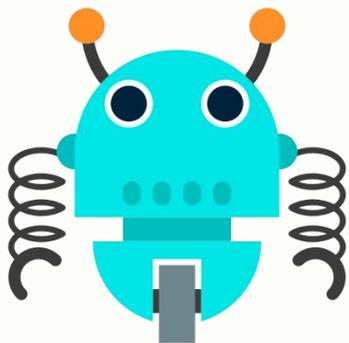


# 第08讲 迁移学习技术

欧新宇



# 迁移学习技术



- 迁移学习概述 (动机、定义、分类)
- 迁移学习的微调策略 (预训练模型、微调技术)
- 迁移学习在视觉任务中的应用
- 小结





---

# 迁移学习概述

---

**“如果需要训练/使用CNN，你需要有大量的训练数据。”**

破下

## 迁移学习的动机

**利用在相关任务上训练好的模型来帮助当前的任务。**

### 方法

- 特征抽取器 (Word2Vec, ResNet50特征, I3D特征)
- 在相关任务上训练模型并直接进行重用
- 在预训练模型上进行微调

### 相关领域

- 半监督学习
- 极致情况: Zero-Shot/Few-Shot Learning
- 多任务学习, 每个任务都有一些带标签的可用数据

# 迁移学习概述



迁移学习



迁移学习



迁移学习



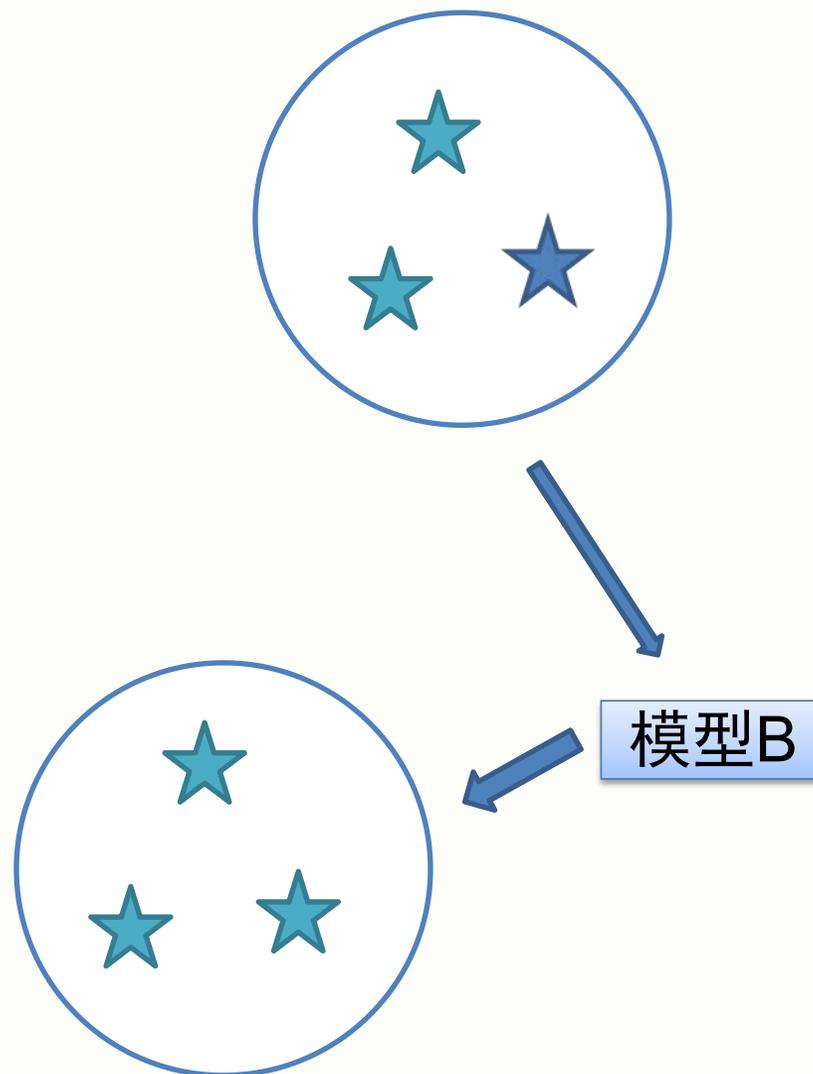
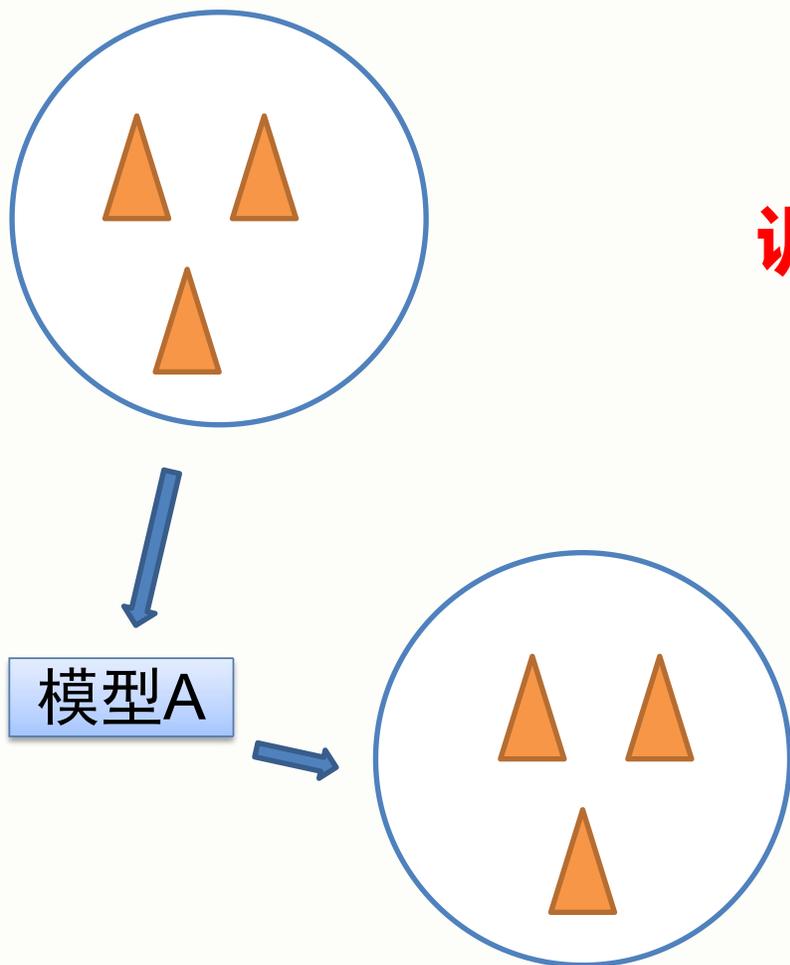
- 婴儿学会爬行，距离学会走路就不远了
- 会骑自行车后，学骑摩托车也就很简单了
- 会打羽毛球，再学打网球也就没那么难了

人类很自然就具备了举一反三的迁移能力。  
计算机又如何获得迁移能力呢？？？

# 迁移学习概述

## 传统的机器学习

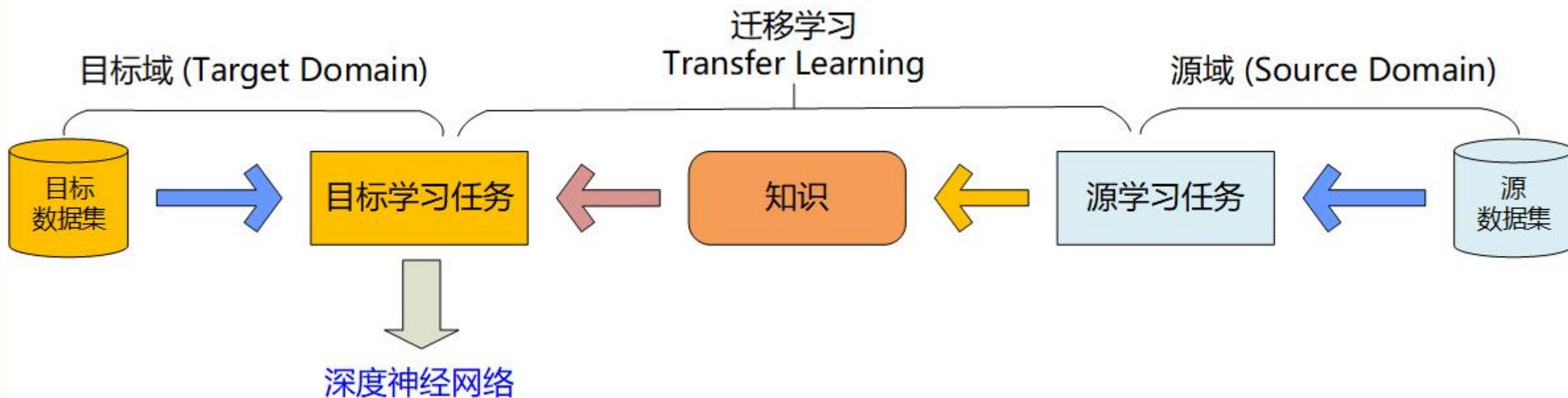
训练和评估都在同一个任务或域中



## 迁移学习的定义

## 迁移学习 (Transfer Learning)

把一个领域 (源领域, Source Domain) 的知识迁移到另一个领域 (目标领域, Target Domain), 使得目标领域能够获得更好的学习效果。



## 迁移学习的分类

## 特征迁移

- 在源领域学习特征表示，把学到的知识以特征的形式进行编码，并从源领域传递到目标领域，提升其任务效果。

## 样本迁移

- 源领域中数据的某一部分可以通过调整权重的方法重用，用于目标领域的学习。

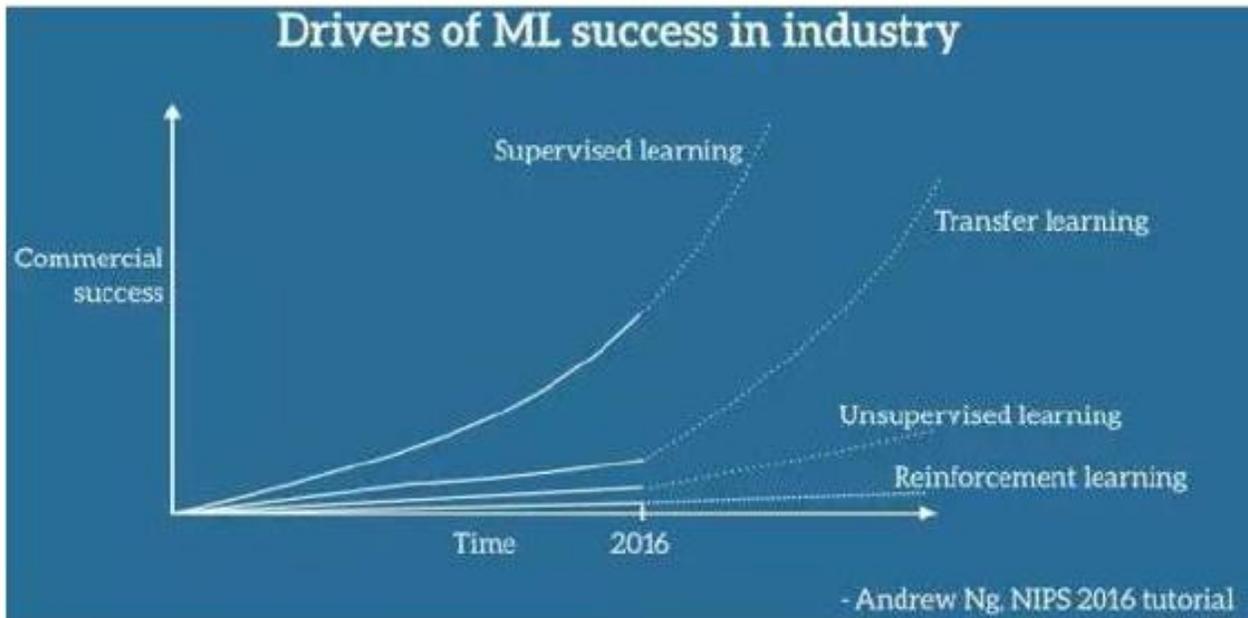
## 模型迁移

- 任务之间共享相同的模型参数或者是服从相同的先验分布。

## 关系知识迁移

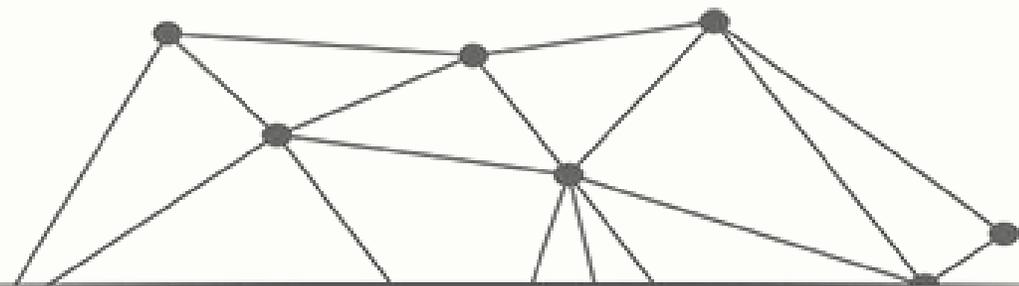
- 相关领域之间的知识迁移。

## 迁移学习的研究意义

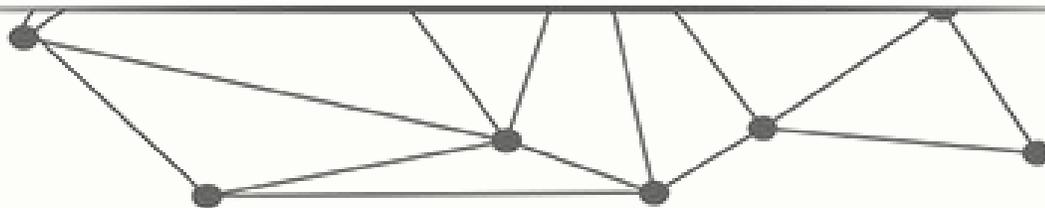


在NIPS2016上，吴恩达表示：“在监督学习之后，迁移学习将引领下一波机器学习技术商业化浪潮。”

- 目前，大多数成功的工作都依赖于大量有标签数据，复用现有知识或数据，可以使已有的大量工作不至于完全丢弃；
- 不需要花费巨大代价去重新采集和标注新数据集，很多学习任务很难获得大量有标签数据
- 对于快速出现的新领域，能够快速迁移和应用，体现时效性的优势



## 课堂互动 7.5.1





---

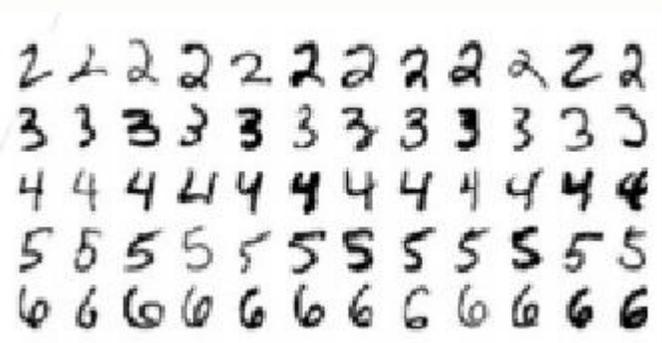
# 迁移学习的微调策略

---

# 迁移学习的微调策略

## 计算机视觉中的知识迁移

在计算机视觉领域中存在很多大量规模 (Large-Scale) 的标注数据集，特别是在图像分类任务中。**知识迁移**可以实现从这些大规模的数据集中获取**有用的信息**到目标CV任务中，目标任务通常比这些大规模数据集小10-100倍。

	ImageNet	CUB_200_2011	Mnist
			
# images	1.2M	11788	6000
# classes	1000	200	10

# 迁移学习的微调策略

## 预训练模型

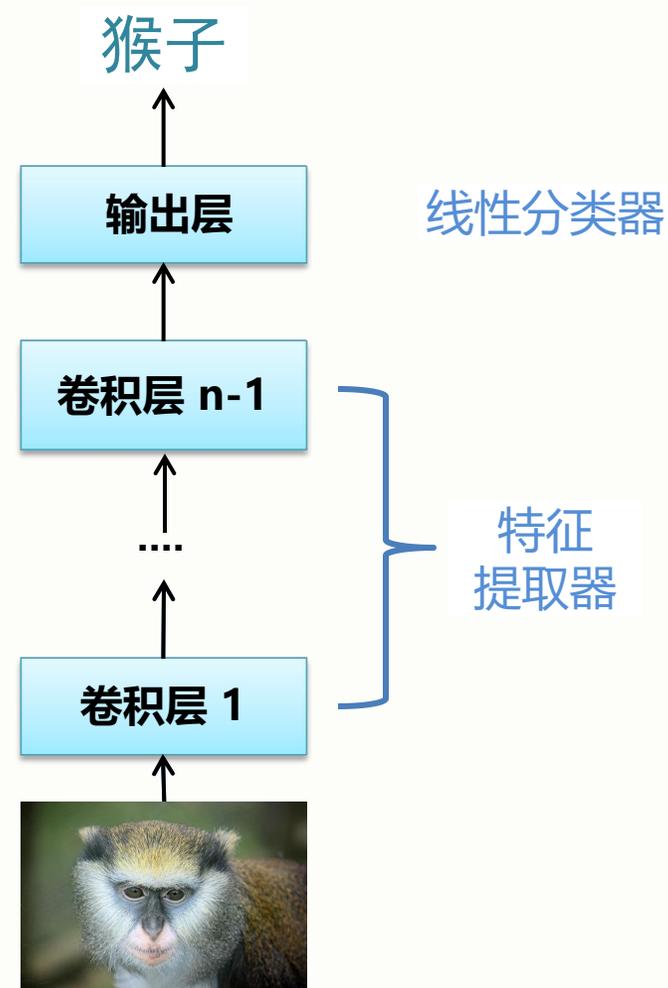
### 神经网络的逻辑构成

- 一个用于将原始像素映射成线性可分离特征的特征提取器 (编码器Encoder)
- 一个用于决策的线性分类器 (解码器Decoder)

### 预训练模型 (Pre-trained Model)

一个在大规模且足够通用的数据集上预先训练好的神经网络模型。该模型的特征提取器能够很好地被应用在：

- ✓ 其他数据集：行人识别、医学图像数据集、卫星图像
- ✓ 其他任务：目标检测、图像分割、图像搜索



# 迁移学习的微调策略

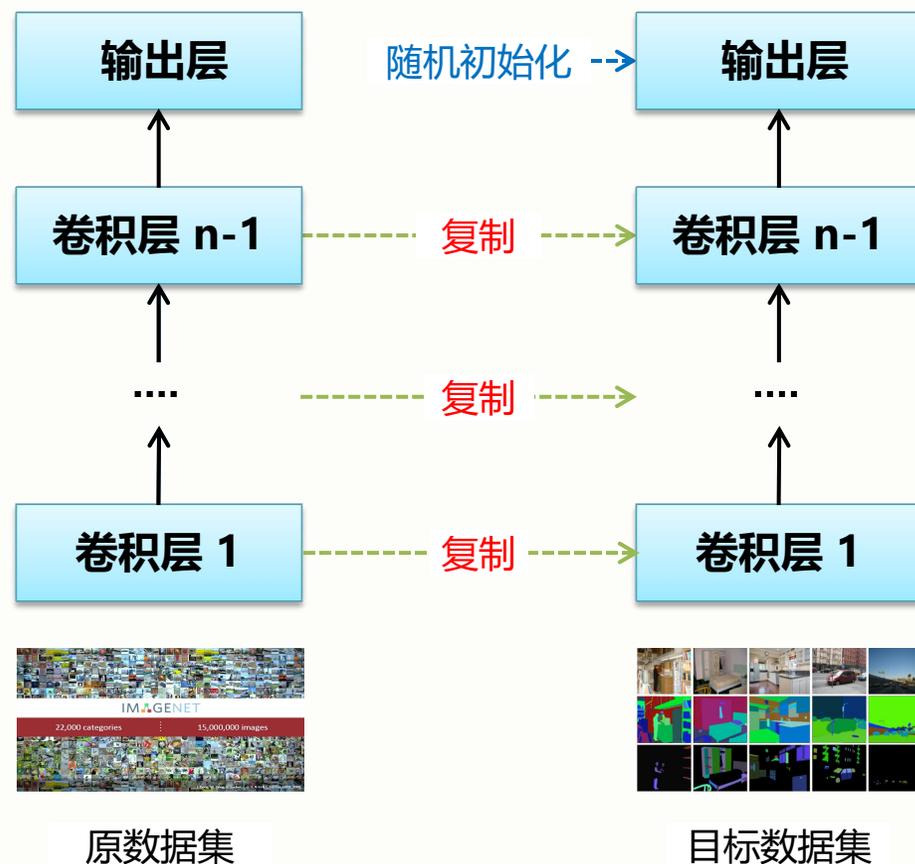
## 微调技术(Fine-Tuning techniques)

### 1. 初始化新模型

- 使用预训练模型的特征提取器参数**初始化**新模型的特征提取器
- 随机初始化输出层
- 启动训练，实现参数的优化

### 2. 使用**较小的学习率**和**较少的周期**进行训练

- 目的是规范搜索空间，让其不要太大
  - ✓ 更小的学习率
  - ✓ 更少的学习周期



# 迁移学习的微调策略

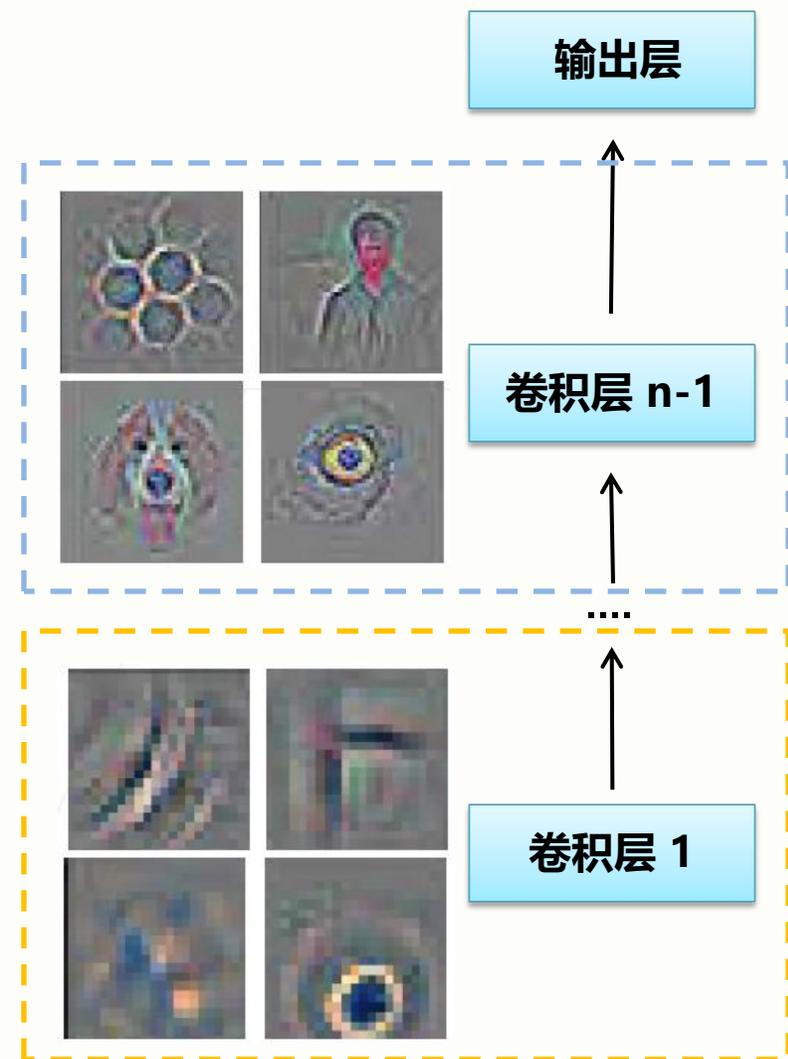
## 冻结底层 (Freeze Bottom Layers)

神经网络学习层次化的特征

- **低级特征**是通用的，泛化性能好，例如：曲线、边缘、气泡；
- **高级特征**更具任务型和数据集偏向性，例如：类别标签。

在微调过程中，冻结底层，并从头训练顶层

- 保持低级通用特征的完整（权重不更新，学习率为0）
- 聚焦于学习任务独特的特征



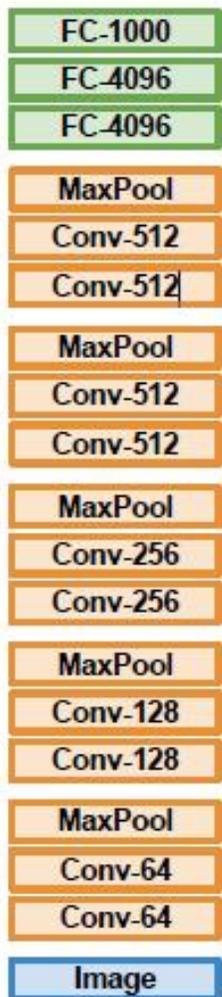
# 迁移学习的微调策略

## 微调策略的选择

数据集  
规模

冻结层

微调层



更特别

更通用

数据集  
相似度

	非常相似的数据集	非常不同的数据集
小数据集	冻结整个卷积部分，微调线性分类器 	1. 微调大部分层，冻结其他层 2. 可能遇到问题，数据增广和收集更多数据会更有效 
大数据集	仅微调小部分层，冻结其他层 	微调整个模型 



# 迁移学习在视觉任务中的应用

# 迁移学习在视觉任务中的应用

## 对于具体项目来说...

Transfer learning 就像:



# 迁移学习在视觉任务中的应用

## 去哪找预训练模型?

深度学习框架通常都会提供预训练模型库 “**Model zoo**” 。

因此，我们不需要自己进行模型预训练。

- ✓ **TensorFlow**: <https://tfhub.dev>
- ✓ **PyTorch**: <https://github.com/pytorch/vision>
- ✓ **PaddlePaddle**: <https://www.paddlepaddle.org.cn/modelbase>

### ● 常见的预训练模型

- Alexnet模型
- 牛津VGG模型
- 谷歌Inception模型
- 微软ResNet模型

```
import paddle
```

```
# 设置输入样本的维度
```

```
input_spec = InputSpec(shape=[None] + train_parameters['input_size'], dtype='float32',  
name='image')
```

```
label_spec = InputSpec(shape=[None, 1], dtype='int64', name='label')
```

```
# 载入官方标准模型, 若不存在则会自动进行下载, pretrained=True/False控制是否使用Imagenet预训练参数
```

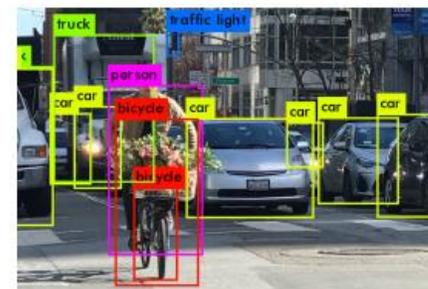
```
network = paddle.vision.models.mobilenet_v2(num_classes=20, pretrained=True)
```

```
model = paddle.Model(network, input_spec, label_spec)
```

# 迁移学习在视觉任务中的应用

## 精度保持与加速收敛

- 微调预训练模型 (on ImageNet) 是计算机视觉的一种流行方法, 它允许以节省时间的方式建立精确的模型, 被广泛应用于在不同的CV应用中
  - ✓ 检测Detection/分割Segmentation (图像相似, 但是目标不同)
  - ✓ 医学图像/卫星图像 (相同的任务, 但是非常不同的图像)
- 微调加速收敛
- 微调并不能总是保证提高精度
  - ✓ 当目标数据集比较大的时候, 从头开始训练也可以得到相似的精度。



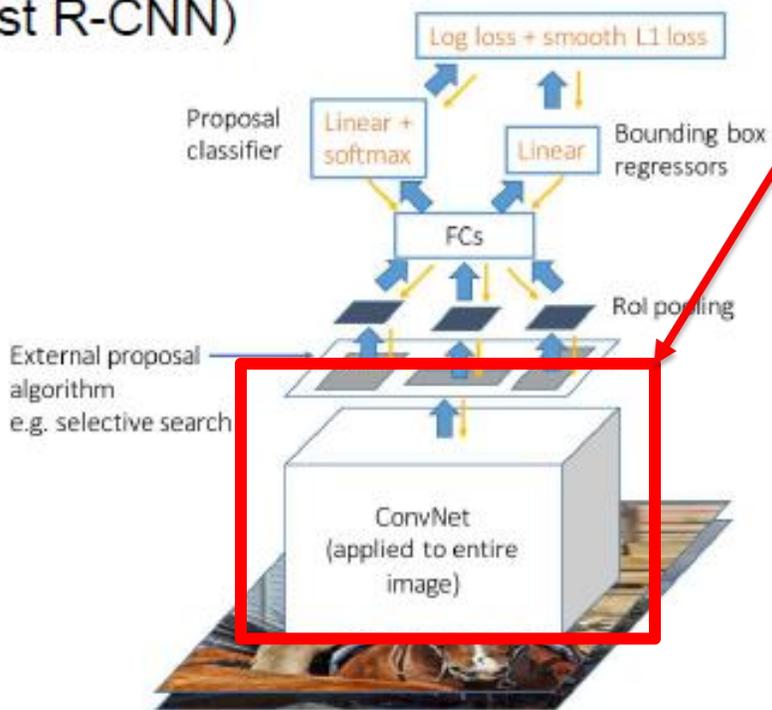
图像来源于李沐CS329

# 迁移学习在视觉任务中的应用

## 从分类到检测，再到图像标注

### 目标检测

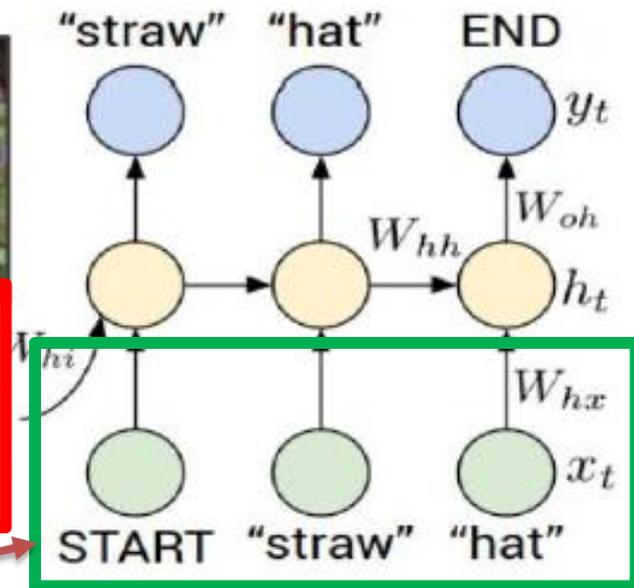
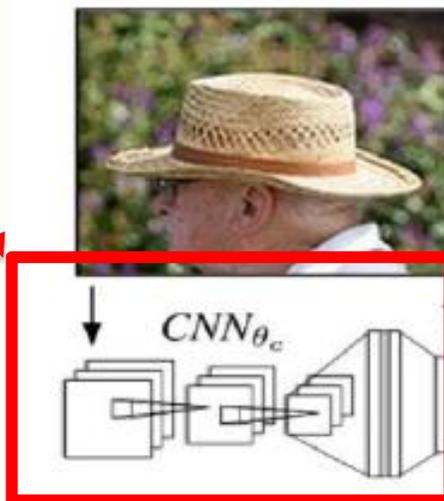
Object Detection (Fast R-CNN)



基于ImageNet  
预训练的CNN

### 图像标注 (加字幕)

Image Captioning: CNN + RNN



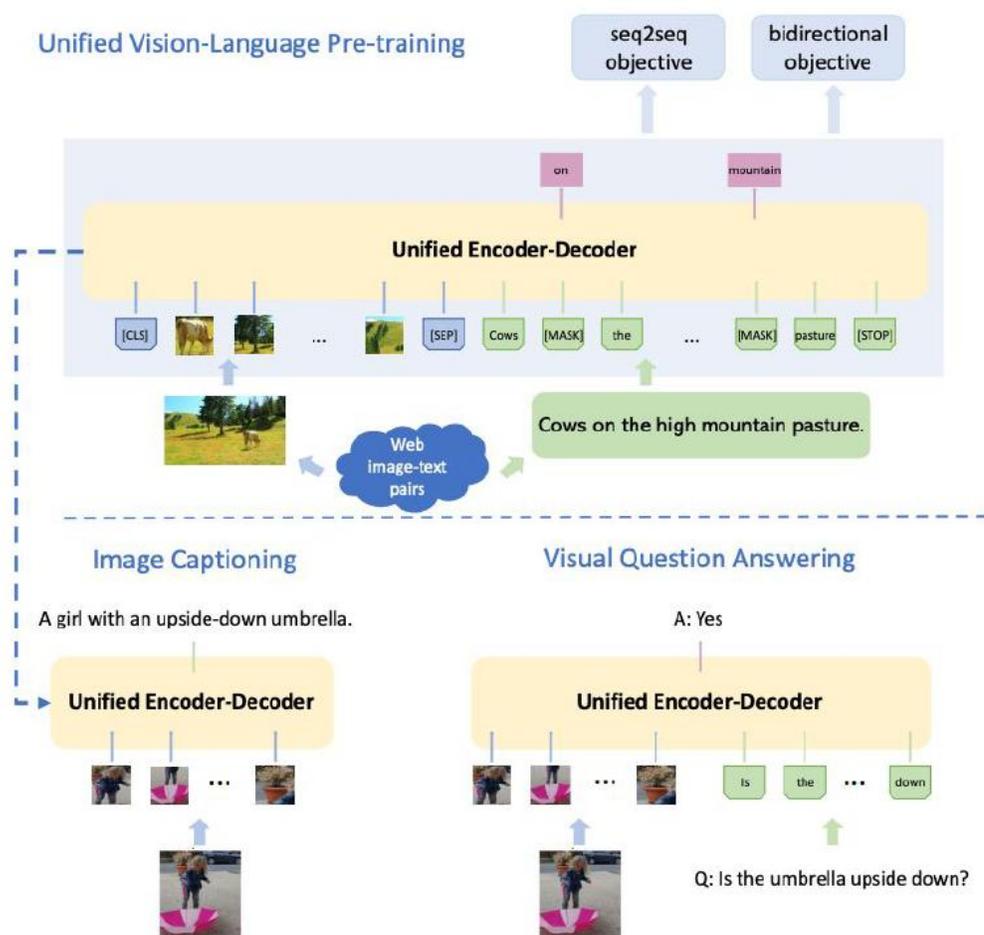
使用word2vec  
预训练的词向量

Girshick, "Fast R-CNN", ICCV 2015  
Figure copyright Ross Girshick, 2015. Reproduced with permission.

Karpathy and Fei-Fei, "Deep Visual-Semantic Alignments for Generating Image Descriptions", CVPR 2015  
Figure copyright IEEE, 2015. Reproduced for educational purposes.

# 迁移学习在视觉任务中的应用

## 从CNN到视觉问答



Zhou et al, "Unified Vision-Language Pre-Training for Image Captioning and VQA" CVPR 2020  
Figure copyright Luwei Zhou, 2020. Reproduced with permission.

1. 在ImageNet数据集上进行CNN预训练
2. 在Visual Genome数据集上执行目标检测的微调训练
3. 在大量文本上训练BERT语言模型
4. 组合 (2) 和 (3) 的模型, 并进行图像/语言模型的联合训练
5. 基于模型 (4) 进行微调训练, 应用在图像标注, 视觉问答等问题。

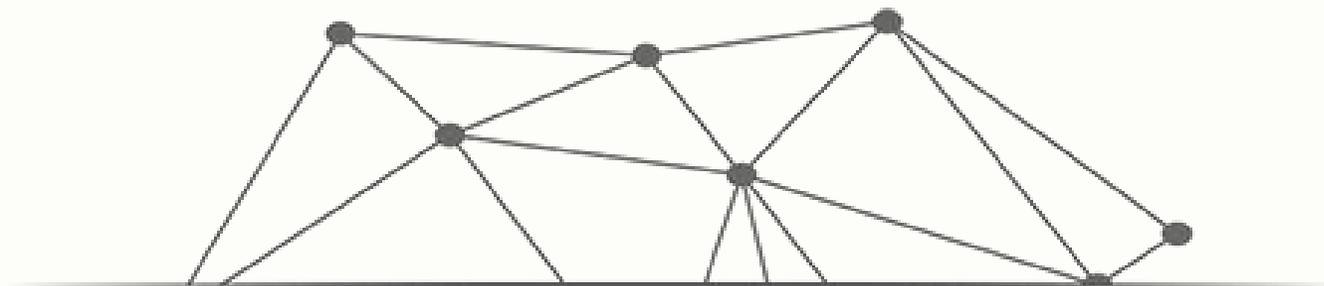
## 小 结

- 找一个与任务数据集相似的，并且非常大的数据集，然后训练一个大型的卷积神经网络。通常，是直接找一个预训练模型，例如在ImageNet图像分类数据集上与训练好的模型。
- 使用预训练模型来初始化目标任务的权重参数
- 在任务数据集上进行微调训练。微调能够加速训练，通常也能改进精度。

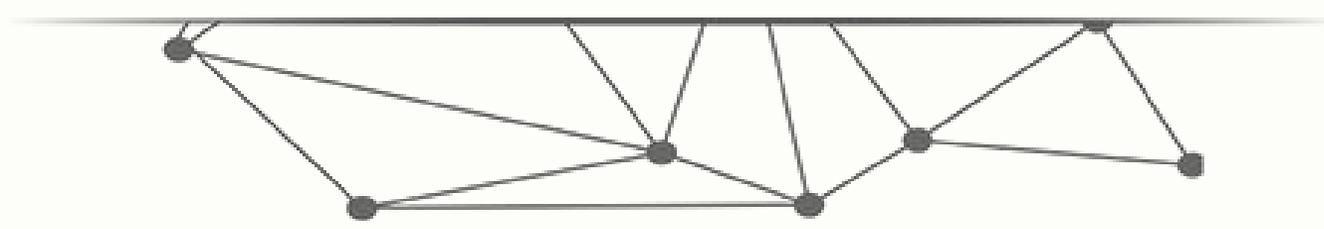
“事上磨”



“知行合一”



## 课堂互动 7.5.2



读万卷书 行万里路 只为最好的修炼



QQ: 14777591 (宇宙骑士)

Email: [ouxinyu@alumni.hust.edu.cn](mailto:ouxinyu@alumni.hust.edu.cn)

Website: <http://ouxinyu.cn>

Tel: 18687840023

地址: 安宁校区 诚远楼201

南院 智能应用研究院A306-2