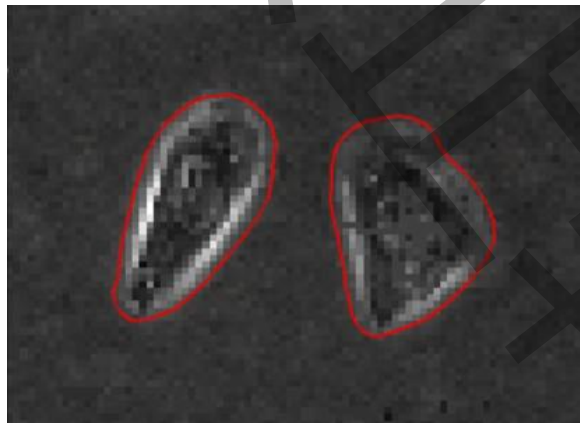
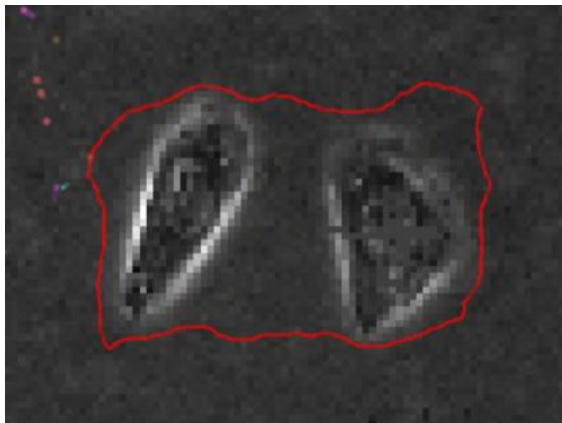


# 图像分割

# 1.基本概念

## 什么是图像分割

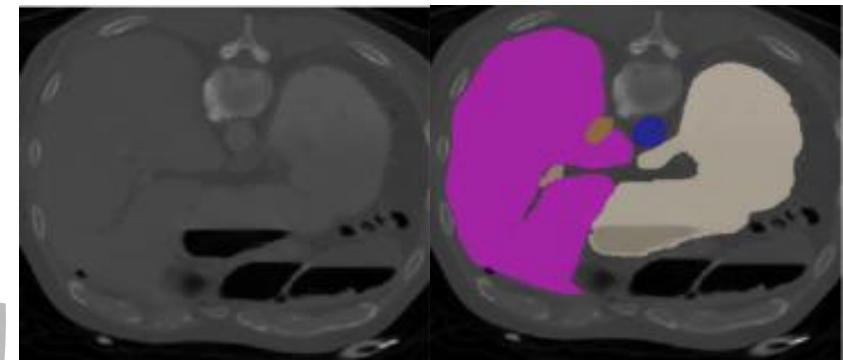
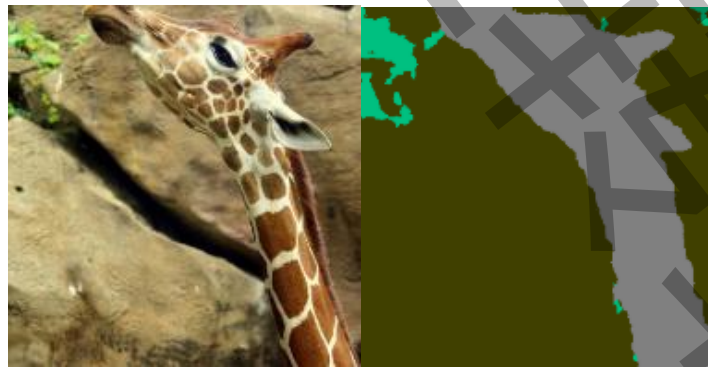
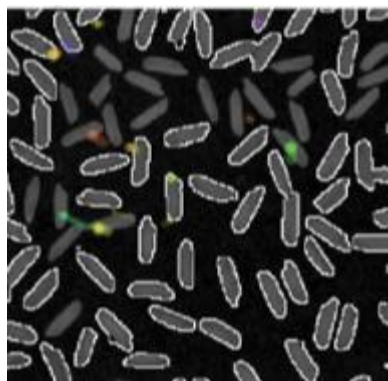
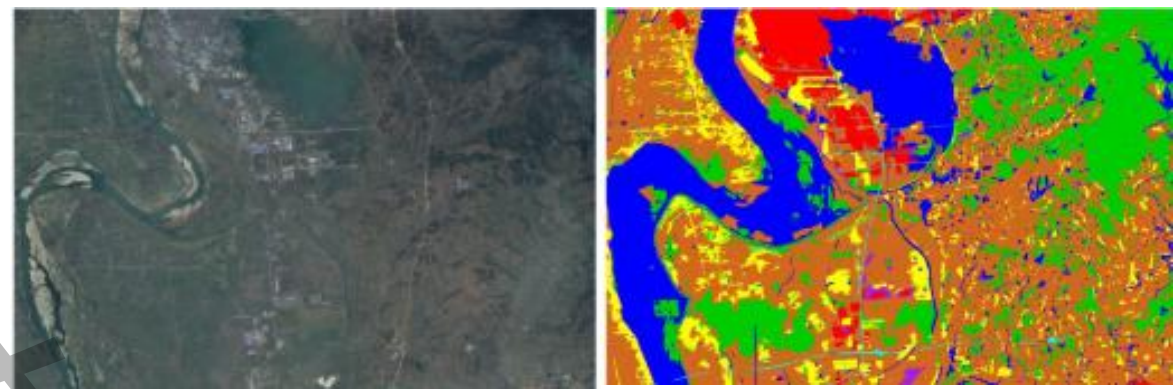
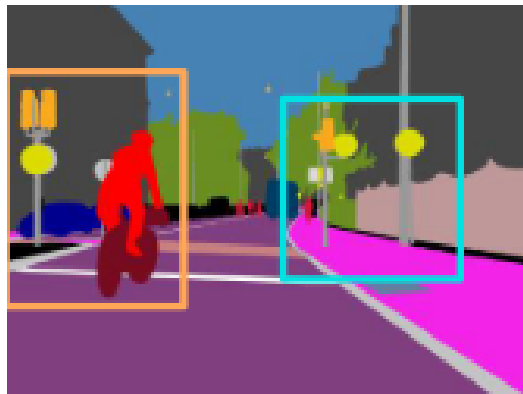
- 预测目标的轮廓



将不同的像素划分到不同的类别(class), 本质上是分类的问题, 但是对像素进行分类

## 图像分割的应用场景

➤ 人像抠图、医学组织提取、遥感分析、自动驾驶等



医学图像分割：医学影像、CT照片等

## 图像分割的前景和背景

- 物体Things: 可数的前景目标 (行人、汽车)
- 背景Stuff: 不可数背景 (天空、树林, 路面、草地)

*Foreground*

*Background*

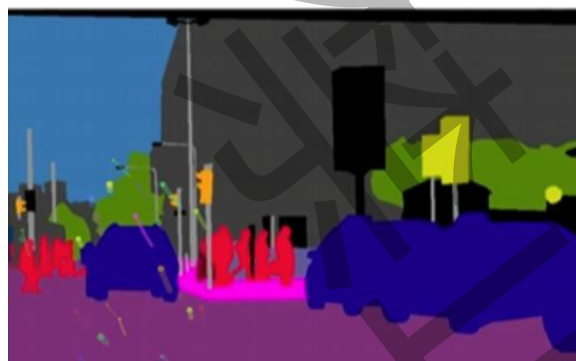




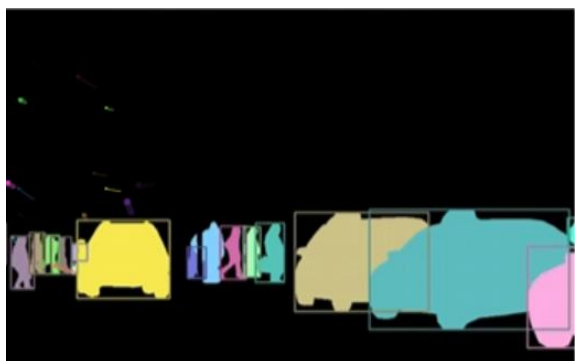
## 图像分割的三个层次



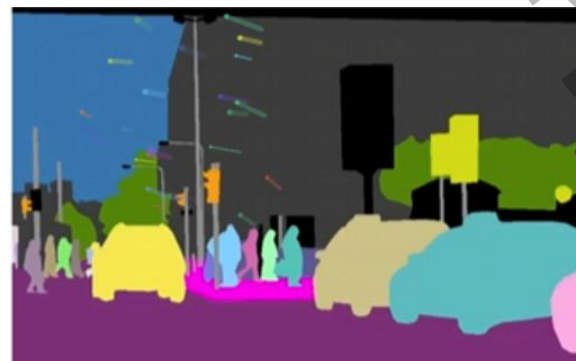
(a)原图像



(b)语义分割



(c)实例分割



(c)全景分割

### ➤ 语义分割 (Semantic Segmentation)

每一个像素只能**属于一类**，预测结果为掩膜

### ➤ 实例分割 (Instance Segmentation)

只预测**前景目标**的类别属性及边框，个体ID，  
每一个像素可以**属于多个ID**

### ➤ 全景分割 (Panoptic Segmentation)

简单理解为**语义分割+实例分割**

## 2.语义分割

### 计算机视觉的任务

分类



猫

图片级

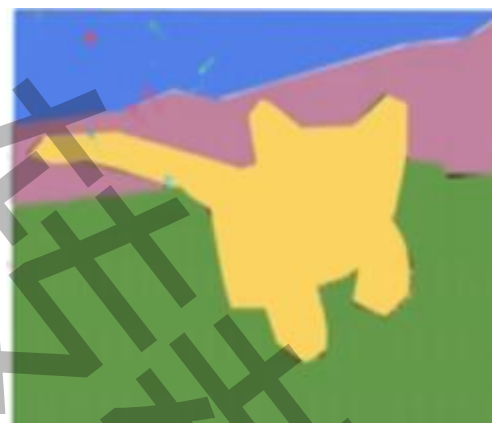
目标检测



猫、狗

区域级

语义分割



猫、天空、草地、树林

像素级

实例分割



猫、狗、狗

挑战

### 3.数据集

- PASCAL Visual Object Classes (VOC项目) , 20类图像, 1类背景, 总共9963张图片, 24640个标注目标



- 从2007年开始引入语义分割和实例分割的标注图像



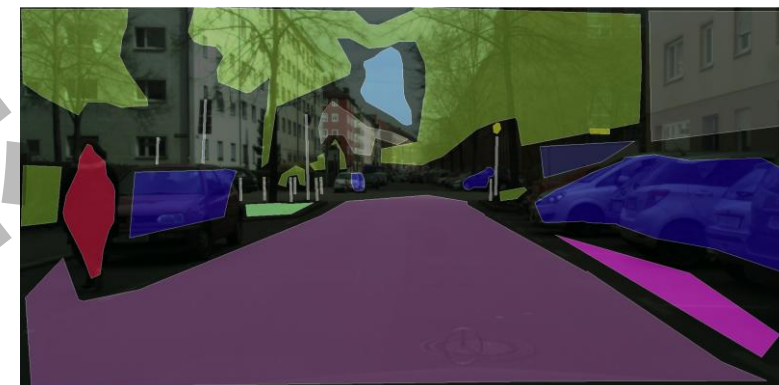
Everingham, M.; and Winn, J. 2012. The PASCAL visual object classes challenge 2012 (VOC2012) development kit. Pattern Anal. Stat. Model. Comput. Learn., Tech. Rep, 2007(1-45): 5.



➤ Cityscapes (用于自动驾驶场景)

50个城市在春夏秋三个季节不同时段，不同场景的街景图

500张精细标注的图像，20000张粗略标注的图像



Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benen-son, R., Franke, U., Roth, S., & Schiele, B. (2016). The cityscapes dataset for semantic urban scene understanding. In IEEE conference on computer vision and pattern recognition (pp. 3213–3223)



## ➤ COCO数据集

Common Objects in Context, 以场景理解为目标, 选取了比较复杂的日常场景

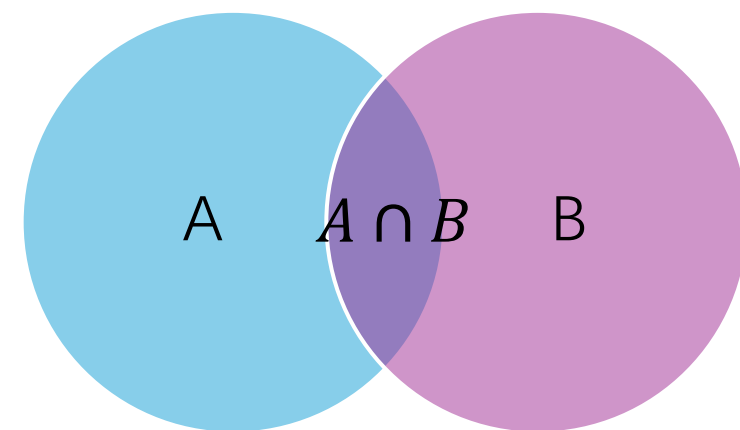


一共有91类, 以人类4岁小孩能够辨识为基准, 其中82类超过5000个实例

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft Coco: Common Objects in Context. In European conference on computer vision (pp. 740–755). Springer.

## 4.评估指标

- Pixel Accuracy: 逐像素分类精度
- Mean Pixel Accuracy: 每个类被正确分类像素的比例
- IoU: 前景目标交并比
- mIoU: 每个类IoU的平均值
- .....



## 5.语义分割常用优化目标

- 交叉熵损失,  $z_k$ 表示网络的输出,  $f(z_k)$ 表示概率

$$f(z_k) = \frac{e^{z_k}}{\sum_j e^{z_j}}, \quad l(y, z) = - \sum_{k=0}^C y_c \log(f(z_k)), \quad l(y, z) = \log \sum_j e^{z_j} - \log e^{z_y}$$

分割损失即所有像素分类损失的累加

- 交叉熵损失的特点

$$\frac{\partial l(y, z)}{\partial z_k} = \begin{cases} f(z_k) - 1 & \text{当 } y = k \\ f(z_k) & \text{其他} \end{cases}$$

可用于多类别, 正负样本梯度稳定



## 5.语义分割常用优化目标

- Dice损失, P、G分别表示预测和真实值

定义Dice相似度:  $S = 2 \frac{|P \cap G|}{|P| + |G|}$        $S_1 = \frac{2 \sum_N p_i g_i}{\sum_N p_i^2 + \sum_N g_i^2}$        $S_2 = \frac{2 \sum_N p_i g_i}{\sum_N p_i + \sum_N g_i}$

Dice损失:  $D(p, g) = 1 - S$

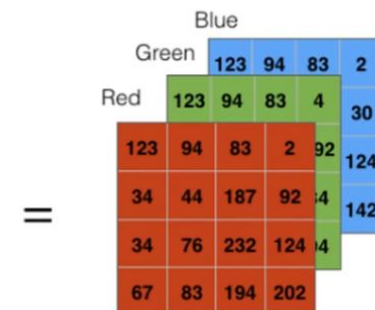
- Dice损失的特点

导数:  $\frac{\partial S}{\partial p_j} = 2 \frac{g_j (\sum_N g_i^2 - \sum_N p_i^2)}{(\sum_N p_i^2 + \sum_N g_i^2)^2}$        $g_i = 0$ , 梯度值为0

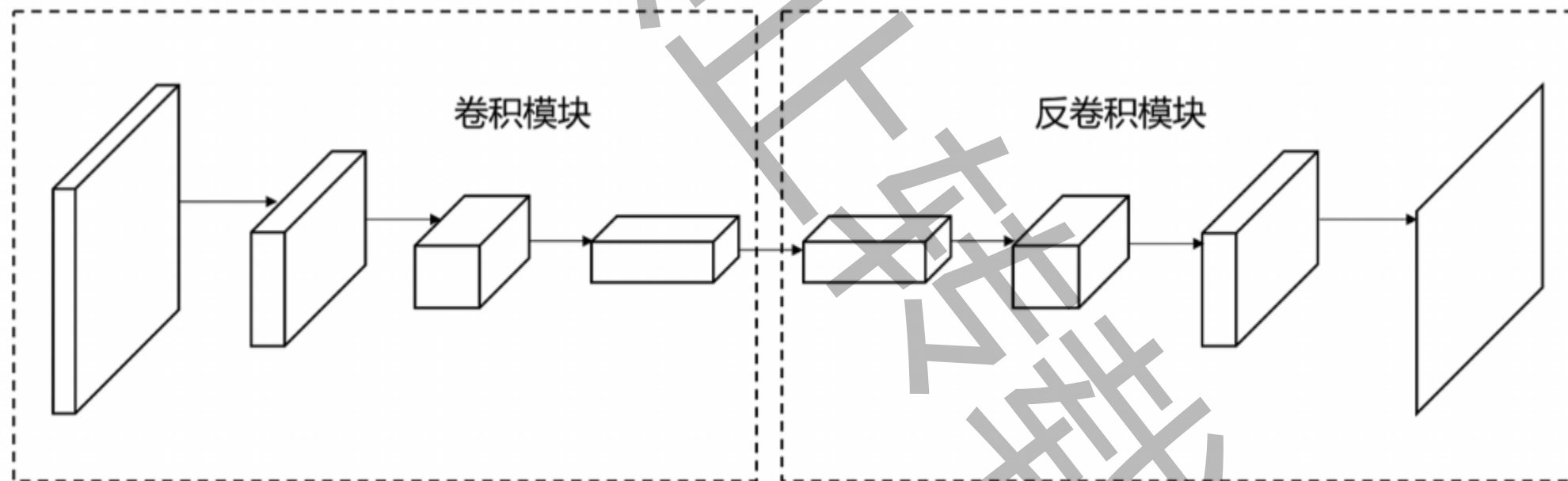
- ◆ 专注正样本, 适合小目标
- ◆ 损失可能并不稳定

## 6.语义分割网络

- 卷积模块：特征提取
- 反卷积模块：上采样恢复原图像尺寸



RGB图像



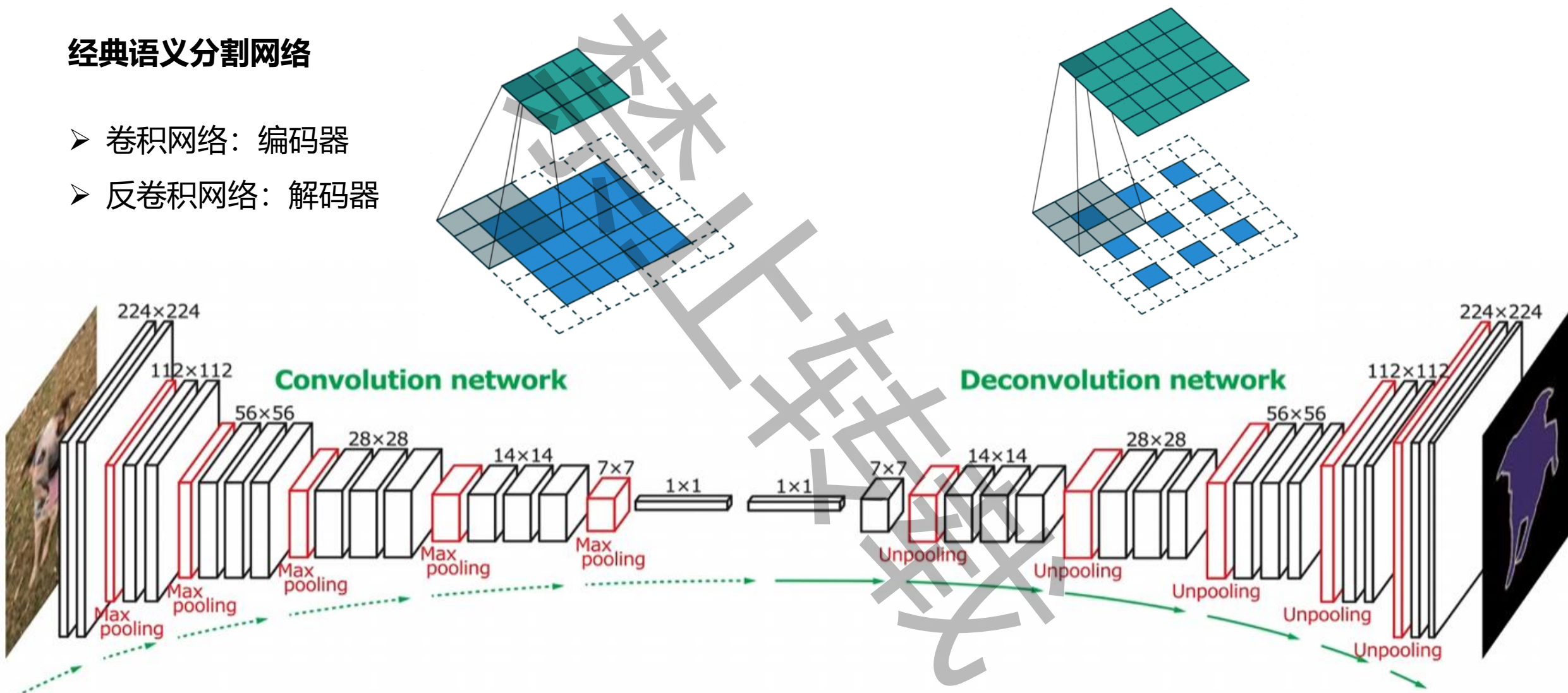
$$Out_H = \frac{H + 2P - FH}{S} + 1, Out_W = \frac{W + 2P - FW}{S} + 1$$

<https://pytorch.org/docs/stable/generated/torch.nn.Conv2d.html#torch.nn.Conv2d>

# 6.语义分割网络

## 经典语义分割网络

- 卷积网络：编码器
- 反卷积网络：解码器



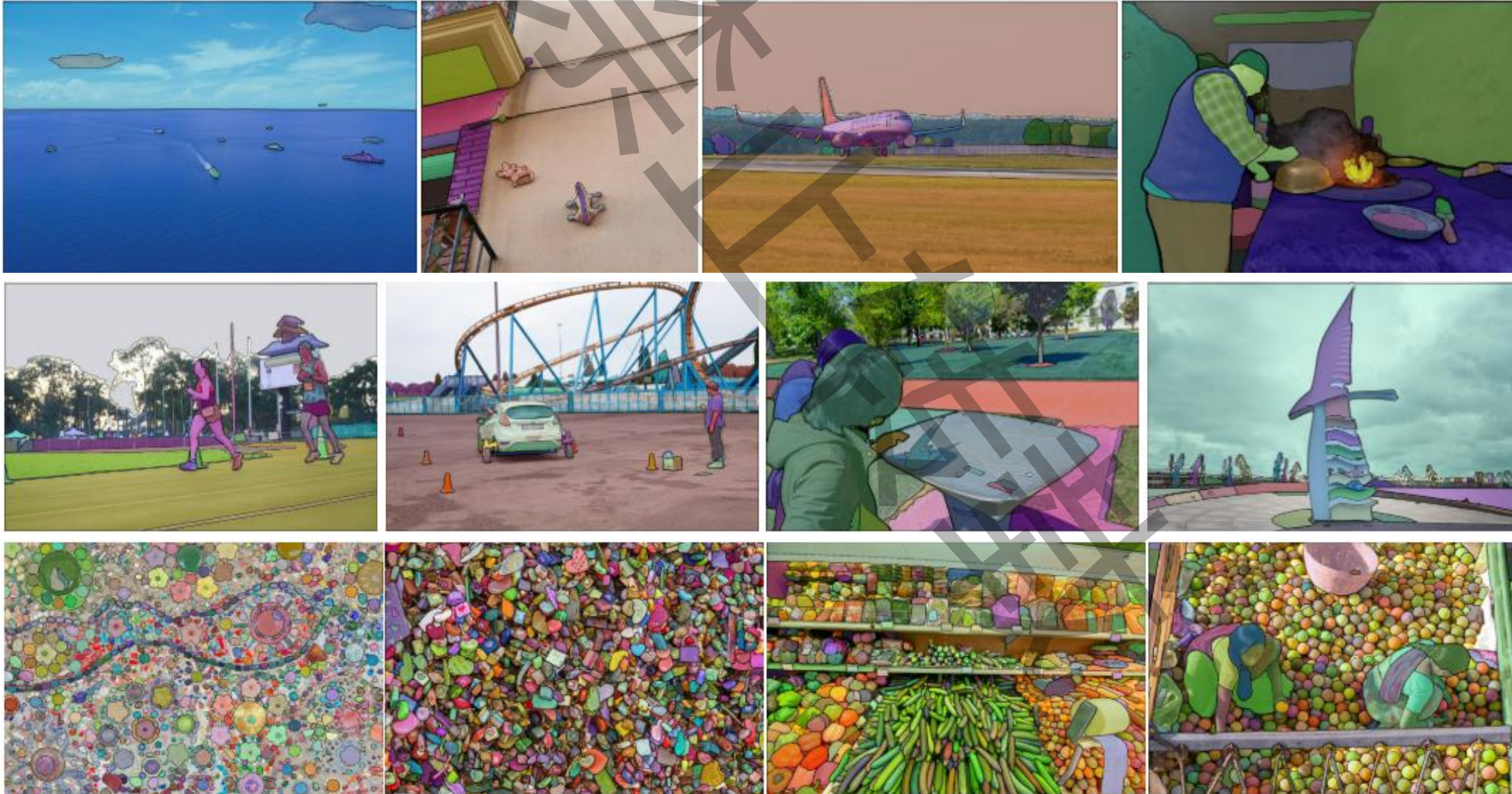
反卷积是一种**特殊的正向卷积**，先按照一定的比例通过补 0 来扩大输入图像的尺寸，再进行普通的卷积。





# 7.Segment Anything

Meta 提出的一个通用分割模型，包含超过**10亿**个分割掩膜和**1100万**张符合许可且尊重隐私的图像<sup>[1]</sup>

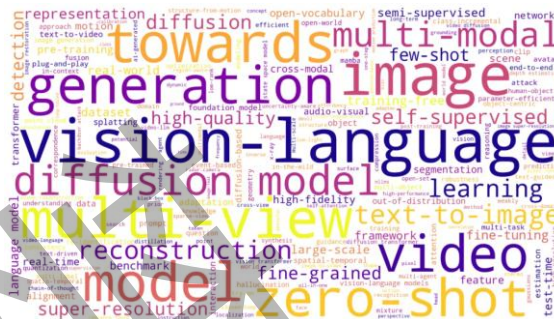




# 7.Segment Anything

## Demo





图像分割是计算机视觉中重要的基础任务，广泛被应用于医学图像处理、自动驾驶、遥感分析等领域，但当前的研究面临的主要问题：**标注大量的数据集耗费人力和财力，模型的泛化能力比较差**。针对上述问题，有研究者采用Few-shot Semantic Segmentation或Zero-shot Semantic Segmentation等方法，尝试在少样本的训练集上实现模型的泛化能力，随着Transformer的推广，其也被应用到图像分割当中。

在未来的研究工作中，Few-shot Semantic Segmentation或Zero-shot Semantic Segmentation有望成功解决上述的两个挑战，推动图像分割的研究进一步发展<sup>[1]</sup>

[1]Catalano N, Matteucci M. Few Shot Semantic Segmentation: a review of methodologies, benchmarks, and open challenges[J]. arXiv preprint arXiv:2304.05832, 2023.



## 近期学习内容 (2.24-3.31)

# 论文阅读（部分）

题目	时间	单位	期刊/会议	关键词	代码
Segment Anything	2023	Meta AI	ICCV	语义分割	<a href="https://github.com/facebookresearch/segment-anything">https://github.com/facebookresearch/segment-anything</a>
Stronger, Fewer, & Superior Harnessing Vision Foundation Models for Domain Generalized Semantic Segmentation (DGSS)	2024	中国科学技术大学、上海人工智能实验室	CVPR	视觉基础模型、域泛化语义分割	<a href="https://github.com/w1oves/Rein">https://github.com/w1oves/Rein</a>
High Quality Segmentation for Ultra High-resolution Images	2022	香港中文大学、Adobe等	CVPR	超高清图像分割	<a href="https://github.com/qqilu/Entity/tree/main/High-Quality-Segmentation">https://github.com/qqilu/Entity/tree/main/High-Quality-Segmentation</a>
Disentangle then Parse: Night-time Semantic Segmentation with Illumination Disentanglement	2023	中国科学技术大学、上海人工智能实验室	ICCV	夜间语义分割	<a href="https://github.com/w1oves/DTP">https://github.com/w1oves/DTP</a>
SED:A Simple Encoder-Decoder for Open-Vocabulary Semantic Segmentation	2024	天津大学，重庆大学等	CVPR	开放词汇语义分割	<a href="https://github.com/xb534/SED">https://github.com/xb534/SED</a>
Prompting Multi-Modal Image Segmentation with Semantic Grouping	2024	中国科学院大学	AAAI	多模态图像分割	-
Prompt-and-Transfer: Dynamic Class-Aware Enhancement for Few-Shot Segmentation	2025	中国科学院、清华大学	TPAMI	少样本分割、语义分割	-

论文阅读（部分）

题目	时间	单位	期刊/会议	关键词	代码
Cross-Domain Few-Shot Semantic Segmentation via Doubly Matching Transformation	2024	南京航空航天大学、西安电子科技大学	IJCAI	少样本语义分割	<a href="https://github.com/ChenJiayi68/DMTNet">https://github.com/ChenJiayi68/DMTNet</a>
Relevant Intrinsic Feature Enhancement Network for Few-Shot Semantic Segmentation	2024	中国科学院大学、中国科学院、阿里巴巴	AAAI	少样本语义分割	<a href="https://github.com/baoxiaoyi/RiFeNet">https://github.com/baoxiaoyi/RiFeNet</a>
Scribble-Supervised Semantic Segmentation with Prototype-based Feature Augmentation	2024	河海大学等	ICML	涂鸦监督语义分割	<a href="https://github.com/TranquilChan/PFA">https://github.com/TranquilChan/PFA</a>
Progressive Feature Self-Reinforcement for Weakly Supervised Semantic Segmentation	2024	之江实验室、西安电子科技大学等	AAAI	弱监督语义分割	<a href="https://github.com/Jessie459/feature-self-reinforcement">https://github.com/Jessie459/feature-self-reinforcement</a>
A Transformer-based Adaptive Prototype Matching Network for Few-Shot Semantic Segmentation	2024	南京信息工程大学、青海师范大学等	IJCAI	少样本语义分割、Transformer	-
Scribble Hides Class: Promoting Scribble-Based Weakly-Supervised Semantic Segmentation with Its Class Label	2024	北京大学、北京大学深圳研究生院	AAAI	基于涂鸦的弱监督语义分割	<a href="https://github.com/Zx119990529/Class-driven-Scribble-Promotion-Network">https://github.com/Zx119990529/Class-driven-Scribble-Promotion-Network</a>



# 代码部分(Pytorch+MMsegmentation)



PyTorch深度学习快速入门教程  
(绝对通俗易懂!)【小土堆】



Pytorch框架与经典卷积神经网络与实战



语义分割与MMSegmentation



Pytorch 搭建自己的DeeplabV3  
+语义分割平台 (Bubbliiiiing 深...)

- [https://www.bilibili.com/video/BV1hE411t7RN/?spm\\_id\\_from=333.1387.favlist.content.click&vd\\_source=d9c11bf42dbabecdf6bb0ea659f7a921](https://www.bilibili.com/video/BV1hE411t7RN/?spm_id_from=333.1387.favlist.content.click&vd_source=d9c11bf42dbabecdf6bb0ea659f7a921)
- [https://www.bilibili.com/video/BV1e34y1M7wR/?spm\\_id\\_from=333.1387.favlist.content.click&vd\\_source=d9c11bf42dbabecdf6bb0ea659f7a921](https://www.bilibili.com/video/BV1e34y1M7wR/?spm_id_from=333.1387.favlist.content.click&vd_source=d9c11bf42dbabecdf6bb0ea659f7a921)
- [https://www.bilibili.com/video/BV1qN411z7hz/?spm\\_id\\_from=333.1387.favlist.content.click&vd\\_source=d9c11bf42dbabecdf6bb0ea659f7a921](https://www.bilibili.com/video/BV1qN411z7hz/?spm_id_from=333.1387.favlist.content.click&vd_source=d9c11bf42dbabecdf6bb0ea659f7a921)
- [https://www.bilibili.com/video/BV173411q7xF/?spm\\_id\\_from=333.1387.favlist.content.click&vd\\_source=d9c11bf42dbabecdf6bb0ea659f7a921](https://www.bilibili.com/video/BV173411q7xF/?spm_id_from=333.1387.favlist.content.click&vd_source=d9c11bf42dbabecdf6bb0ea659f7a921)

**Thank you**