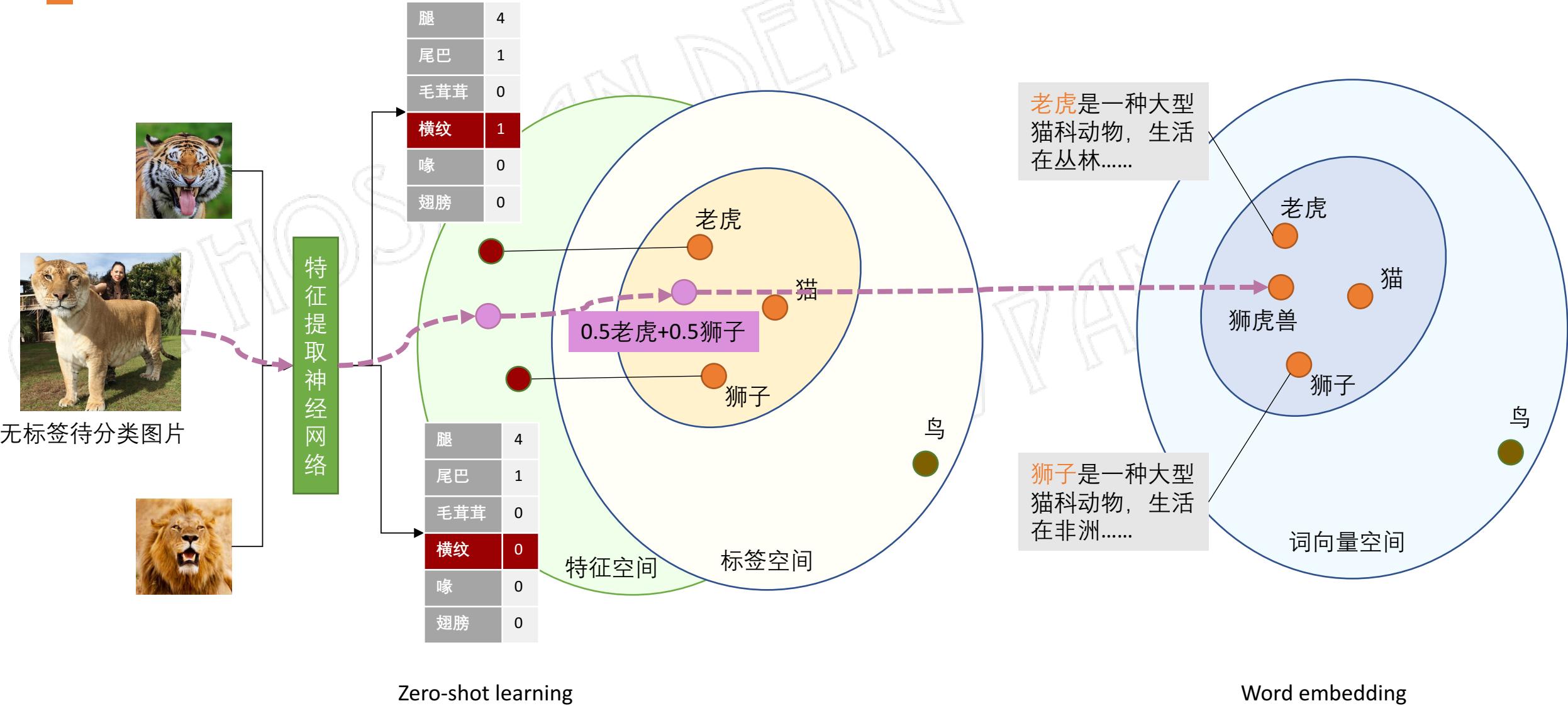


Word embedding

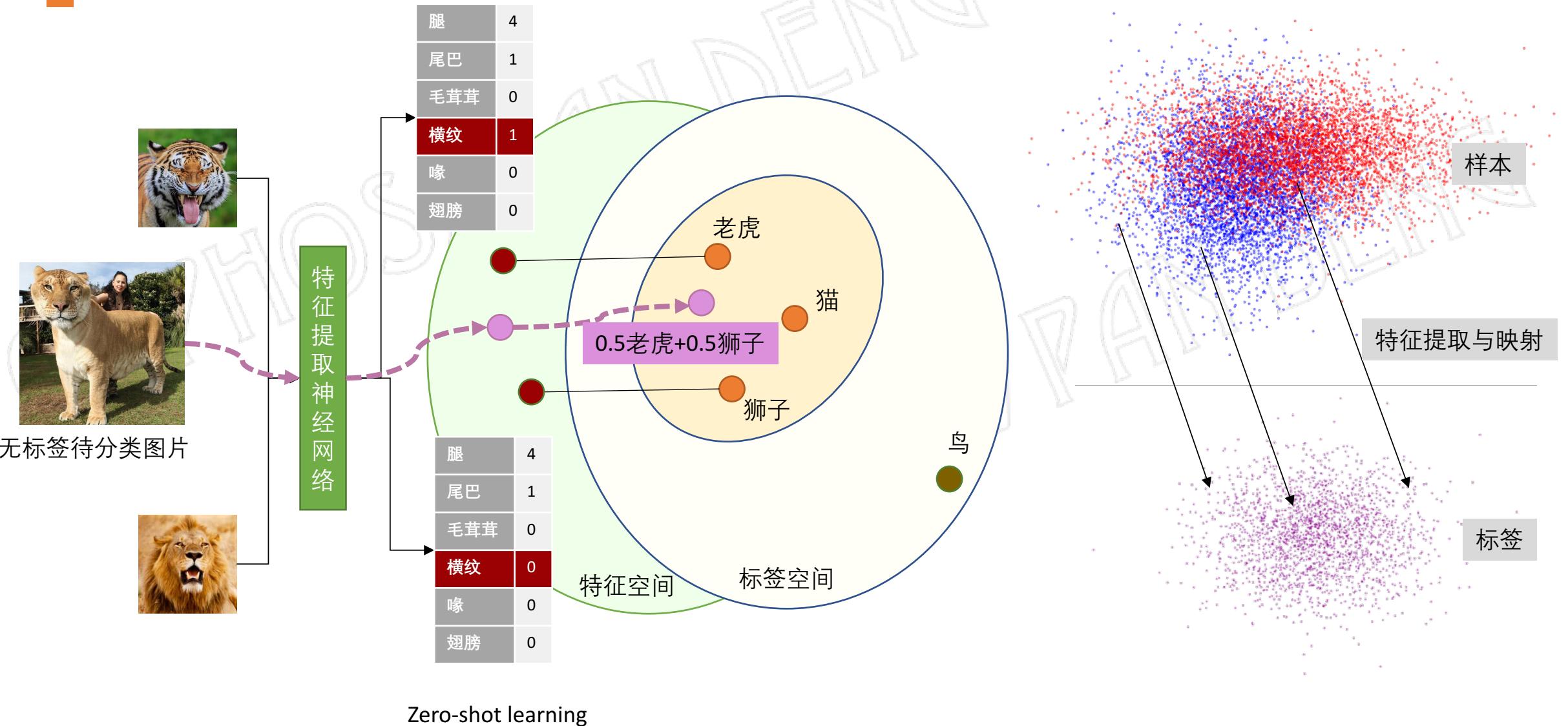
# Zero-shot learning : 特征embedding, 建立特征->标签的映射函数



# Zero-shot learning : 特征embedding, 建立特征->标签的映射函数



# Zero-shot learning : 特征embedding, 建立特征->标签的映射函数



# 自我学习

## Self-taught Learning: Transfer Learning from Unlabeled Data

Rajat Raina  
Alexis Battle  
Honglak Lee  
Benjamin Packer  
Andrew Y. Ng  
Computer Science Department, Stanford University, CA 94305 USA

RAJATR@CS.STANFORD.EDU  
AJBATTLE@CS.STANFORD.EDU  
HILLEE@CS.STANFORD.EDU  
BPACKER@CS.STANFORD.EDU  
ANG@CS.STANFORD.EDU

自我学习不要求未标注数据和已标注数据有同样的分布，而半监督学习要求未标注数据和已标注数据服从同样的分布。

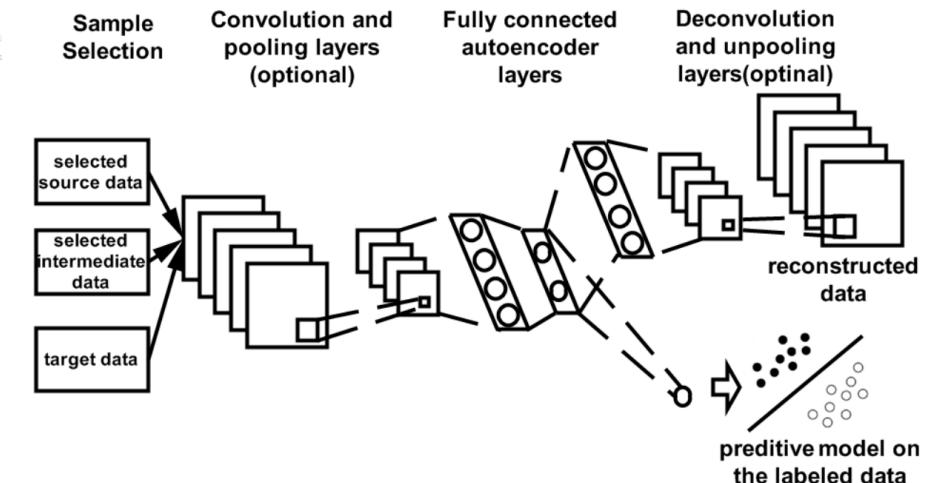
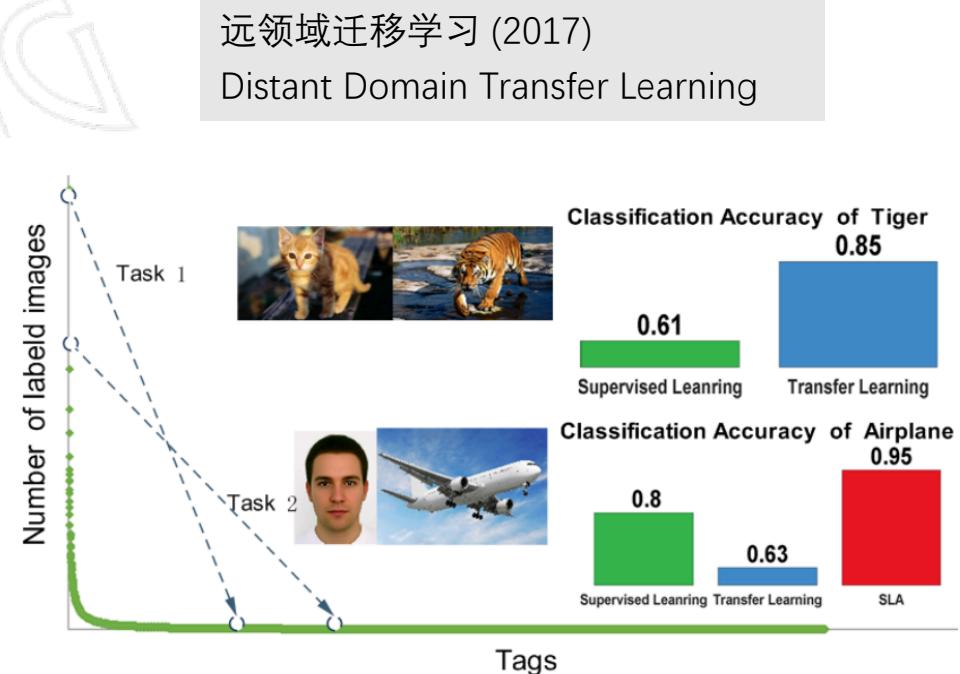
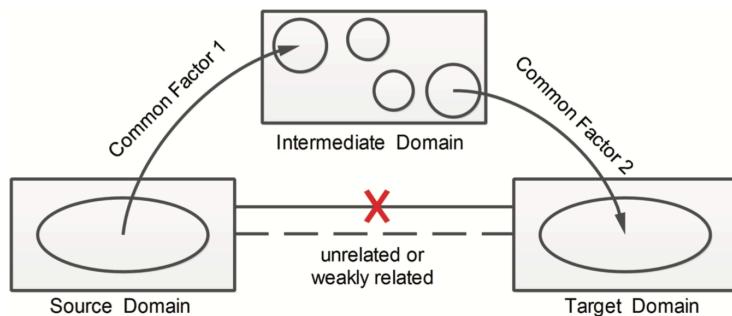
可通过autoencoder等实现

|        | 源域 & 目标域 | 源任务 & 目标任务 | 数据标签 |         | 任务方法  |
|--------|----------|------------|------|---------|---|
|        |          |            | 源数据  | 目标数据    |   |
| 传统机器学习 | 相同       | 相同         | 有标签  | 有标签     | 验证  |
|        |          |            |      | 无标签     | 测试  |
|        |          |            | 无标签  | 无标签     | 聚类  |
| 迁移学习   | 相同/相关    | 相关         | 有标签  | 有标签     | 多任务学习<br>fine-tune<br>自适应<br>TrAdaBoost           |
|        | 相关       | 相同/相关      |      |         |   |
|        | 相同/相关    | 相关         | 无标签  | 无标签/有标签 | Domain-adversarial training<br>Zero-shot Learning |

# 迁移学习的可拓展性：“摸着石头过河”

利用处于（不相似的）源域和目标域之间的若干领域，完成传递式  
的知识迁移

传递迁移学习 (2015)  
Transitive Transfer Learning



谢谢！