机器学习 Lab2 实验报告

陆浩旗 韩金铂

1. **数据处理**
   1. **图片Resize**

观察数据后可以发现，图片的大小不一，为了能够对所有数据进行统一的训练，我们使用Tensor提供的Resize功能来将所有图片修改成统一格式，即256\*256的标准大小。

需要注意的是Resize默认使用的是BILINEAR模式，即线性插值，这在原图的时候因为是RGB格式的所有颜色所以影响不大，但是如果对已经转化成256\*256且每个像素是一个类别的label使用这种插值的会出现没有意义的实数类型，因此我们应该选择NEAREST模式或者说强制转成整数来进行处理。

* 1. **数据集划分**

如果对所有图片进行训练并且进行验证容易产生过拟合的问题，因此在开始训练前我们先把数据集按照4:1的比例划分为训练集和测试集两个集合。我们使用测试集对模型进行训练，训练完成之后将模型应用在测试集中并计算IOU评估。

1. **系统设计**
   1. **模型设计**

我们整体选用UNet作为我们使用的语义分割模型，在经典的UNet模型上我们做了一些修改：主要是将卷积层加上了为1的padding，保证卷积之后图片大小不变，之后再使用2\*2的MaxPool进行池化，将图片缩小，总体架构如图2.1所示：

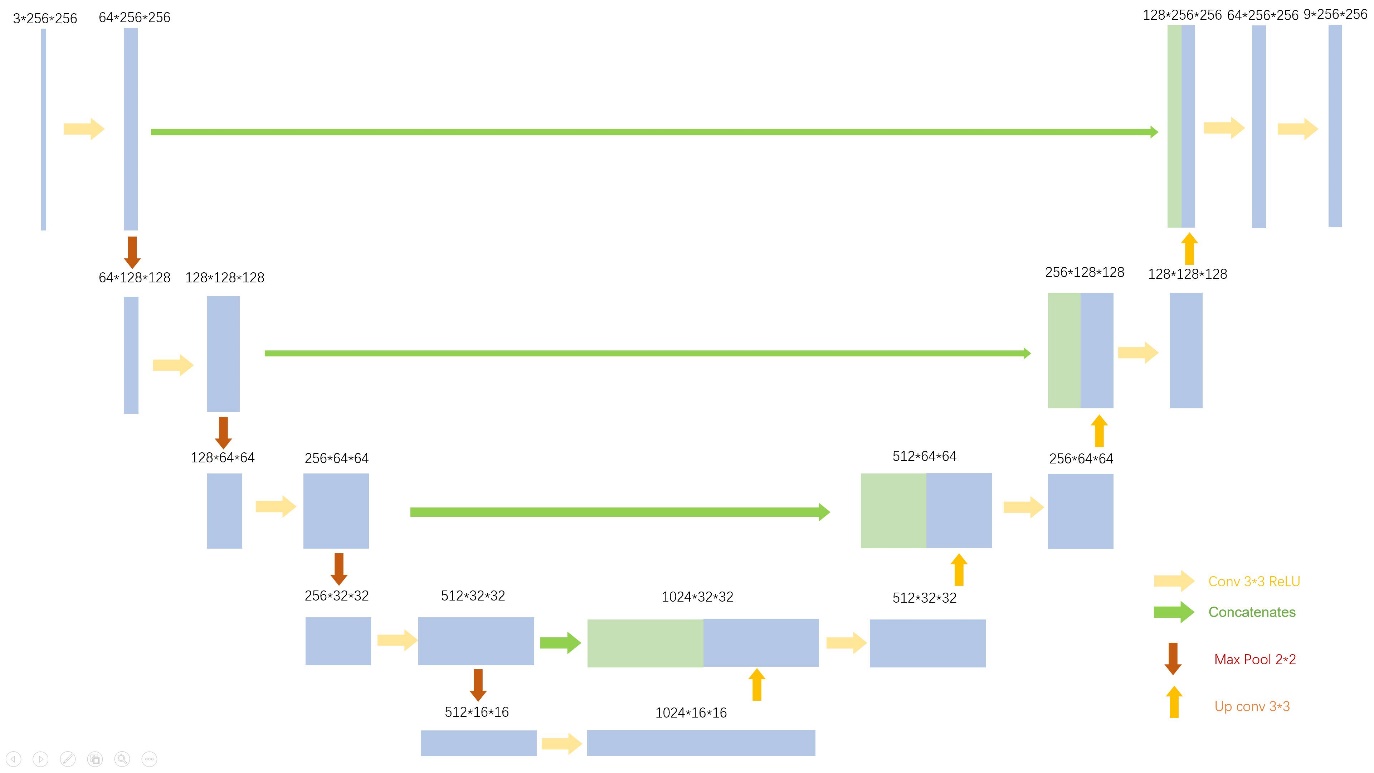


图2.1 本实验采用的修改过后UNet模型

我们输入为3\*256\*256的RGB图片，第一步卷积到channel为64的卷积层，之后的步骤如图2.1所示，最后将64\*256\*256的图片进行一次卷积输出为9\*256\*256的图片，其中9代表我们的类别数目（class number），代表每个像素点为各种类别的概率分布。

* 1. **训练方法**

经过调研发现，我们的任务为多类别图像分割，提供的label为每个像素点所属的类别，因此接近一个多分类任务，这种任务适宜采用交叉熵作为损失函数，之后再用梯度下降法进行训练。训练过程中我们根据GPU的显存大小设置了不同的batch size，如8、16、32，同时根据epoch的不同设置了不同的learning rate，如1e-2、1e-3、1e-4等，并做了对比试验，详见第三部分。

1. **实验结果**
   1. **评价指标**

本实验中我们采用交叉熵作为损失函数，但是最终的评价指标是IOU（Intersection over Union），该指标常常用来衡量目标检测任务中，定义如图3.1：

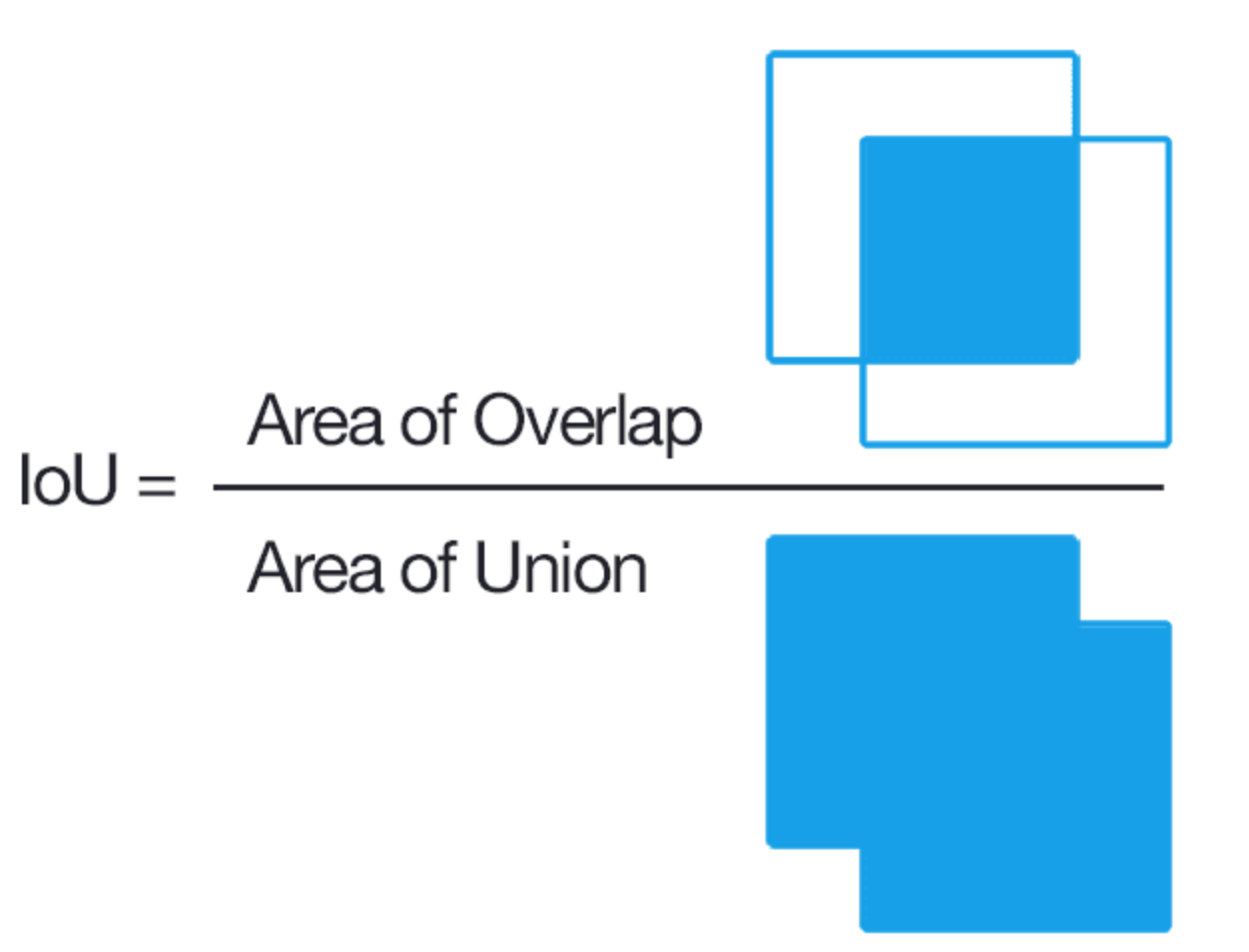


图3.1 IOU定义

因为我们是多分类任务，因为我们可以对每种类别单独计算其IOU，并在最终取平均求出Mean IOU，即MIOU。

* 1. **最终预测准确率**
  2. **调参实验及结果分析**
     1. **Learning Rate**

我们初始选用的learning rate为1e-2，最终的效果如图3.2~3.4所示：

可见效果并不理想，loss函数与IOU值都在后期进行震荡，分析后我们觉得应该是learning rate太大，因此我们采用的调整learning rate的方法：在batch size为32的条件下，1-20轮lr为1e-3，21-100轮lr为1e-4，101-150轮lr为1e-5，在这种情况下的loss曲线如**图3.5**所示。

可以看到loss曲线在1~100轮总体呈现下降趋势，在接近100轮时发生震荡，此时我们降低学习率后继续稳步下降，说明在继续学习，此时再来观察我们各类别的IOU指标，如**图3.6**所示。可以观察到虽然训练集的IOU仍然在稳步上升，在测试集中的IOU保持不变甚至有所下降，这说明我们已经过拟合，学习率不宜到1e-5的数量级进行学习，前100轮已经接近拟合，class 0,1,2,5,7的IOU都在0.6~0.7的水平，已经比较优秀。

如**图3.7**所示，我们给出了一些预测后的图片，可以发现模型对于一些大的轮廓学习的还行，但是对于其中的小物体如人，车，船等黄色的label识别还是不够准确，具体的原因我们将在下一节进行探讨。

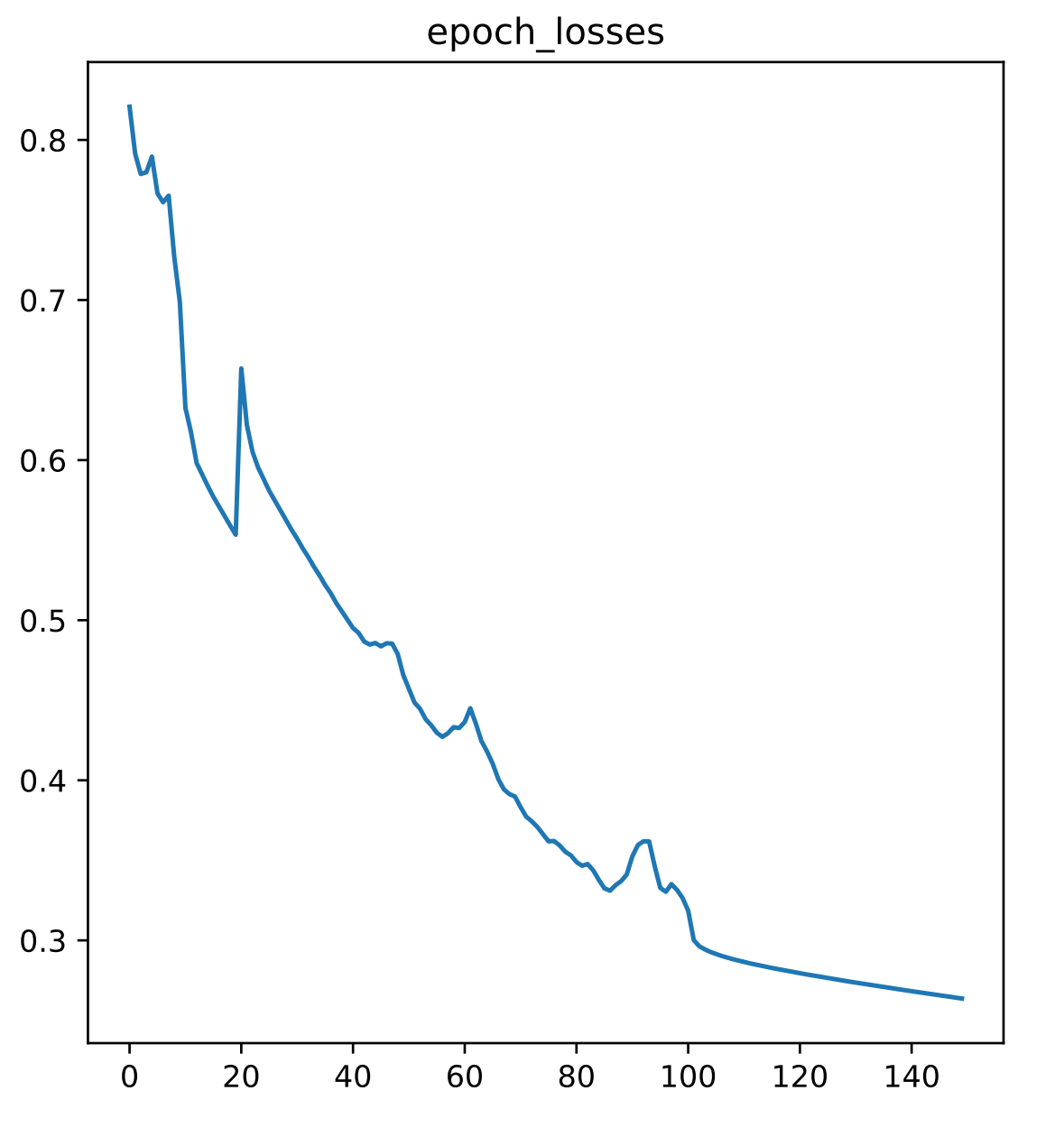


图3.5 自适应learning rate下的loss曲线图

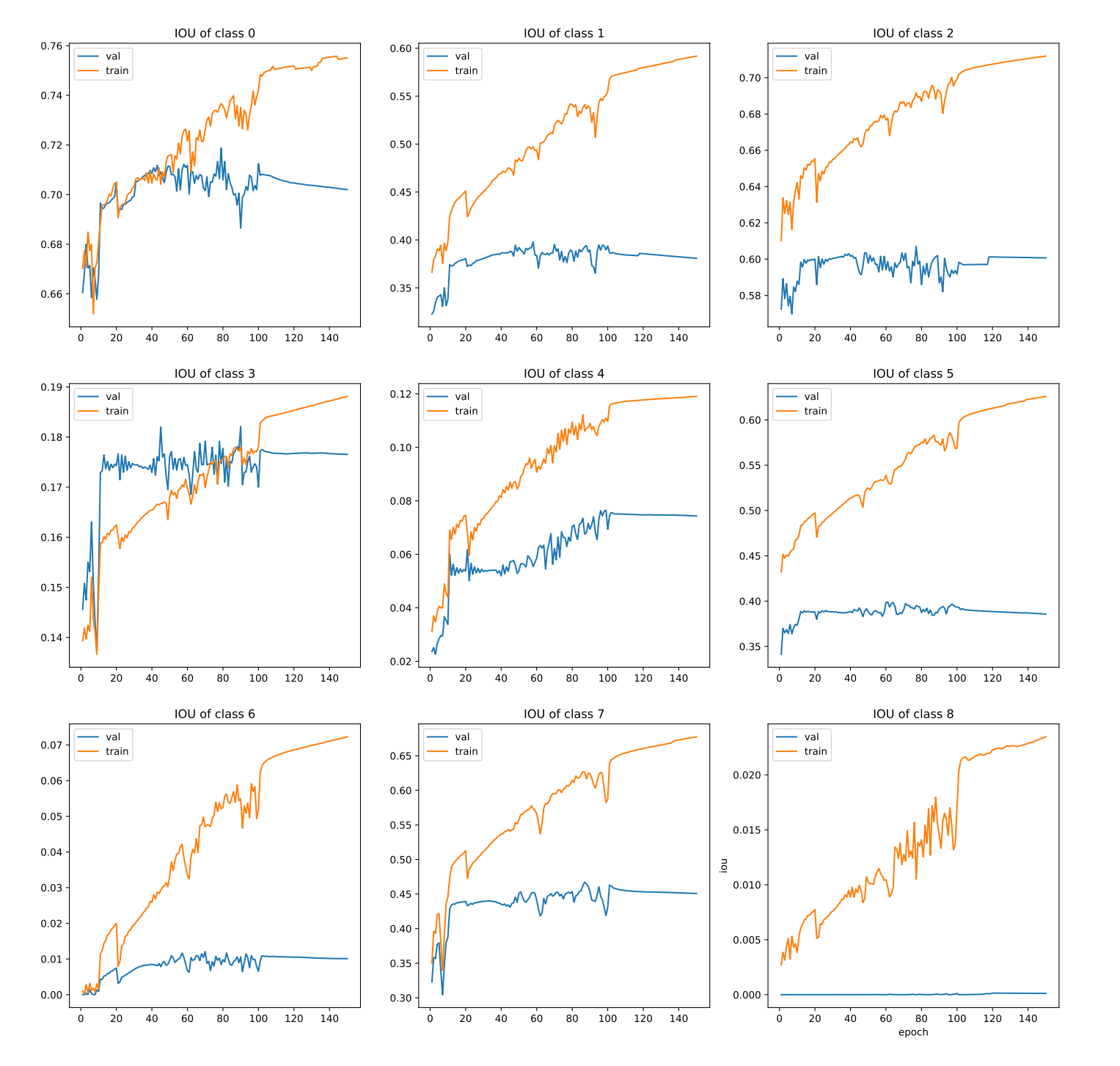


图3.6 自适应learning rate下的各类别IOU图

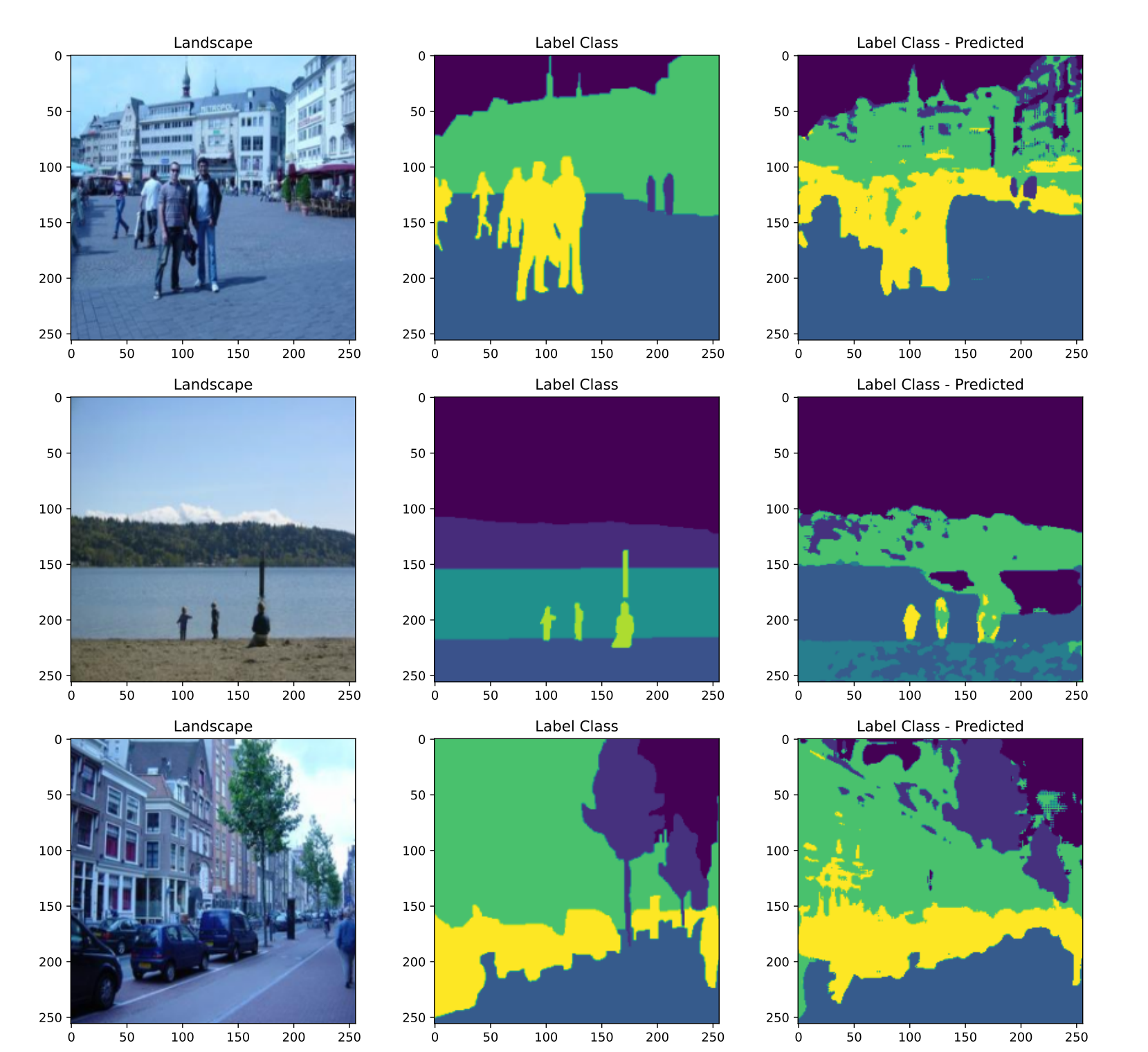


图3.6 自适应learning rate下模型的预测展示

* + 1. **关于各类别数据集分析**

根据3.3.1中各种情况下的训练结果我们可以发现class 6,8两个类别的IOU基本接近0，class3,4两个类别的IOU也不高，在0.1左右，我们将各类别在数据集中出现的次数按像素点进行统计后发现了一些规律，如图3.7，图3.8所示：

可以看出各类别出现次数相当不均匀，而且训练效果的IOU与数据集中的出现次数呈现正相关，这也很好解释，像class6和8分别代表mountain和unknown，出现次数太少模型学习到的程度也相对较低。

图3.7 数据集中各类别出现次数

图3.8 训练集中各类别所占比例

* + 1. **BatchSize**

BatchSize是机器学习中的一个重要参数。在一定范围内，小的BatchSize会引入更多随机性，但可能难以达到收敛；大的BatchSize会使梯度下降方向变得更加准确，减小震荡，但过大的BatchSize可能会产生局部最优的情况。

固定初始学习率lr为0.001，采用自适应的学习率调整方法，分别对BatchSize为2，4，8，16，32进行100次迭代，对比实验结果。

对比图3.10~3.15可得，当BatchSize取2时，模型不收敛；当BatchSize在合理范围内增大时，训练集的IOU会以更快速度上升，验证集的IOU基本保持不变。

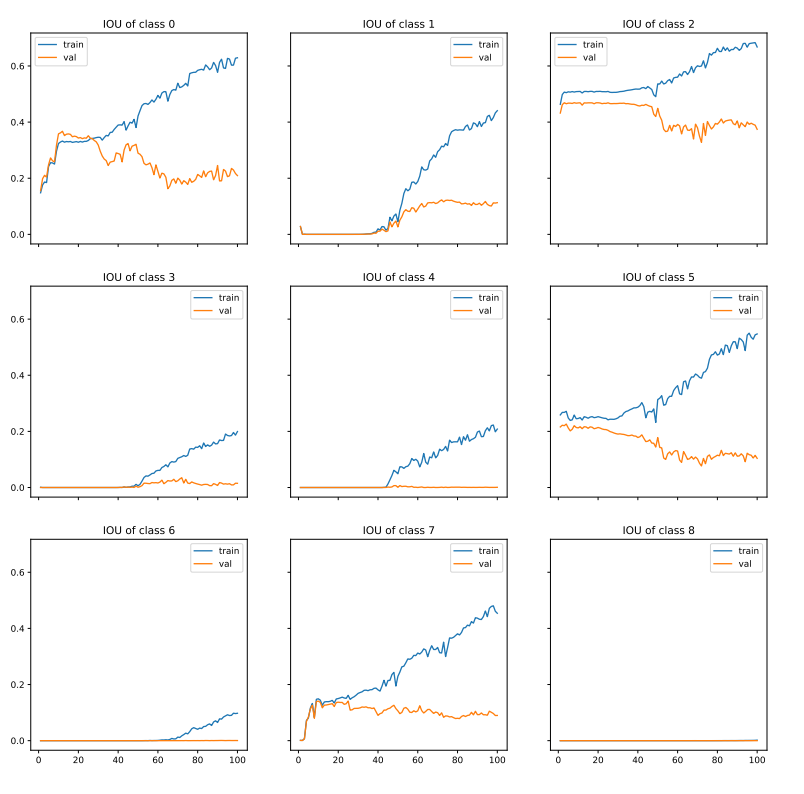


图3.10 BatchSize=2时各类别IOU图

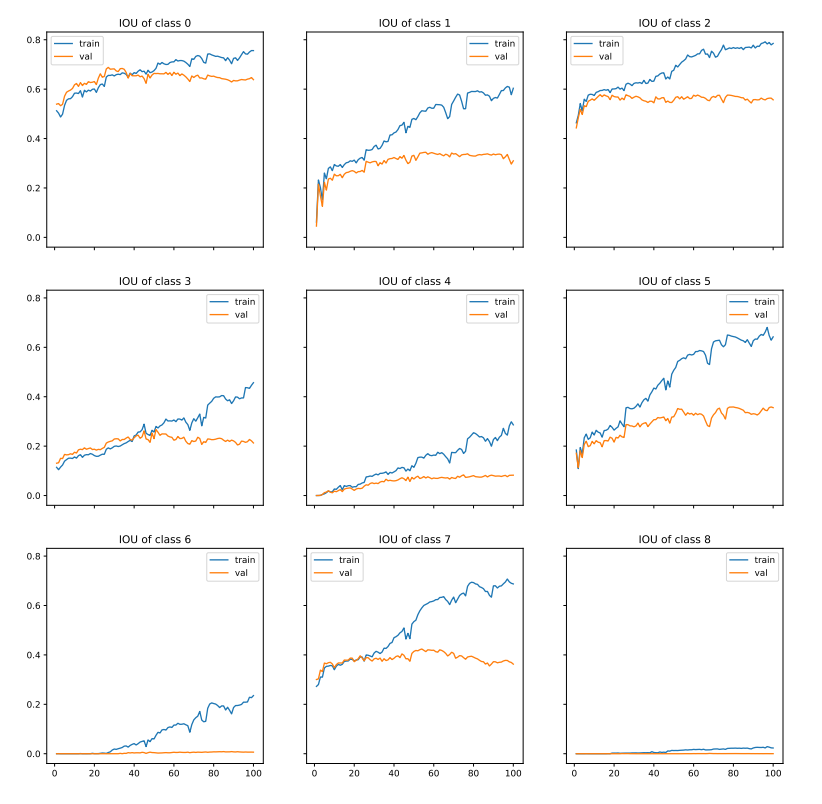


图3.11 BatchSize=4时各类别IOU图