

SRTP第二轮训练总结汇报

- 数据指标及可视化
- 标签图片与预测图片对比
- 不足之处的思考

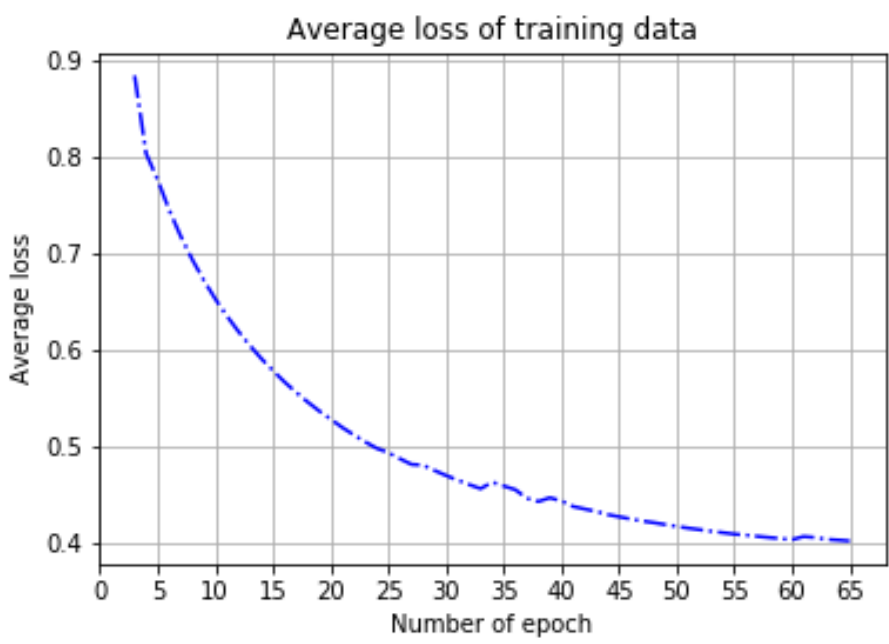
1. 数据指标及可视化

训练与测试数据指标

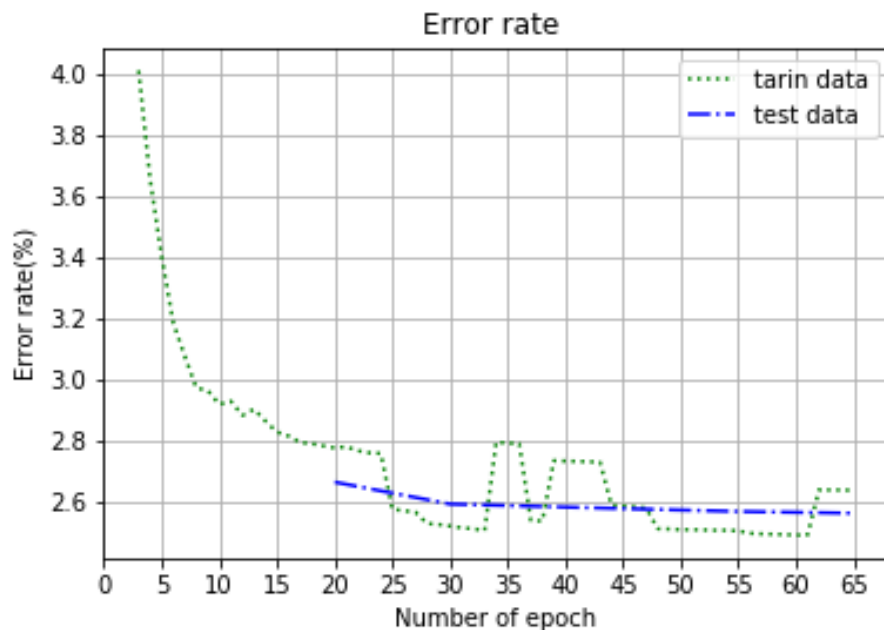
| Epoch | 训练集的平均loss | 训练集的验证误差 | 测试集的验证误差 |
|-------|------------|----------|----------|
| 20 | 0.5286 | 2.690% | 2.662% |
| 30 | 0.4818 | 2.563% | 2.590% |
| 45 | 0.4331 | 2.577% | 2.575% |
| 55 | 0.4095 | 2.502% | 2.566% |
| 65 | 0.4024 | 2.635% | 2.560% |

指标可视化

1. 训练集在不同epoch下的平均loss



2. 训练集和测试集在不不同epoch下的错误率指标

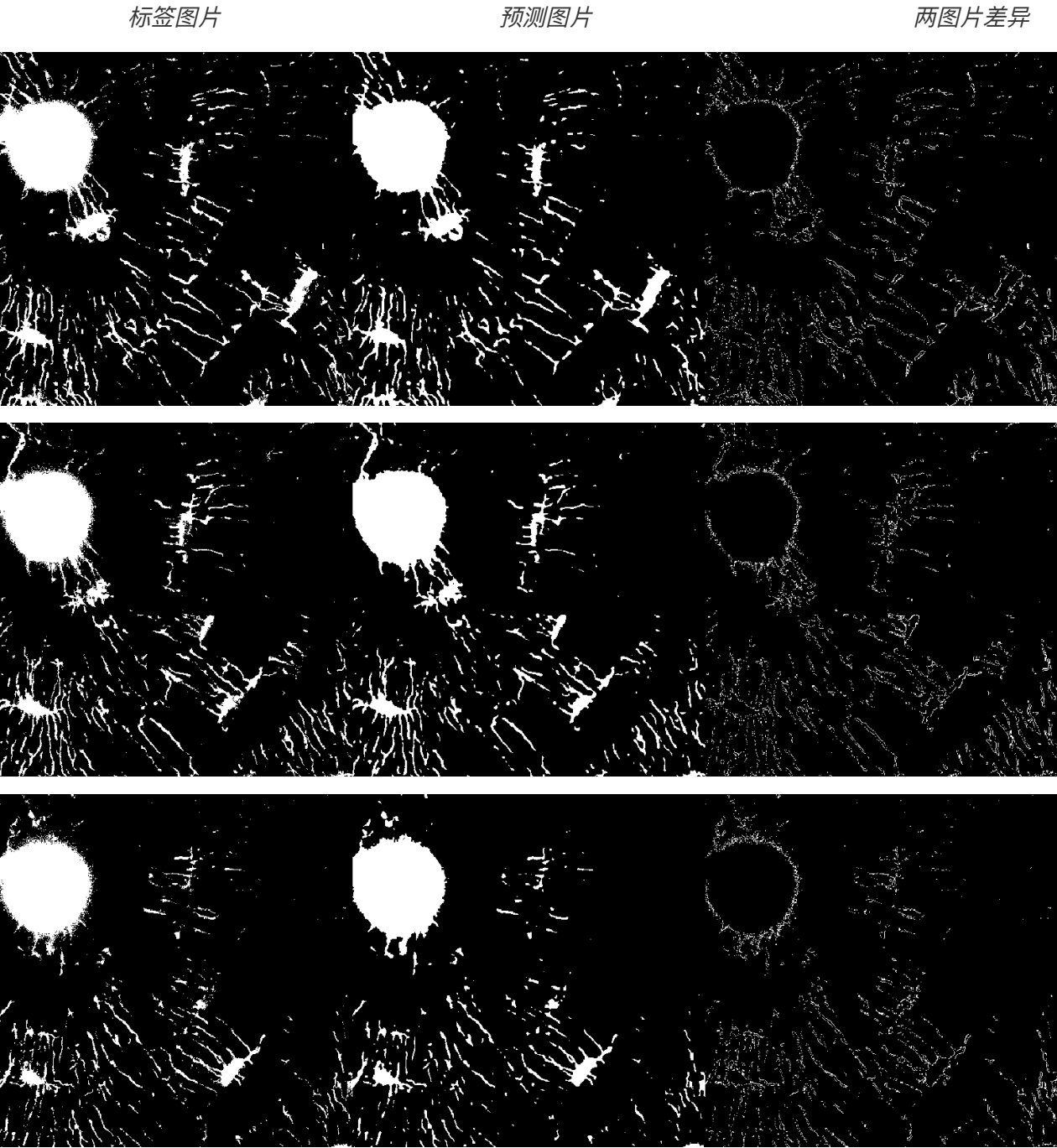


基本说明

- 本次训练模型是在上一次训练20个 $epoch$ 的模型基础上继续训练得来的，总共训练了65个 $epoch$ ，每个 $epoch$ 依然为2000次迭代。
- 测试集误差计算方法：将预测图像保存为二值图后，与标签图片对比算出来的 L_1 误差。
- 此外，由于开始训练的第一个 $epoch$ 的 $loss$ 值与误差值相比比较大后面过大，故删除第一个 $epoch$ 下的两个计算值再来作图。
- 训练中途有中断，恢复训练后相邻的两个 $epoch$ 下会有十分相近的 $loss$ 值和误差值，我们认为没有什么参考意义，所以删除了其中五个 $epoch$ 下的这两个计算值再进行画图。

2. 标签图片与预测图片对比

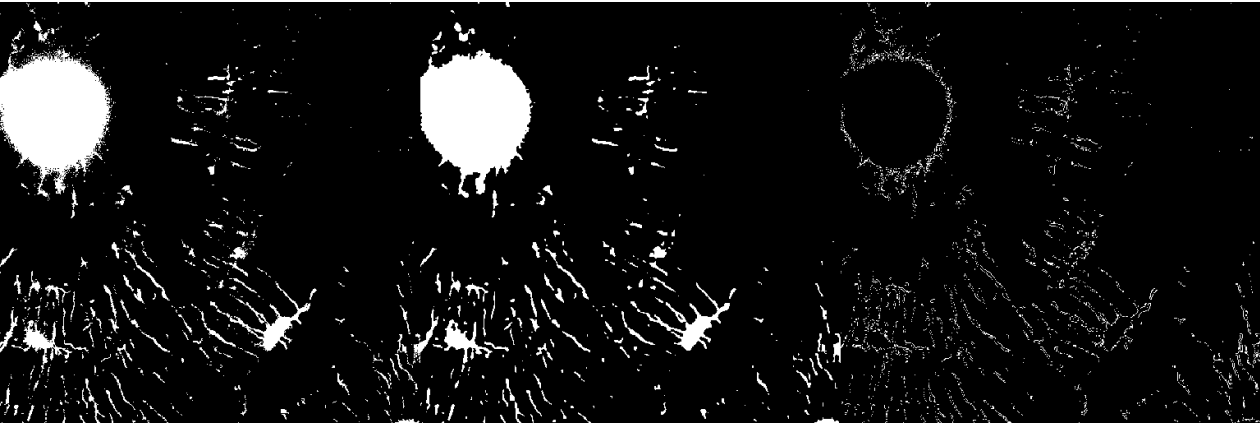
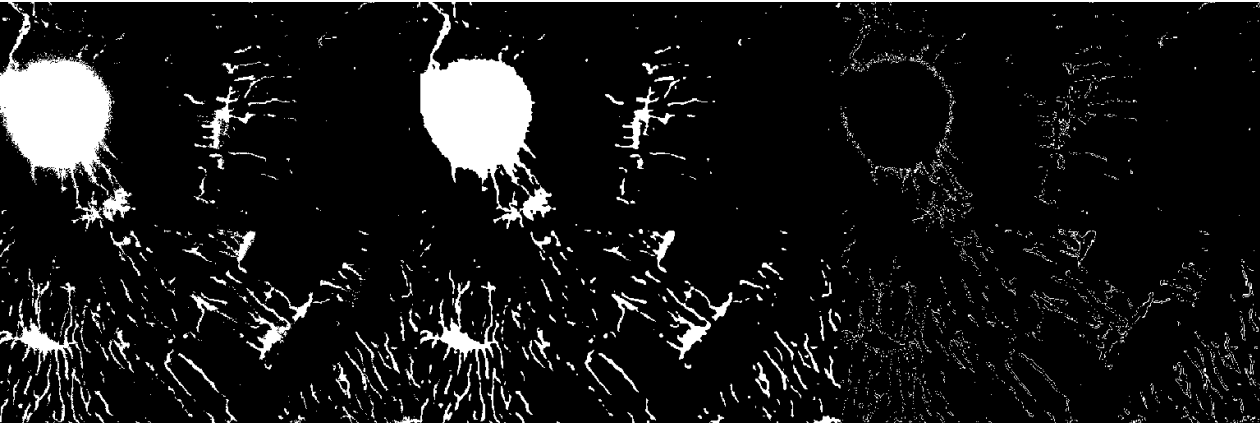
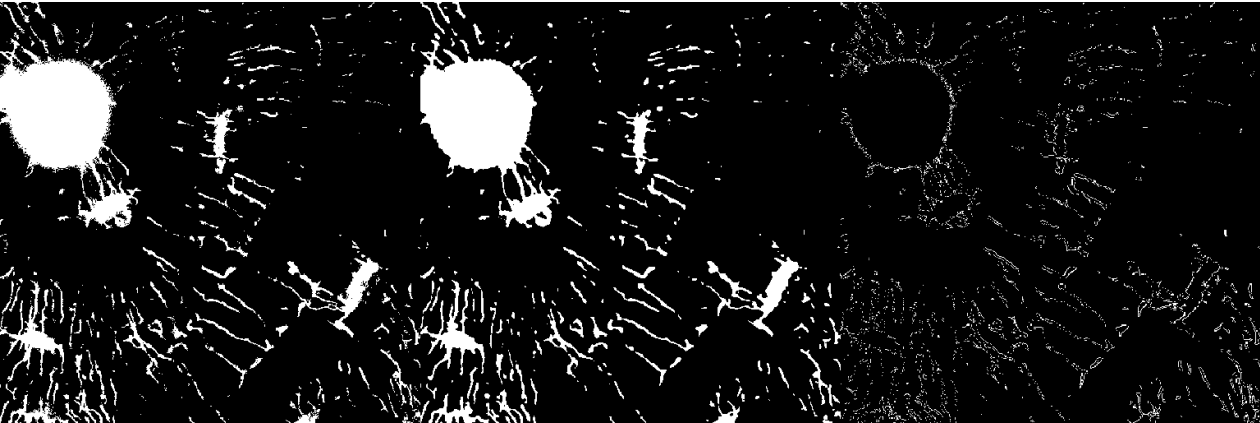
小说明：求标签图片和预测图片的差异使用PIL库中的`ImageChops.difference()`函数。该函数的输入为两张需要比较的图片，输出为两张待比较图片所有不一致的像素点构成的图片，即对两张图进行异或操作。

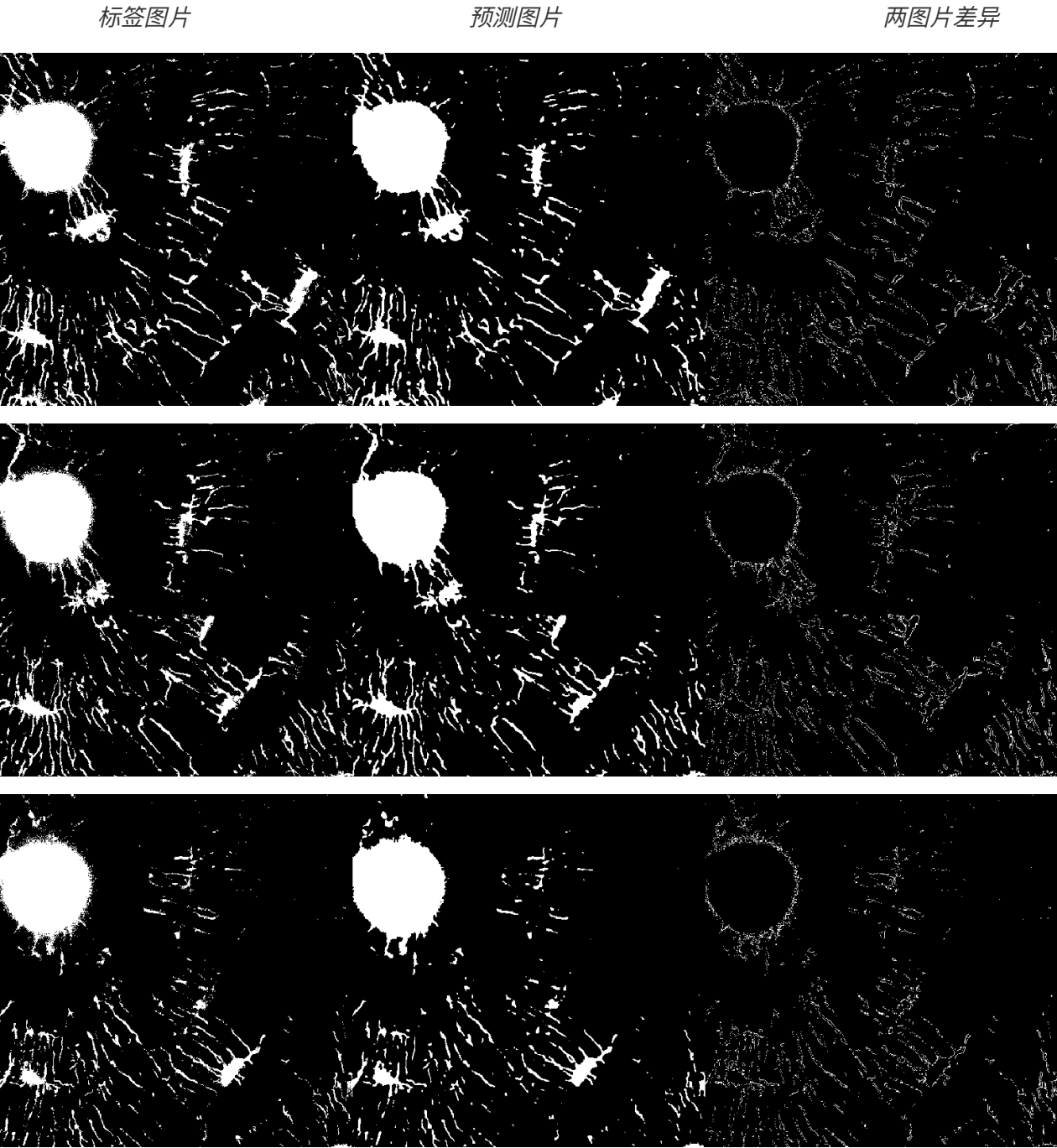


标签图片

预测图片

两图片差异

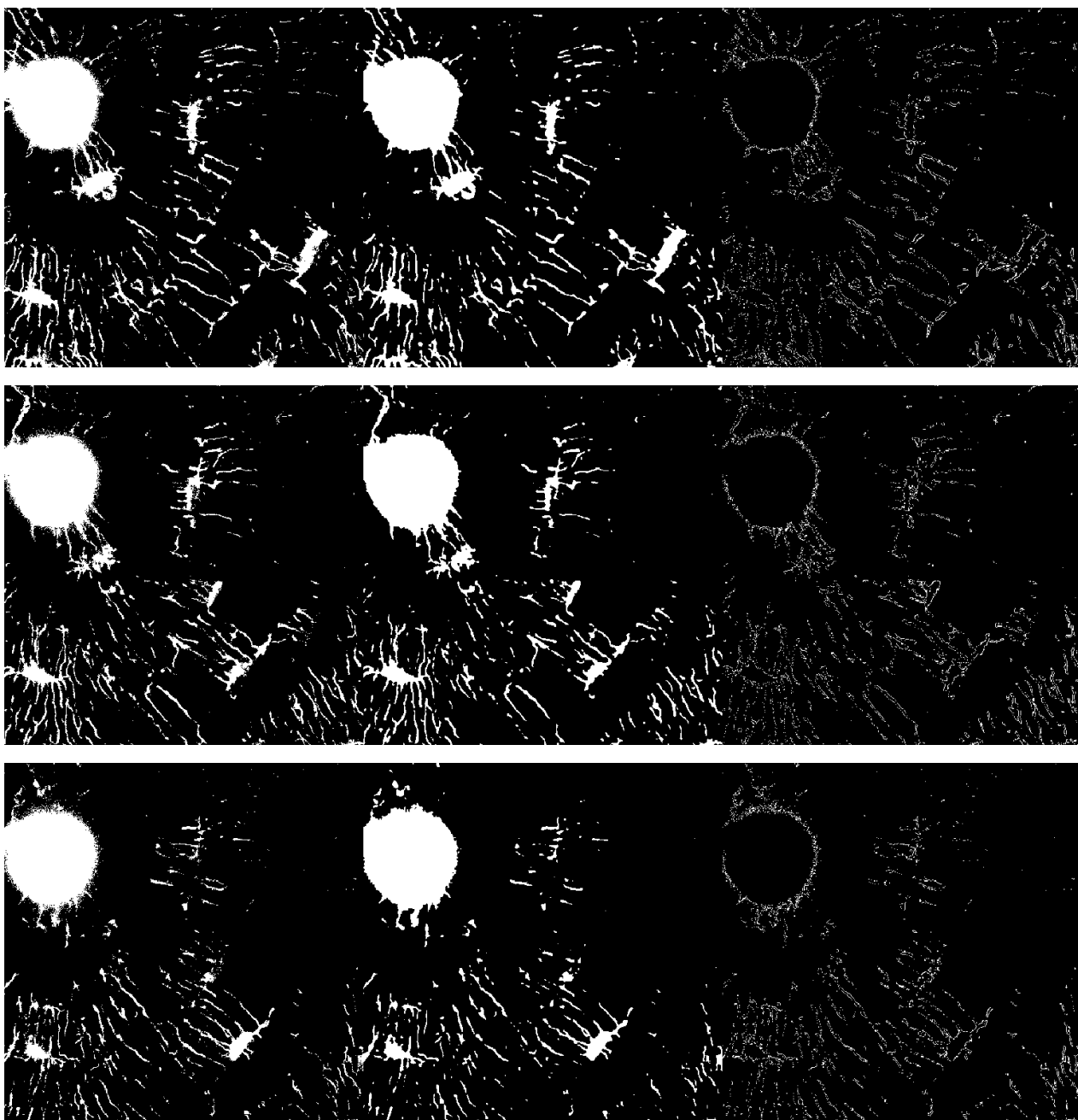




标签图片

预测图片

两图片差异



3. 不足之处的思考

- 关于第一次报告提出问题“训练集的平均误差会随着设置阈值的变大而减小，然而，设置较大阈值的导出图像明显地和标签图像不太吻合。反而经过较小阈值处理的图像能够较好地对应标签图像”的一个思考，与对应发现的问题：卷积神经网络 U_{net} 输出的图片大小相对于输入的图片尺寸有所减小（图像尺寸从原来的 512×512 缩小为 472×472 ），我们考虑是没有在神经网络第一层卷积设置合适的 $padding$ 层或步长设置不当导致的（刚开始训练的时候没有考虑到这些因素）。这就解释了之前的疑惑：当阈值设置得较大，图像中占大多数的黑色更多，而骨细胞核等主体结构依然为白色，这与标签图片重叠较多。反过来对比视野缩小的预测图片，相当于在有限的画布上放大了中间主体部分-----也就是细胞核等白色部分，这无疑增加了白色像素点的占比，故和标签图片重叠更少，得到的验证误差更大。