

## 第一周 机器视觉疾病检测

### 计算机视觉在医疗诊断中的应用

1. 皮肤科Dermatology: 皮肤癌诊断
2. 眼科Ophthalmology: 糖尿病视网膜病变
3. 组织病理学Histopathology: 确定癌症蔓延的程度。由于输入图片太大，需要切分成大量小图片

### 如何处理不平衡分类以及小数据集

1. 类别不均衡：加权损失，重采样

$$L(X, y) = \begin{cases} w_p \times -\log P(Y = 1|X) & \text{if } y = 1 \\ w_n \times -\log P(Y = 0|X) & \text{if } y = 0 \end{cases}$$

$$w_p = \frac{\text{num negative}}{\text{num total}} \quad w_n = \frac{\text{num positive}}{\text{num total}}$$

#### Weighted Loss

2. 多任务：多标签/多任务损失

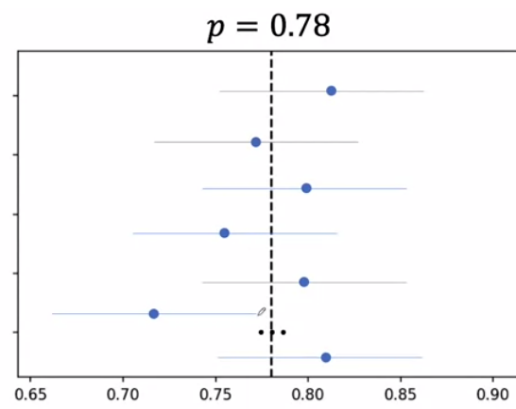
$$\text{Multi-Task} \quad L(X, y_{\text{mass}}) + L(X, y_{\text{pneumonia}}) + L(X, y_{\text{edema}})$$

3. 数据集大小: 迁移学习+数据增强

### 模型性能评估

1. 数据集划分: 训练，验证，测试
2. 三个挑战
  - a. 患者重叠Patient Overlap。解决方法: 保证病人仅仅出现在3个数据集中的某一个
  - b. 数据集采样Set Sampling。解决方法: 按比例采样
  - c. Ground Truth, 观察者意见不统一。解决方法: 协商一致投票; 更多测试 ( 进行CT扫描 )

- ### 合理解释置信区间



**Interpretation of 95% confidence**

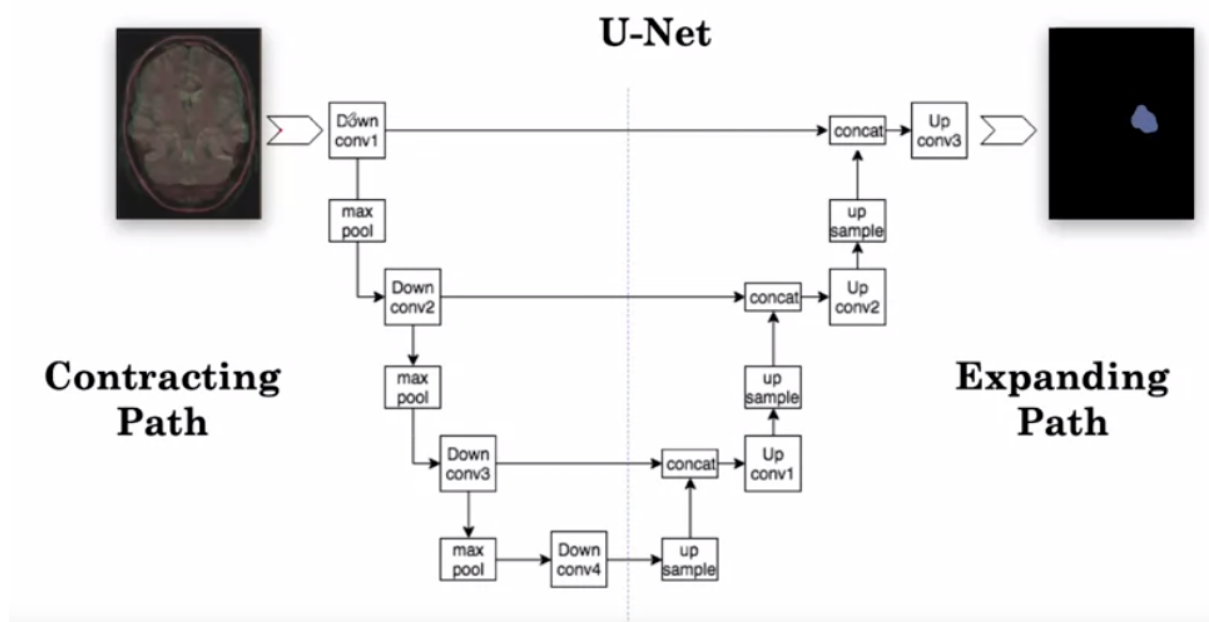
In repeated sampling, this method produces intervals that include the population accuracy in about 95% of samples

影响因素: 样本大小

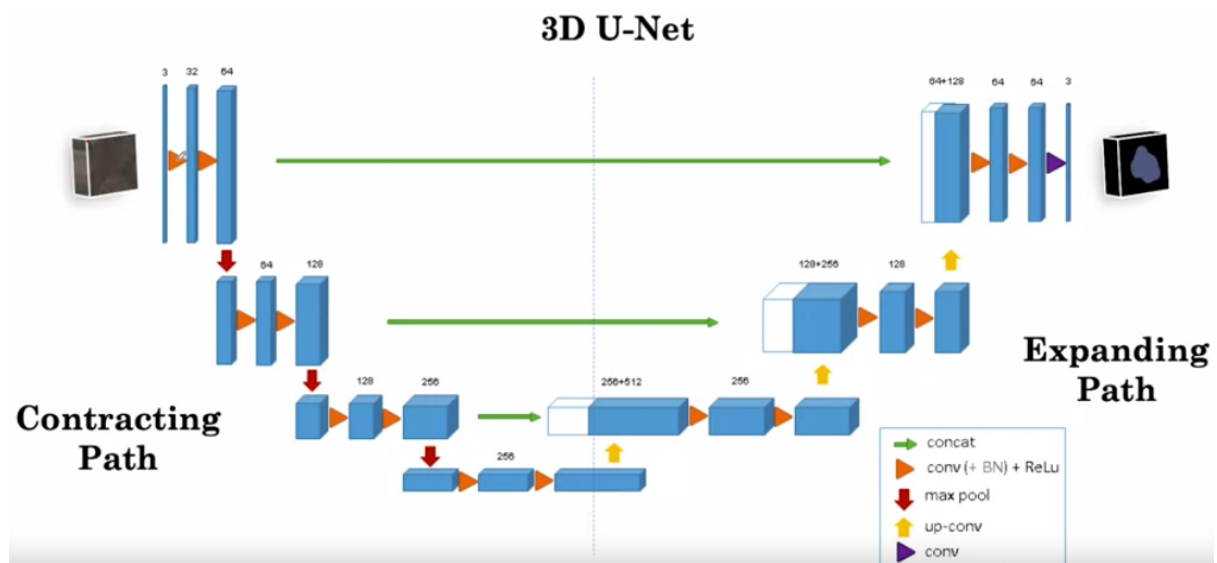
### 第三周 核磁共振成像的图像分割

#### 图像分割

1. 图像配准Image Registration：旋转图像来对齐
2. 图像分割方法
  - a. 2D:切片, 处理, 组合。缺点：失去3D上下文信息
  - b. 3D:切块。缺点：失去相邻空间信息
3. U-Net
  - a. 2D U-Net



b. 3D U-Net



4. 数据增强
  - a. 区别一：对输入旋转了90度，意味着对输出也要旋转90度
  - b. 区别二：3D增强
5. 损失函数

软骰子损失Soft Dice Loss：在不平衡数据上性能好

## Soft Dice Loss

$$L(P, G) = 1 - \frac{2 \sum_i^n p_i g_i}{\sum_i^n p_i^2 + \sum_i^n g_i^2}$$

实践思考

1. 不同人群和诊断技术
  - a. 不同国家的病症
  - b. 不同的检测设备
2. 使用外部验证集来检验模型的泛化能力。调整模型以适应真实世界数据
3. 实际效果衡量
  - a. 决策曲线分析Decision Curve Analysis
  - b. 随机对照
  - c. 分析对人口子群体的影响：找寻算法偏差
  - d. 模型可解释性