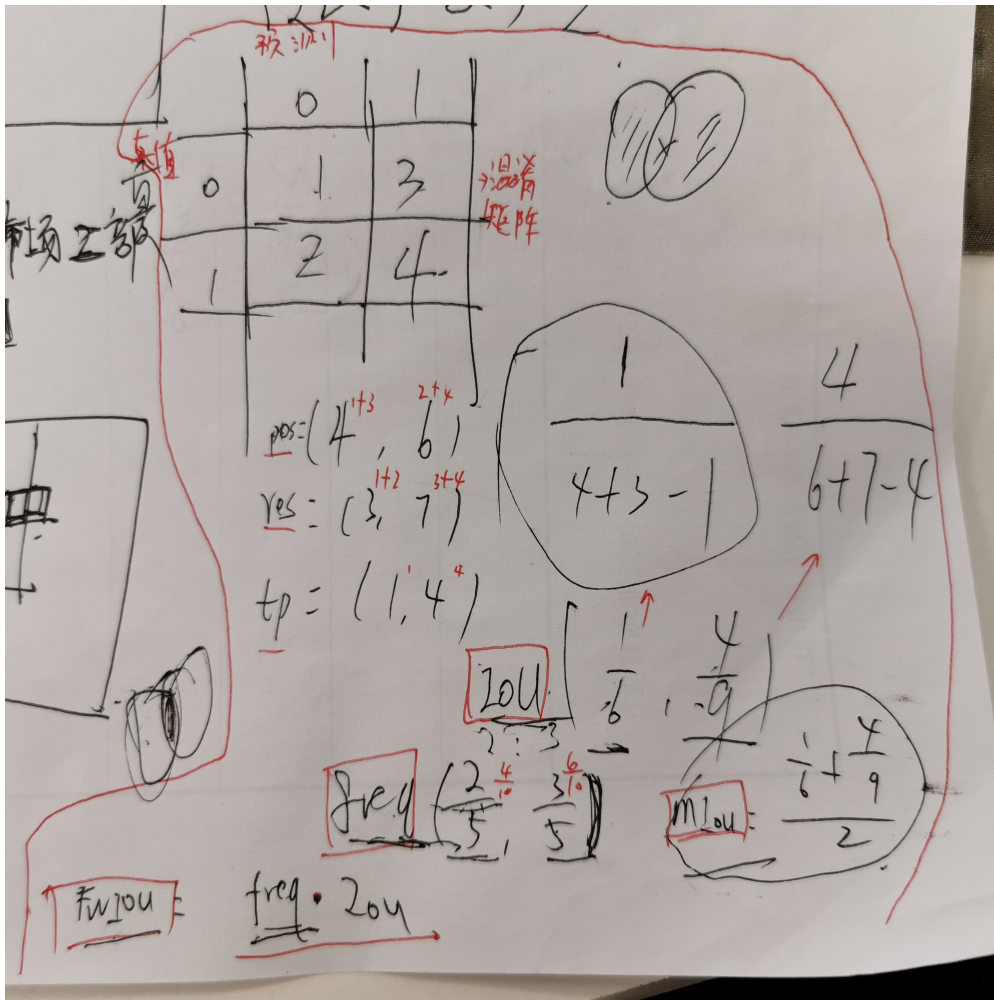


1 语义分割

1.1 评价指标-miou/fwiou

基于混淆矩阵计算



1.2 模型原理

语义分割发展史：

1. 深度学习方法碾压传统方法
2. 在全卷积网络FCN之前，深度学习处理语义分割的思路是patch classification，即对图像切块来做像素分类
3. FCN是具有里程碑意义，卷积后接反卷积（上采样）而不是全连接层来实现逐像素分类。上采样损失细节信息，可以通过跳跃连接来有效还原部分细节信息。
4. 归属于encoder-decoder的unet是在fcn的思想上的网络；用于医学影像分割；特征提取用的不是vgg等预训练网络，而是自由扩展深度的cnn；unet还有一个特点是捷径连接。
5. 空洞卷积取代pooling来扩大感受野且能够保持图像尺寸。

PSPNet：

1. 场景分析
2. 基于fcn的方法缺乏恰当的手段来利用全局场景类别线索。

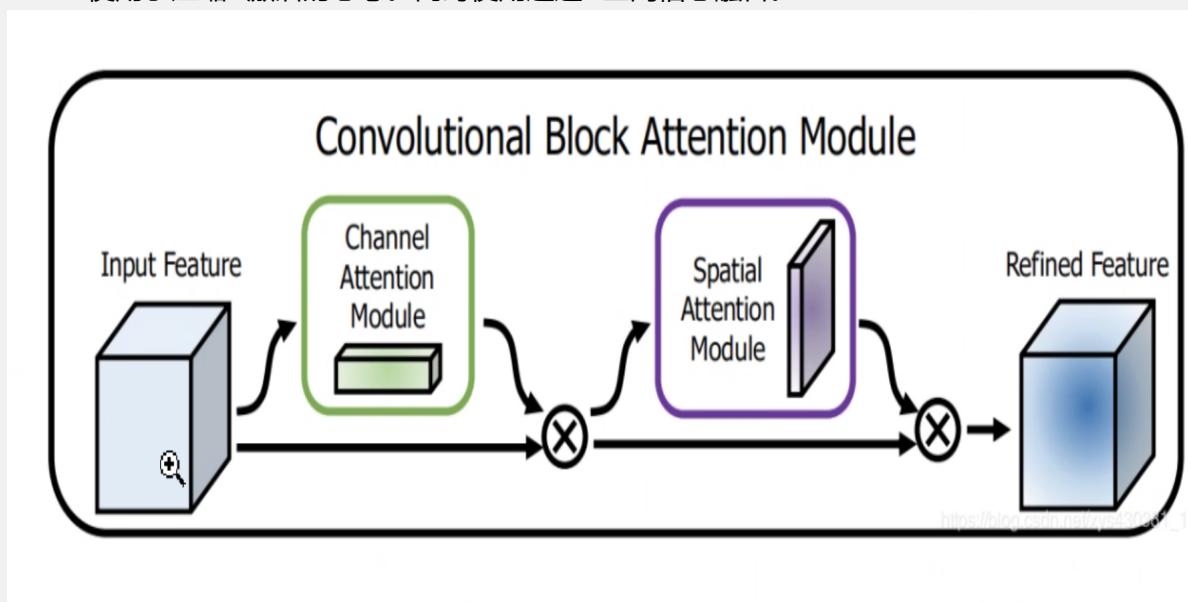
3. 空间金字塔池化对于捕捉全局信息效果很好，空间金字塔网络进一步加强了这一能力。
4. 除了结合全局和局部的线索，该论文提出结合【深度监督损失】的优化策略。

deeplab v3+:

1. modified xception
2. xception骨干网络 input flow调用三次block函数
3. 深度可分离卷积的实现用到了pytorch的pad函数
4. 网络设计理念：模块化设计
5. encoder-decoder 有助于提取锐利边缘？（deeplabv3+摘要中描述，理解有误）
6. Unet和fcn都属于编码-解码结构
7. pytorch实现：sequential 定义块
8. pytorch forward 函数是模型定义的逻辑流

注意力语义分割（信息融合）：

1. nonlocal 保边滤波，保持类间方差的同时，减少类内方差的效果。
2. PSANET 两路attention，相当于transformer中的两个head，两路分别起到collect和distribute的作用。
3. Senet 通道信息融合
4. Cbam 使用了压缩+激活的思想。同时使用通道+空间信息融合。



5. danet是cbam和non-local的结合
6. gcnet优化danet，优化no-local的时间复杂度
7. Emanet是期望最大值的注意力机制

2 优化算法

2.1 梯度下降

important kind of search in current AI

based on concepts from multivariable calculus

search in continuous state space

```

gradientDescent(L){ //损失函数输入
  S = [s1, s2, ..., sn]; // 初始变量
  repeat{
    G = gradient(L,S); //求L在S处的梯度G
    S = S - e*G; //更新S, 速率是e
  } until(termination condition)
  return S;
}

```

3 损失函数

3.1 交叉熵损失函数

对于两个概率分布,一般可以用交叉熵来衡量它们的差异. 标签的真实分布 \mathbf{y} 和模型预测分布 $f(\mathbf{x}; \theta)$ 之间的交叉熵为

$$\mathcal{L}(\mathbf{y}, f(\mathbf{x}; \theta)) = -\mathbf{y}^T \log f(\mathbf{x}; \theta) \quad (2.17)$$

$$= -\sum_{c=1}^C y_c \log f_c(\mathbf{x}; \theta). \quad (2.18)$$

<https://nndl.github.io/>

比如对于三分类问题, 一个样本的标签向量为 $\mathbf{y} = [0, 0, 1]^T$, 模型预测的标签分布为 $f(\mathbf{x}; \theta) = [0.3, 0.3, 0.4]^T$, 则它们的交叉熵为 $-(0 \times \log(0.3) + 0 \times \log(0.3) + 1 \times \log(0.4)) = -\log(0.4) = 0.39$