

# 实验2: TinySeg图像分割模型对比与优化研究

谢翰然 M202473150 覃佐睿 M202473146 日期: 2025/3/24

该实验围绕图像分割模型的核心架构展开对比研究, 基于小型数据集TinySeg(128×128像素), 可使用PyTorch、MindSpore等成熟深度学习框架, 要求学生实现并优化三种主流模型PSPNet、DeepLabv3(含ASPP模块)和CCNet(交叉注意力机制), 通过数据增强、长周期训练(≥50 epoch)及多维度评估(mIoU、Dice系数、边缘F1分数、混淆矩阵), 探究不同解码器结构对分割精度、计算效率和长程依赖建模能力的影响, 最终通过可视化错误样本和参数对比, 深化对语义分割模型设计原理的理解, 同时培养工业级深度学习任务的工程实践能力, 为真实场景中的实时分割任务提供选型参考。

## 1. 模型实现

### 1.1 PSPNet

- 架构: ResNet18(骨干)+PSPNet(解码器)
- PSP模块需包含4级金字塔池化(1×1, 2×2, 3×3, 6×6)
- 最终特征图通过双线性上采样恢复至128×128

### 1.2 DeepLabv3

- 架构: ASPP模块
- branches = [Conv2d(in\_channels, out\_channels, 1), # 1×1 卷积
- Conv2d(in\_channels, out\_channels, 3, padding=6, dilation=6), # 空洞率 6
- Conv2d(in\_channels, out\_channels, 3, padding=12, dilation=12), # 空洞率 12
- Sequential(GlobalAvgPool2d(), Conv2d(...), Upsample(...))
- 最使用1×1卷积将4分支特征相加后输出

### 1.3 CCNet

- 特征映射: 输入特征图 $(X \in \mathbb{R}^{C \times H \times W})$
- 生成Query/Key
- 注意力计算:
  - 对每个位置 $((i,j))$ , 计算水平和垂直方向的关联权重
  - 两次交叉注意力(Criss-Cross)迭代
- 输出融合:  $(Y = \text{Conv}(X + \text{Attention}(X)))$

## 2.训练与评估配置

### 2.1 训练参数

- 训练周期：至少50个epoch（早停条件：验证集mIoU连续10个epoch无提升）
- 优化器：AdamW（lr=0.001,weight\_decay=1e-4）
- 学习率调度：CosineAnnealingLR（T\_max=20,eta\_min=1e-5）
- 损失函数：CrossEntropyLoss+DiceLoss(权重比1:1)

### 2.2 评估指标

- Pixel Accuracy:  $\left( \frac{\text{正确像素数}}{\text{总像素数}} \right)$
- mIoU:  $\left( \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FP_c + FN_c} \right)$
- Dice系数:  $\left( \frac{2 \times TP}{2 \times TP + FP + FN} \right)$

### 3.模型结构图

#### 3.1 PSPNet

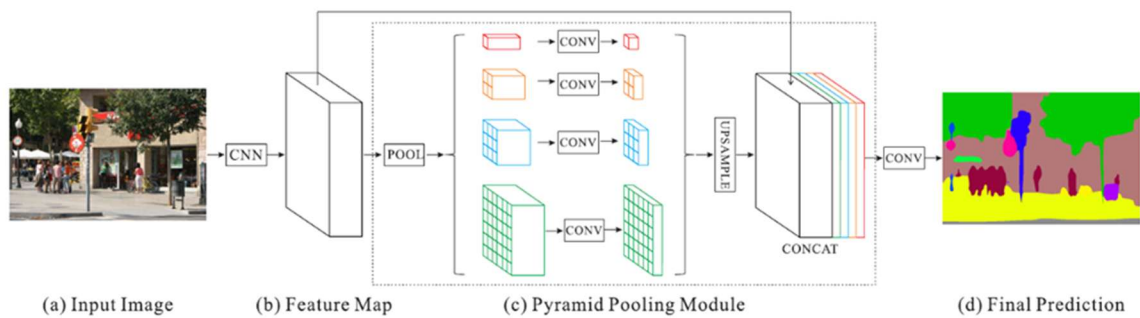


图1 PSPNet模型结构图

#### 3.2 Deeplabv3

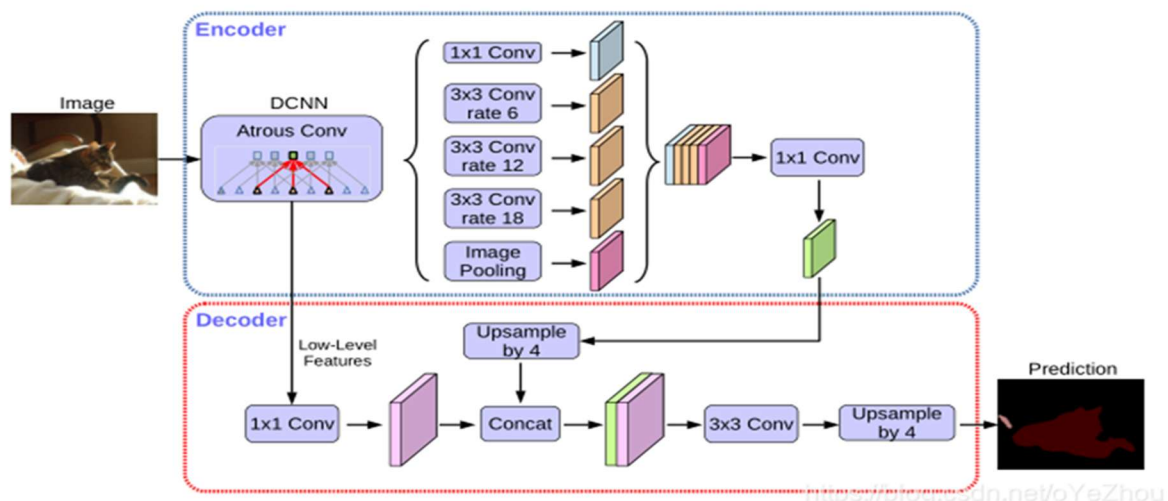


图2 DeepLabv3模型结构图

#### 3.3 CCNet

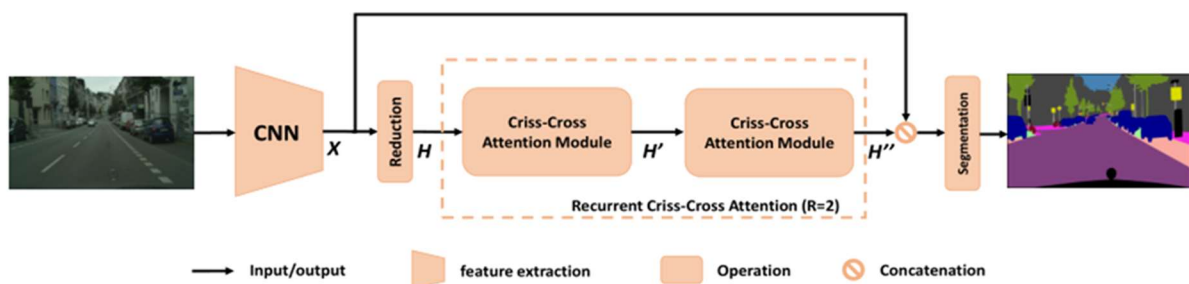


图3 CCNet模型结构图

# 4.数据增强消融实验

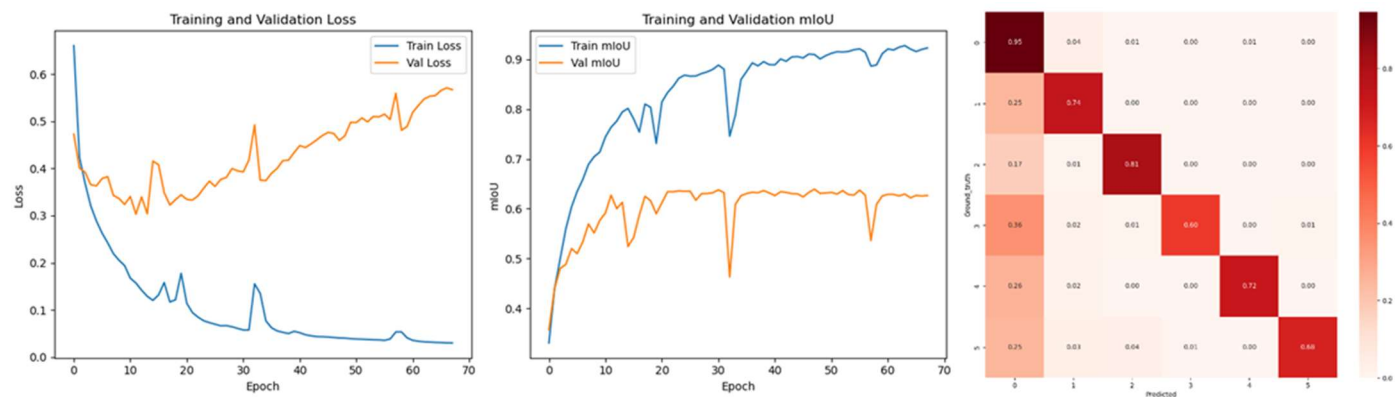
## 4.1 消融实验设定

对于三种主流模型—PSP Net、 DeepLabv3（含ASPP模块）和CC Net（交叉注意力机制），按照下表配置，构建不同的实验组，每个实验组分别使用不同的数据增强方法，其他所有实验条件都保持相同。在每个组上使用相同的训练数据和验证数据来训练模型，并使用相同的评估指标来评估模型的性能：

配置	增强组合
无增强	仅归一化
基础增强	base_aug
完整增强	base_aug + advanced_aug

## 4.2 训练过程可视化结果

### 4.2.1 PSPNet



PSPNet无增强LOSS和mIoU随epoch的变化曲线和混淆矩阵

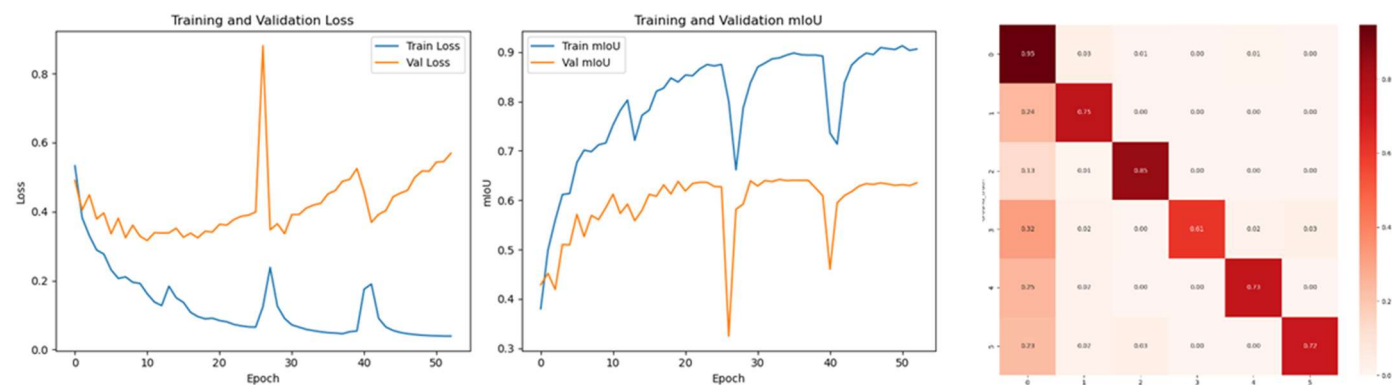


PSPNet基础增强LOSS和mIoU随epoch的变化曲线和混淆矩阵



PSPNet完整增强LOSS和mIoU随epoch的变化曲线和混淆矩阵

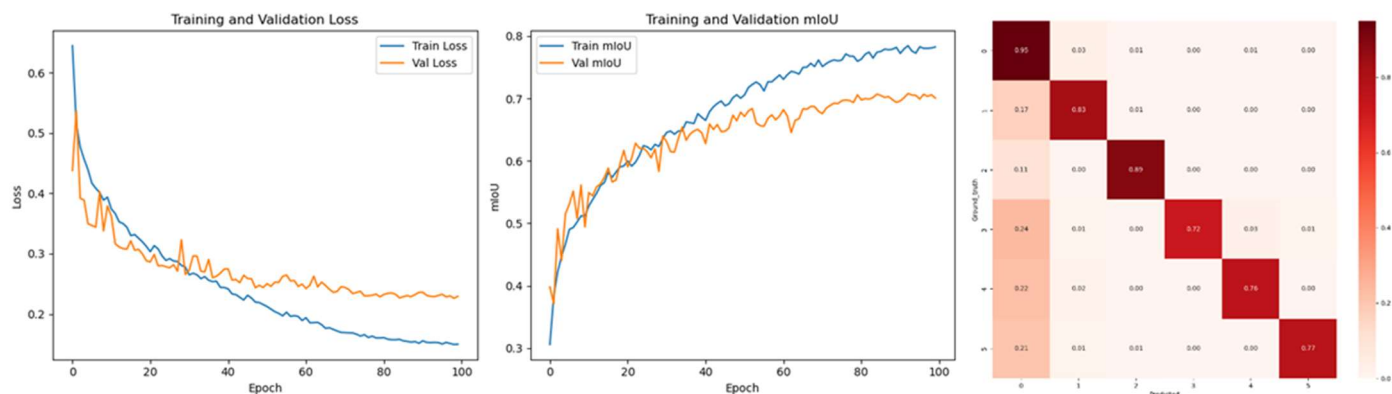
## 4.2.2 Deeplabv3



Deeplabv3无增强LOSS和mIoU随epoch的变化曲线和混淆矩阵

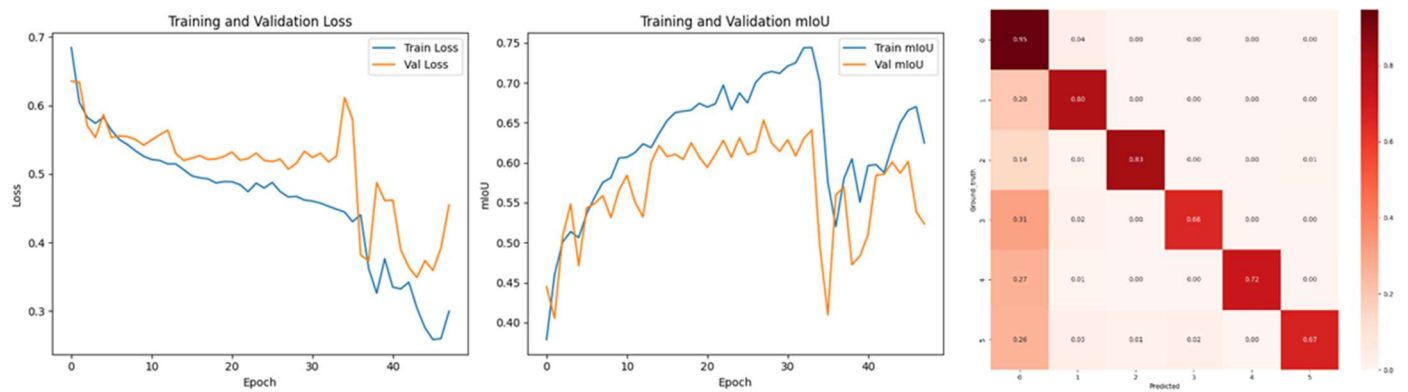


Deeplabv3基础增强LOSS和mIoU随epoch的变化曲线和混淆矩阵



Deeplabv3完整增强LOSS和mIoU随epoch的变化曲线和混淆矩阵

## 4.2.3 CCnet:



## 5.模型对比表格

### 5.1 PSPnet

	Pixel Accuracy	mIoU	Dice系数
无增强	0.8929	0.6102	0.6779
基础增强	0.8983	0.66	0.7152
完整增强	0.9128	0.6914	0.7475

### 5.2 Deeplabv3

	Pixel Accuracy	mIoU	Dice系数
无增强	0.8767	0.6445	0.7038
基础增强	0.9059	0.6896	0.7363
完整增强	0.9195	0.7081	0.7491

### 5.3 CCNet

	Pixel Accuracy	mIoU	Dice系数
无增强	0.8915	0.6474	0.6900
基础增强	0.9076	0.6763	0.7072
完整增强	0.9047	0.6951	0.7440

## 6.实验结果分析

### 6.1 实验结论

- Pixel Accuracy与mIoU差距较大，可能因部分类别（如小目标）分类不稳定。
- 完整增强显著提升边缘分割能力
- 完整增强导致CCNet性能下降（mIoU从0.6463→0.6051），可能因交叉注意力机制对复杂增强敏感，或训练时未充分收敛。
- DeepLabv3可能是因为ASPP模块擅长边缘建模，效果最好



## 6.2 优化建议

- CCNet训练策略：减少高级增强强度（如降低弹性变换的 $\alpha$ 参数）增加注意力头数（从2→4），增强局部特征捕捉能力。
- DeepLabv3小目标改进：添加辅助损失函数（如边缘感知损失）；调整ASPP空洞率（增加小目标对应的空洞卷积分支）。
- PSPNet效率优化：替换骨干网络为轻量级结构（如MobileNetV3）；减少金字塔池化层级（从4→2），降低计算量。

## 7.预测结果可视化

### 7.1 PSPnet

#### 7.1.1 样本1



PSPnet无增强预测结果



PSPnet基础增强预测结果



PSPnet完整增强预测结果



### 7.1.2 样本2



PSPnet无增强预测结果



PSPnet基础增强预测结果



PSPnet完整增强预测结果

### 7.1.3 样本3



PSPnet无增强预测结果



PSPnet基础增强预测结果



PSPnet完整增强预测结果

## 7.2 Deeplabv3

### 7.2.1 样本4



Deeplabv无增强预测结果



Deeplabv基本增强预测结果



Deeplabv完整增强预测结果

### 7.2.2 样本5



Deeplabv无增强预测结果

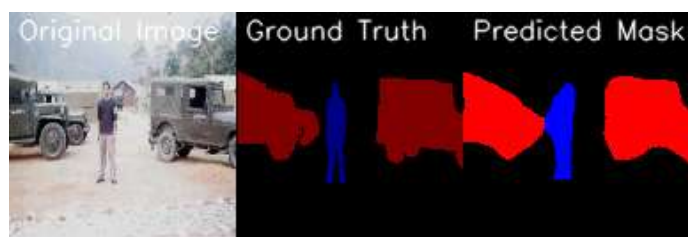


Deeplabv基本增强预测结果

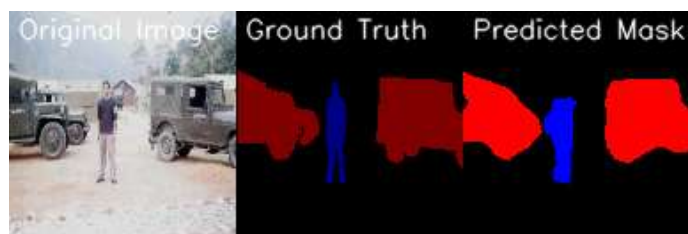


Deeplabv完整增强预测结果

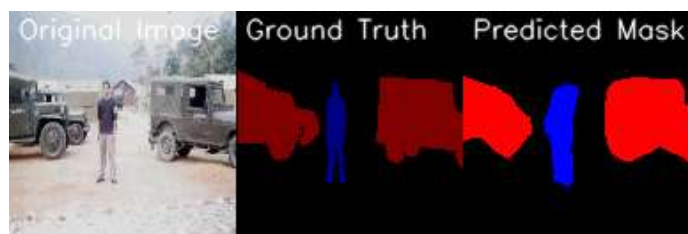
### 7.2.3 样本6



Deeplabv3无增强预测结果



Deeplabv3+基本增强预测结果



Deeplabv3++完整增强预测结果

## 7.3 CCNet

### 7.3.1 样本7



CCNet无增强预测结果



CCNet+基本增强预测结果



CCNet++完整增强预测结果

### 7.3.2 样本8



CCNet无增强预测结果



CCNet基本增强预测结果



CCNet完整增强预测结果

### 7.3.3 样本9



CCNet无增强预测结果



CCNet基本增强预测结果



CCNet完整增强预测结果