# MICCAI2D牙齿分割赛道第四名技术方案

## 1整体流程

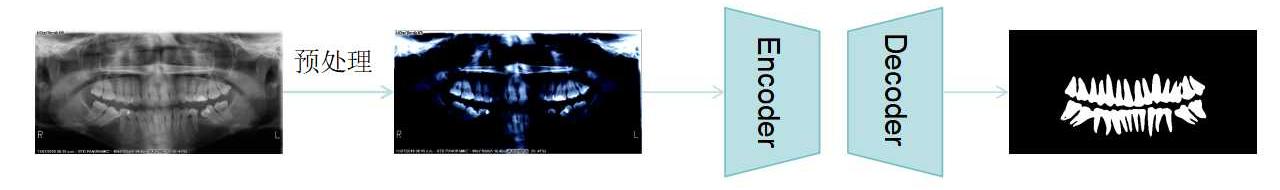


图1 整体流程

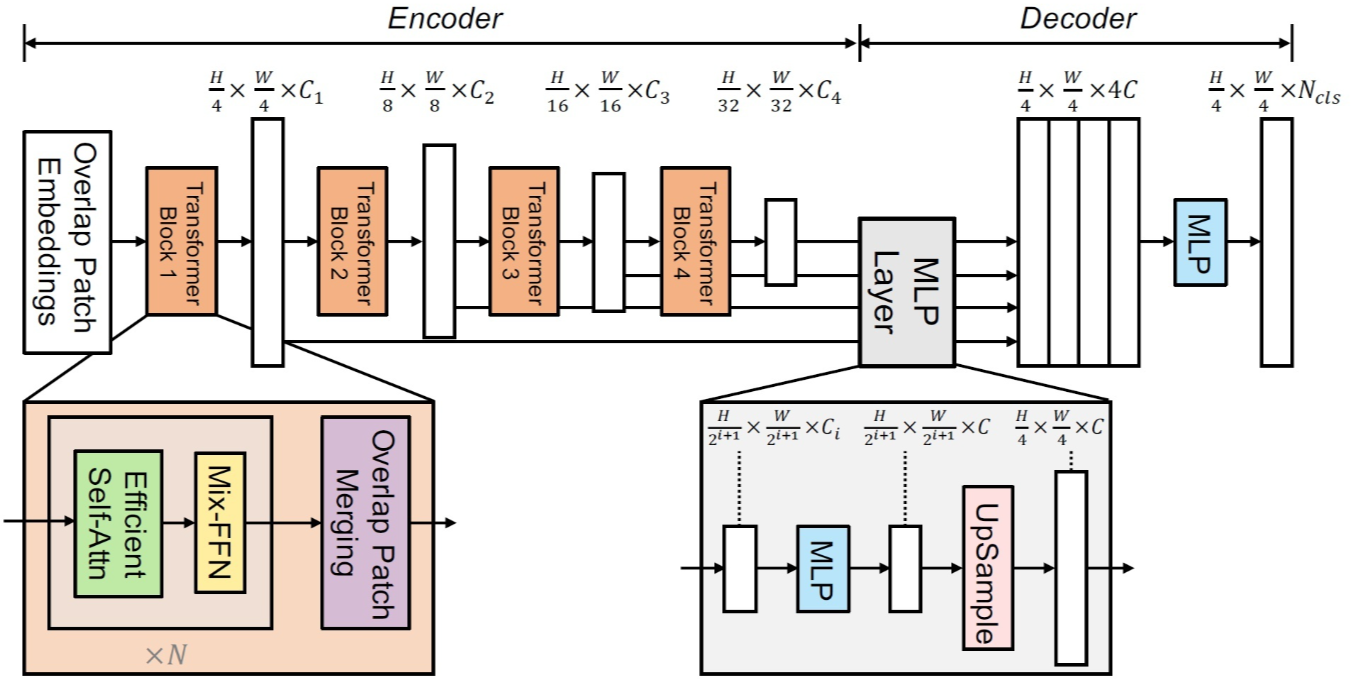
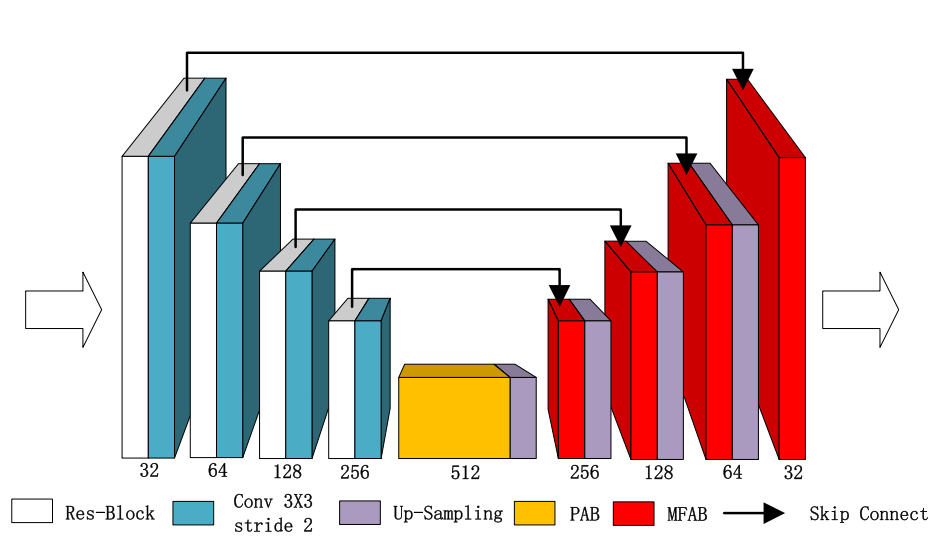
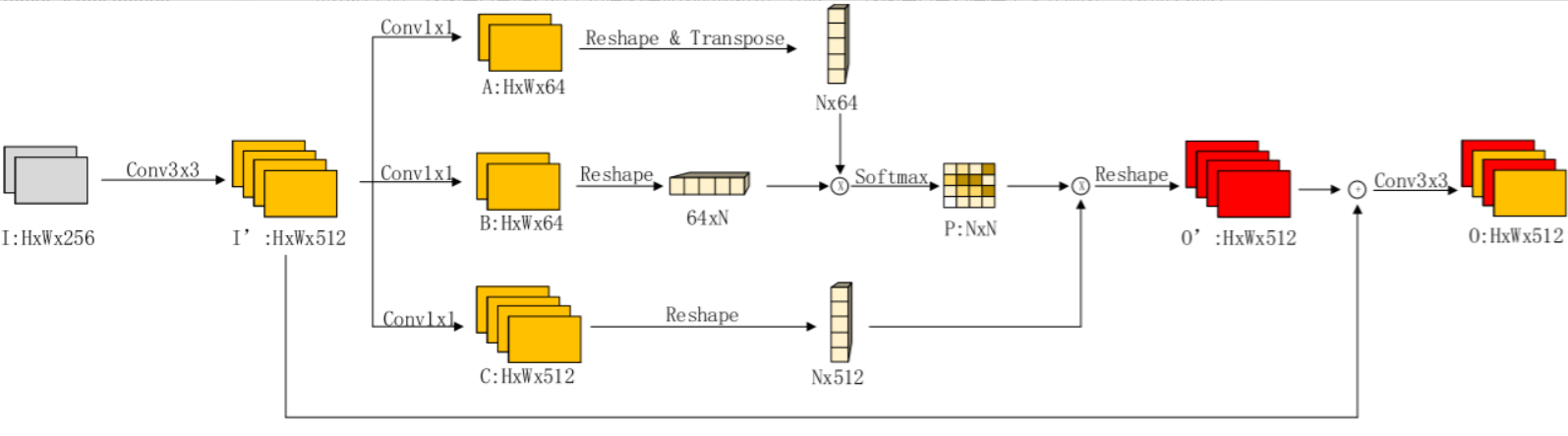


图2 Encoder内部架构

整体流程如图1所示，原始2D牙齿图像经过数据增强与预处理归一化后送入MixVisionTransformer Encoder(图2)提取多不同阶段的多尺度特征，假设原始图像，H和W为图片的尺寸，在本赛道中分别固定为320和640，C为图像通道数等于3，经过MixVisionTransformer Encoder提取后的特征四个不同阶段的特征图尺寸分别为、、、，其中C1、C2、C3、C4分别为64、128、256和512。



**PAB**



**MFAB**

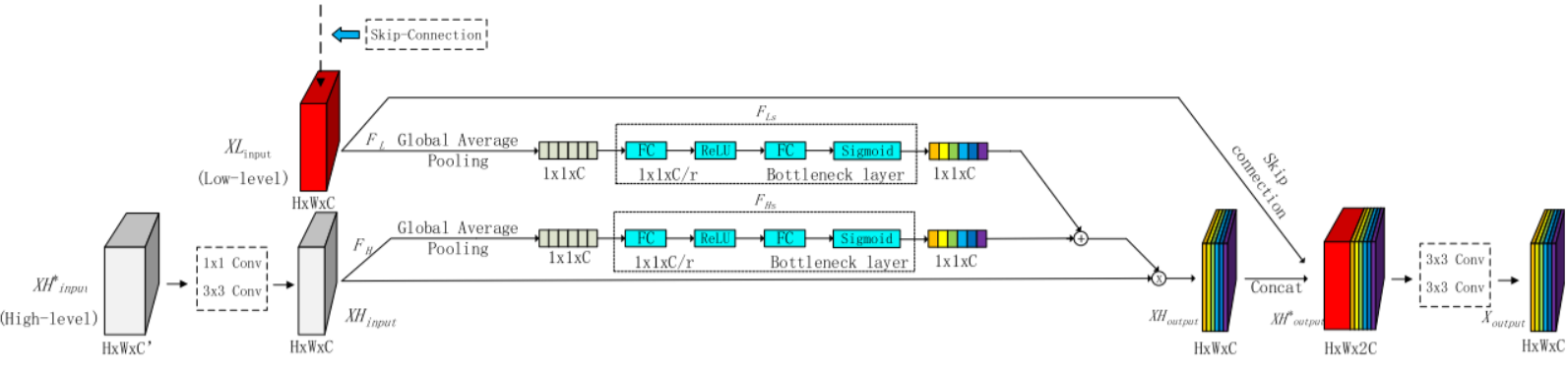


图3 Decoder内部架构

Decoder部分采用了Manet，包含了两种不同的注意力机制模块分别称作Position-wise Attention Block (PAB) 和 Multi-scale Fusion Attention Block (MFAB)，它们分别被用于捕获空间和通道级别的注意力特征图。

## **2实验细节**

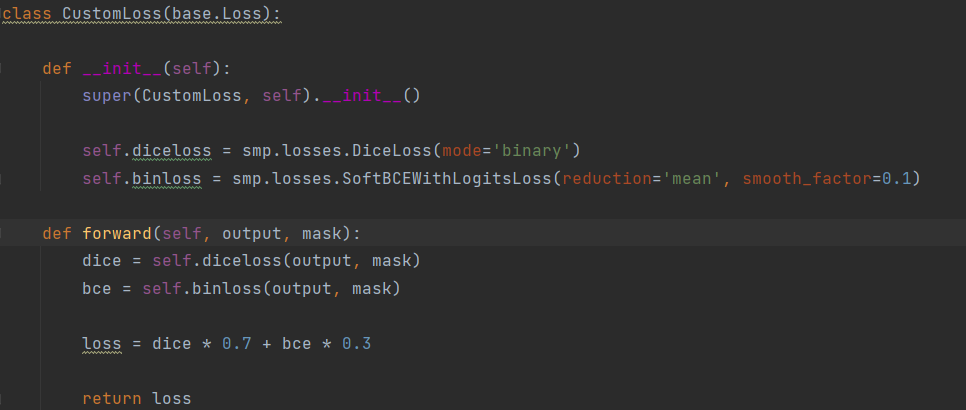
### 2.1数据增强部分

使用RandomFlip,ShiftScaleRotate,ElasticTransform,GaussianBlury以及一些颜色对比度变化如 ColorJitter和RandomBrightnessContrast。由于图片尺寸固定为320\*640，因此没有使用Crop和Resize操作。

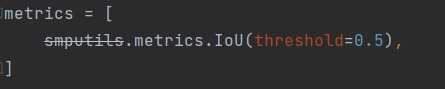


### 2.2损失函数和评价指标部分

使用了Dice和SoftBCEloss的组合，比例系数为7:3。

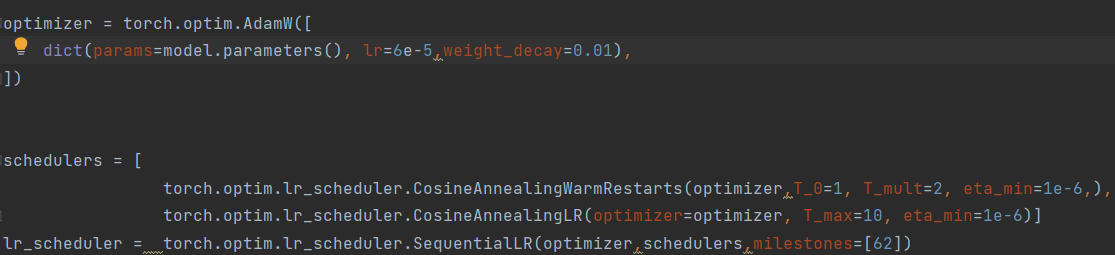


评价指标使用Iou指标阈值设置为0.5



### 2.3训练

使用五折交叉训练，Epoch为200，Batchsize为8，优化器使用AdamW，初始学习率为6e-5，采用余弦退火组合学习率策略如下:



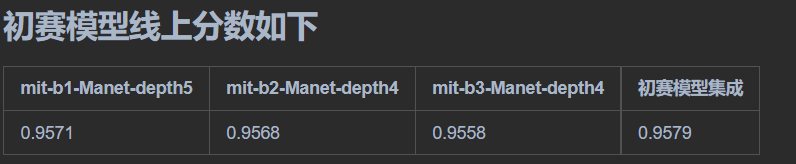
### 2.4 推理

推理过程中采用了Test Time Augmentation(TTA)，对原始图像分别进行水平镜像和垂直镜像预测，然后再翻转回去，最后取三次预测结果的平均作为最终结果。

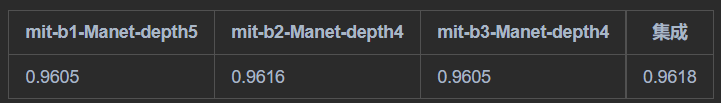


## 3 实验结果

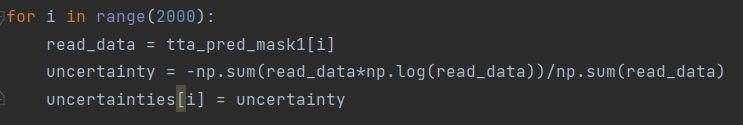
初赛取得了第一名，初赛数据包含2000张有标签数据与500张测试数据 Backbone部分使用了MixVisionTransformer(mit-b1、mit-b2、mit-b3)



复赛与初赛不同，包含900张有标签数据、2000张无标签数据和500张测试数据直接使用初赛集成模型在复赛排行榜上可以达到0.9599，已经是一个不错的分数了。一开始我并没有使用无标签数据，只使用900张有标签数据训练的分数如下:



然后我使用mit-b2-Manet-depth4模型对无标签数据进行推理，同时为了减少低质量的伪标签选择了置信度（通过计算熵值）前Top1000张伪标签加入训练集,伪代码如下:



由于提交次数限制和时间关系，我只选择了mit-b2-Manet-depth4这个模型， 最后在加入伪标签数据之后，5折交叉训练的5个mit-b2-Manet-depth4模型融合的结果在复赛测试集上达到了0.9621(复赛第四)。

## 改进思路

对二分类阈值进行调参，这次比赛我都是用0.5的默认阈值，也许最优阈值不一定为0.5。

考虑多尺度推理TTA以及滑动窗口推理TTA。

尝试使用后处理，如去除细小孔洞等。

由于时间关系，伪标签只迭代了一次，迭代更多次可以提高伪标签的质量。