2019年1月12日星期六

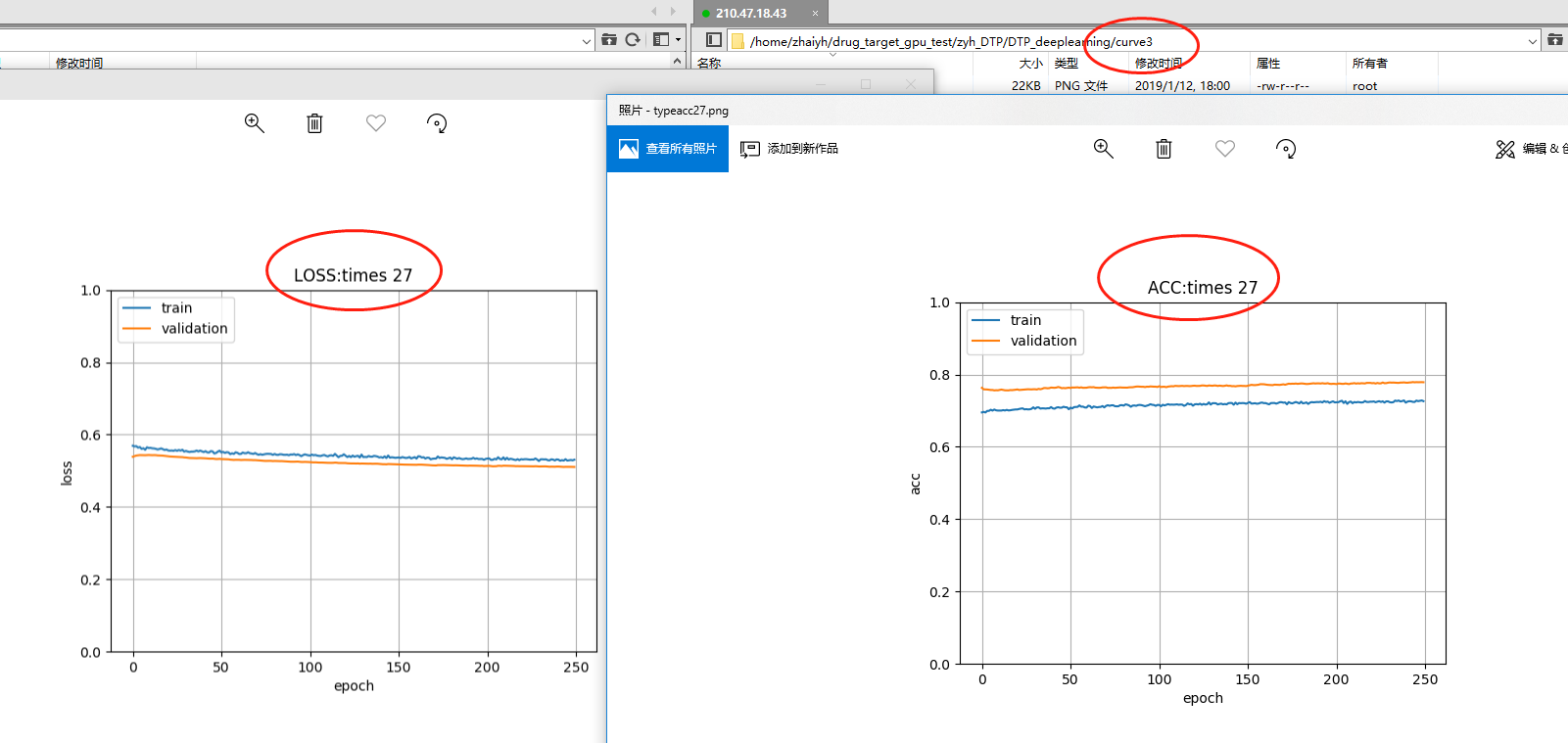
根据泳冰的建议，今天完成两部分内容：1调试batchsize和epoch的关系，查看影响。2增加recall等曲线，查看数值。

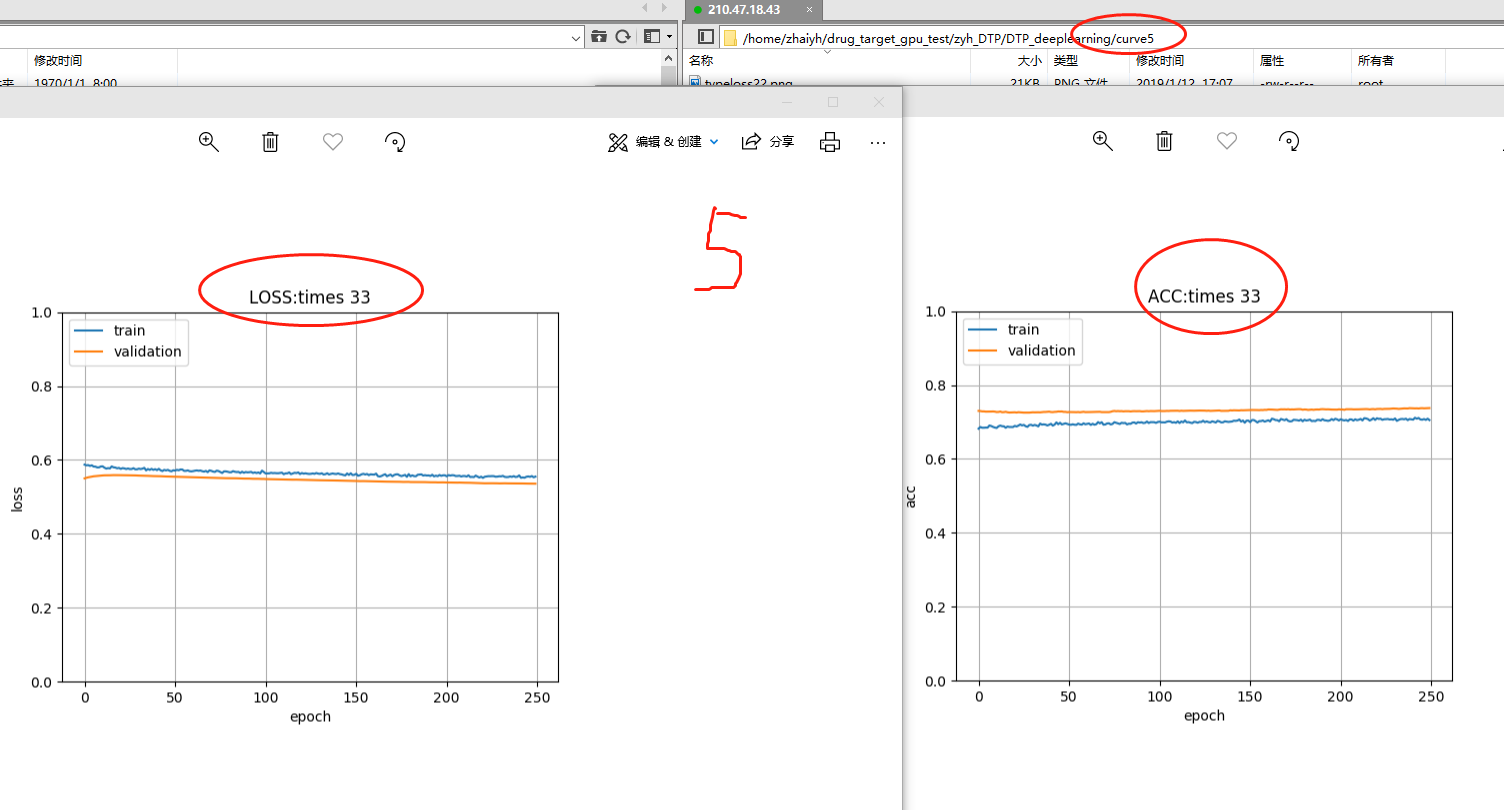
数据正负样本比为1：33。所以正样本与第一份负样本一比一进入训练器，训练epoch次，之后输入正样本与第二份负样本，训练epoch次。（这样训练了33轮 \* epoch 轮后，完成了一次训练。这种训练方法和全循环，之后再增加epoch，有什么不同么？）

1. 调试batchsize,它表示每次放入训练的数目，和机器所能承受的内存有直接关系。很显然，机器可以承受承受更大的batchsize.
2. Epoch,同一批数据训练轮次。并不是一批数据放入训练一遍就结束，而是放入很多遍（放入epoch次）。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 127 | 255 | 511 | 1023 |
| 250 | 3号 |  | 5号 |  |
| 500 |  | 2号 | 4号 | 本机 |
| 750 |  |  |  | 7号 |

对比1：训练250次的比较中，我们查看3-5的结果。





由此可见：5号在同样多的时间里训练了33轮数据，相比27轮数据速度快。

**Batchsize增大可以增大训练速度。**

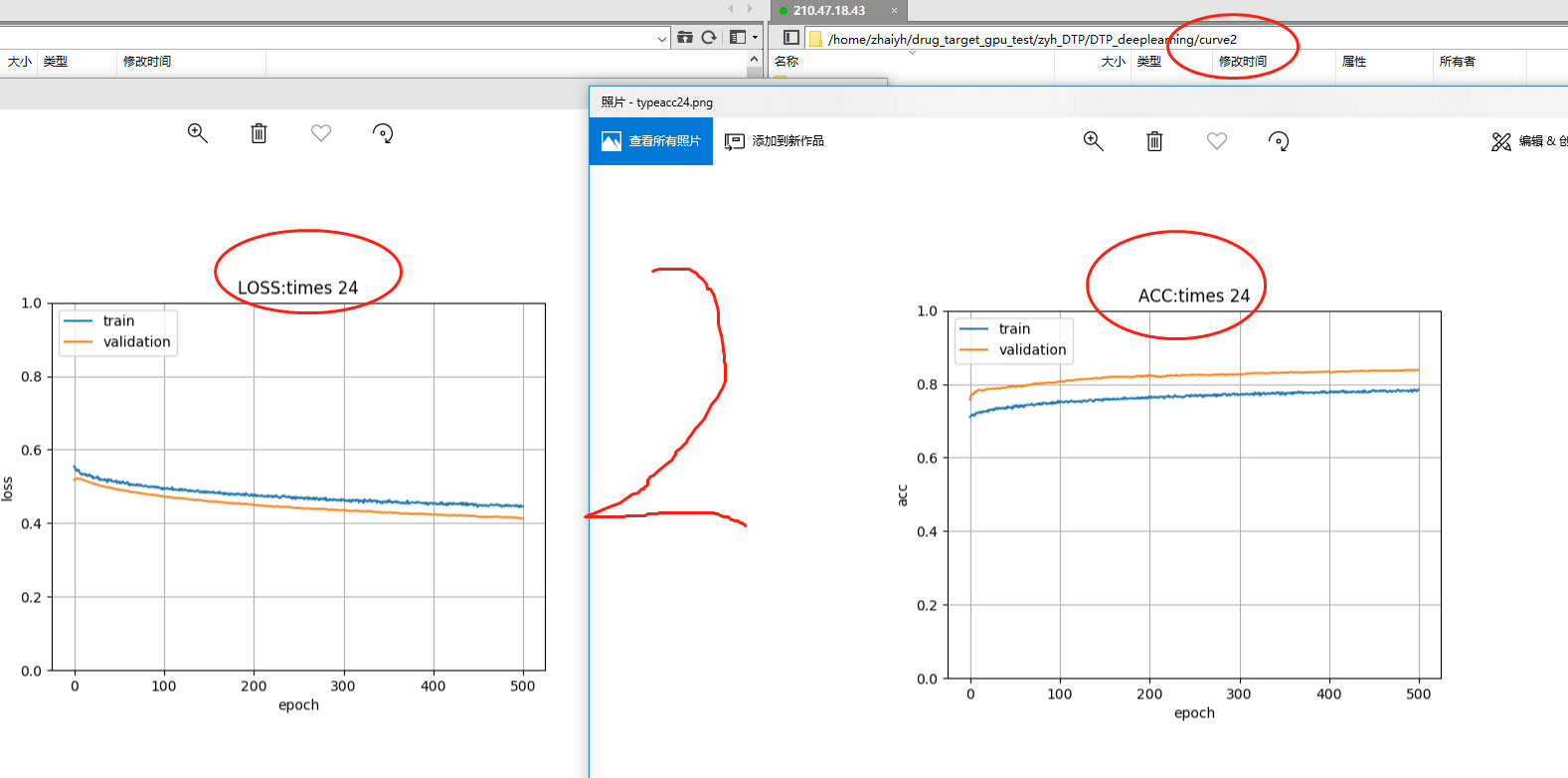
5号图像中acc与val\_acc相距较近，

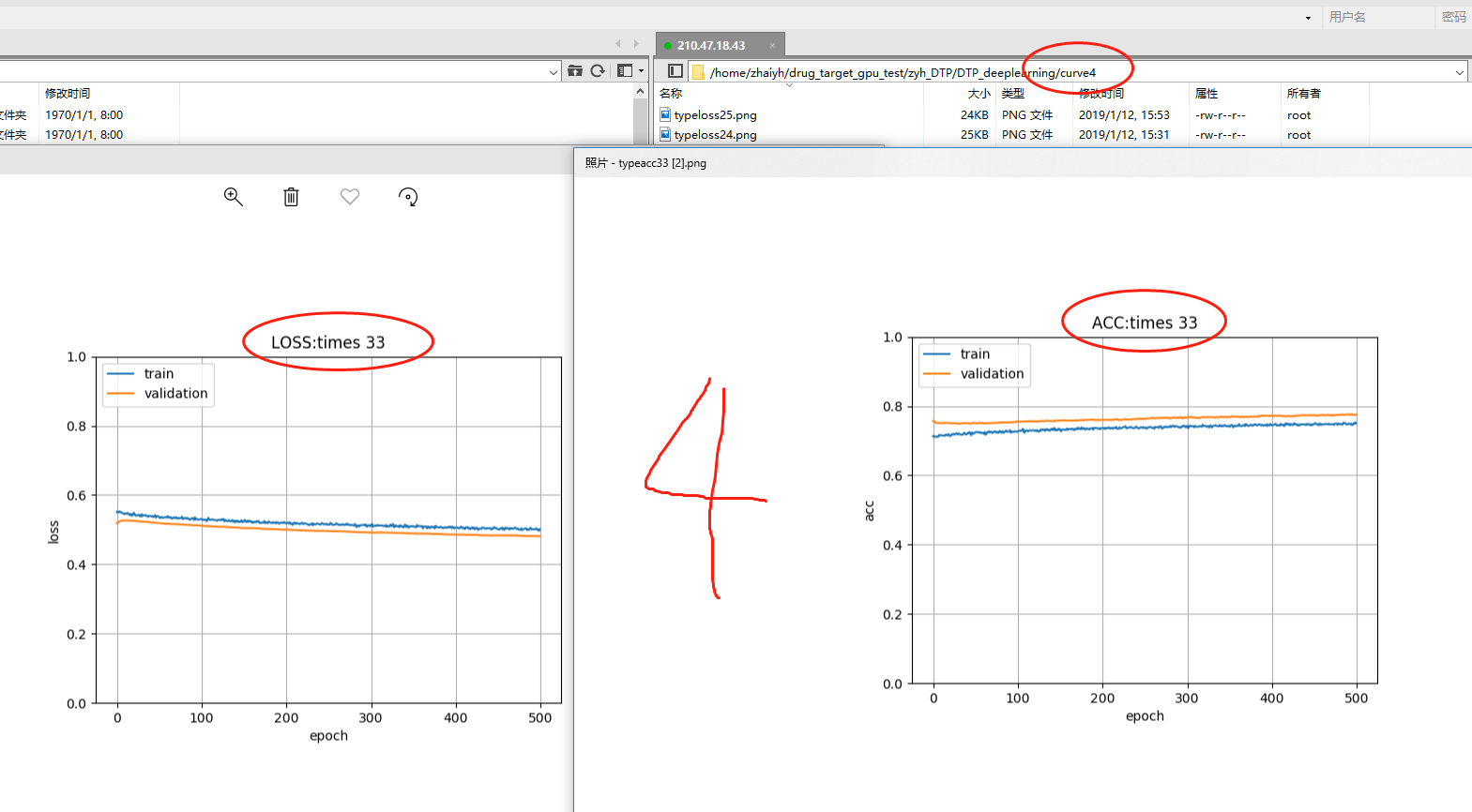
**Batchsize增大，可以降低过拟合（acc中val比train大，说明过拟合？）总之越近越好**

3号图线val\_acc更高,loss更低

**Batchsize减小可以获得更高的val\_acc，可以获得更低的loss**

对比2，我们在训练500次的比较中，我们查看2-4的结果。





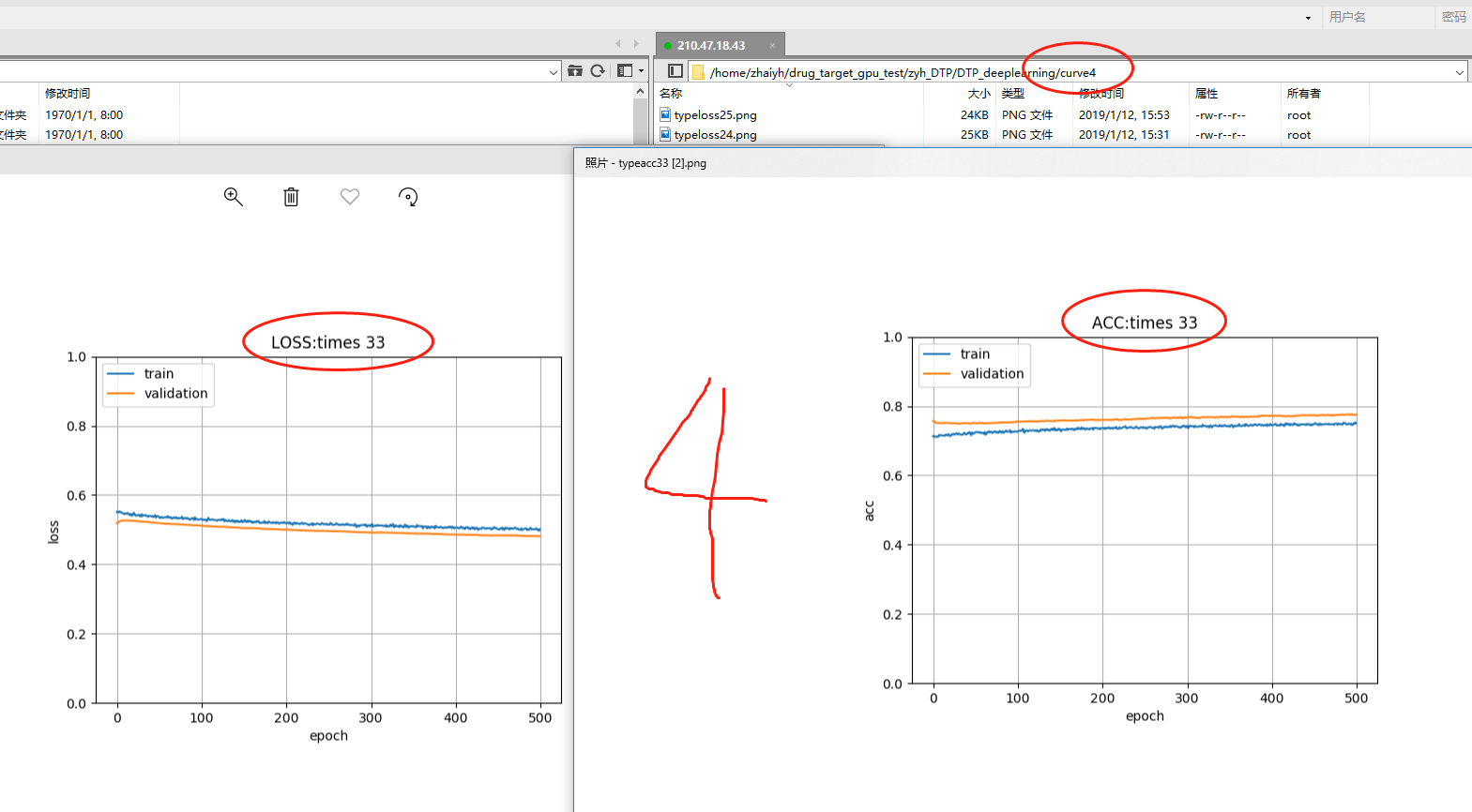
1. **batchsize增大可以训练更多的数据。这条成立。**

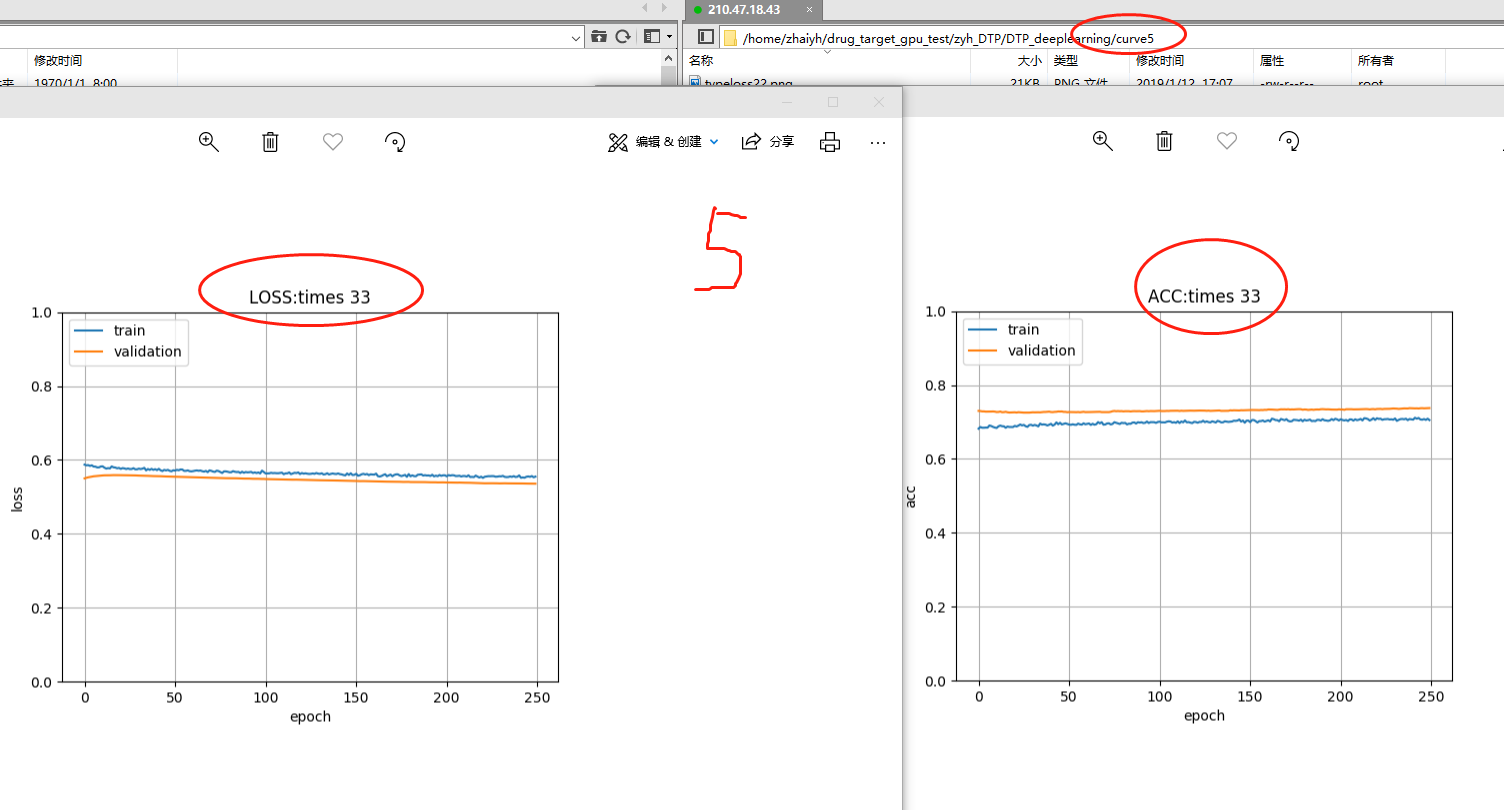
**2，batchsize增大，acc与val\_acc拟合更接近。**

1. **batchsize减小，可以获得更高的val\_acc,成立。(第一次上80%)**

**综上对比，我们发现，batchsize减小，从某个角度来说，会获得更好的acc.**

**Batchasize相同，epoch不同，可以查看4-5对比。**

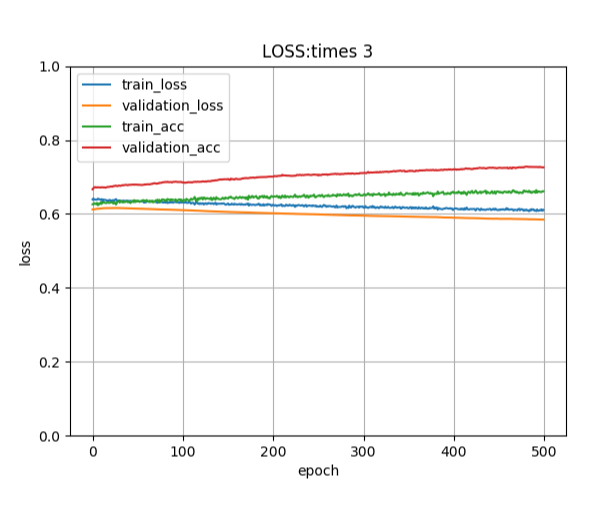


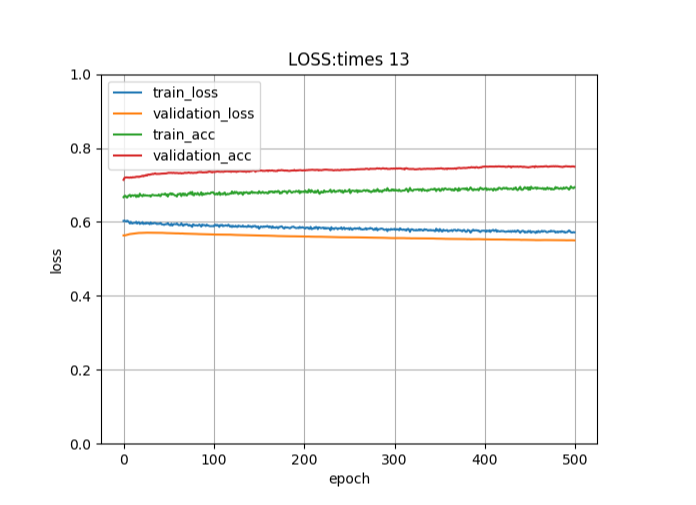


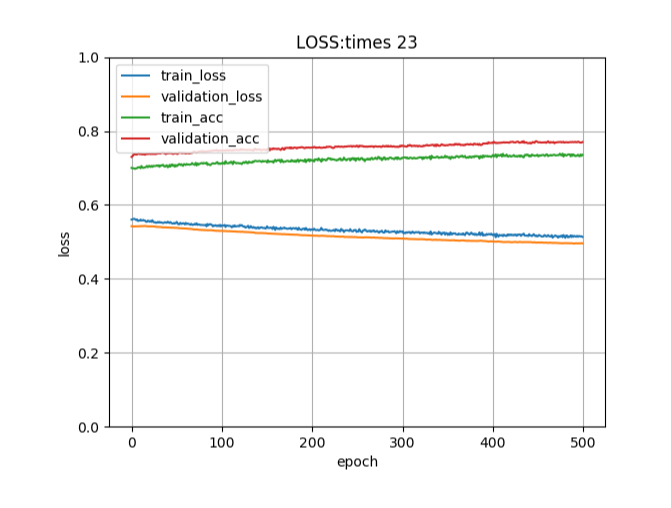
**大体趋势相同，4号稍微高一点。说明，增大训练轮次，看起来并没有巨大的改变。**

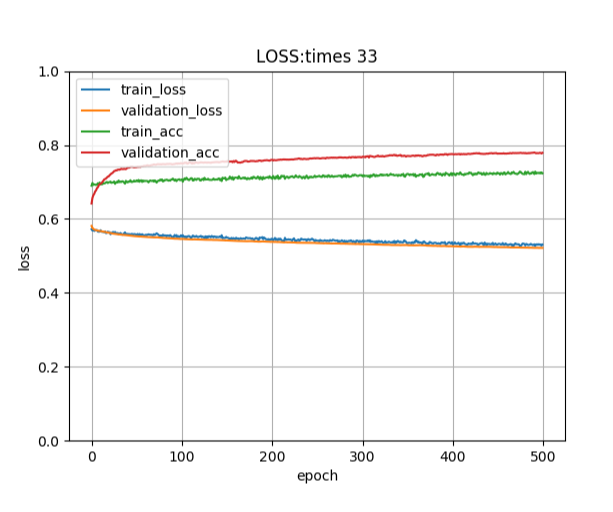
**以上是前几个实验的对比结果。**

**实验6，**



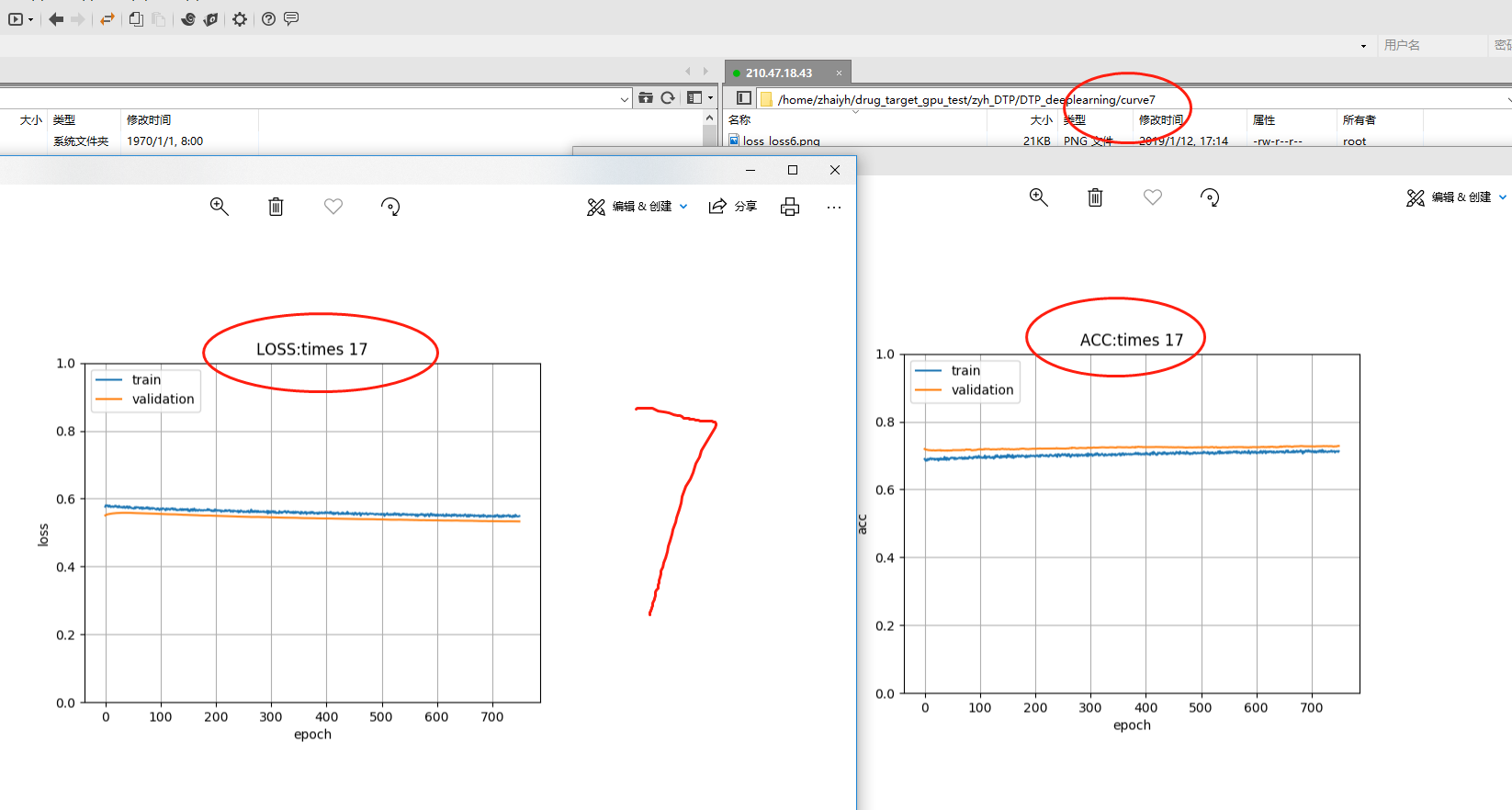


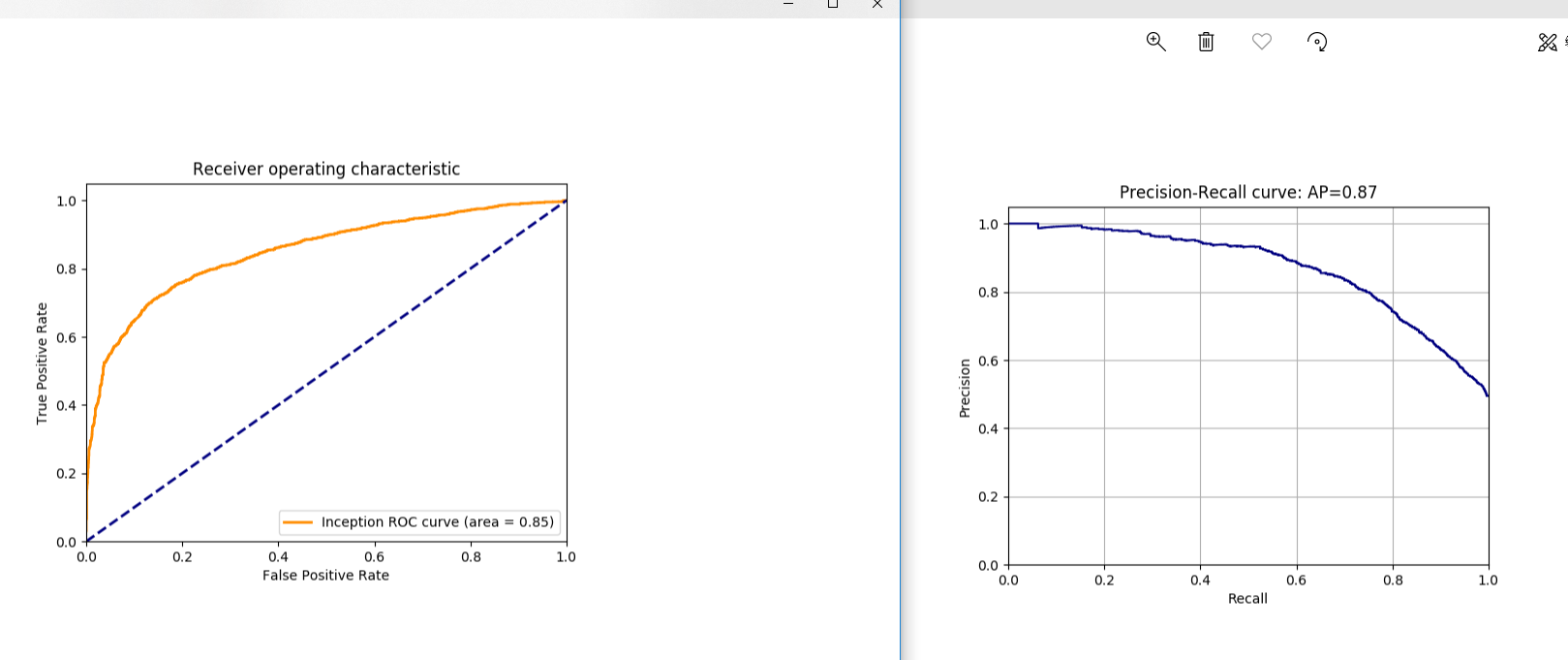


**画在一起了。**

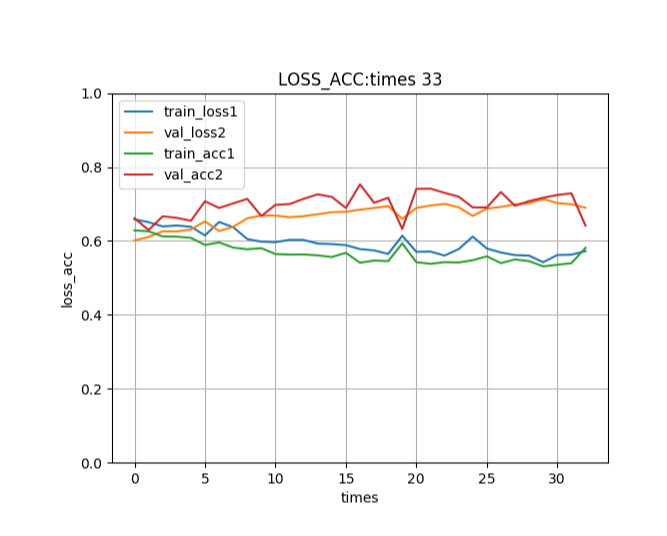
**设置有点忘记了。**

**7号**





**用于画这个，还不会看。**



**该图显示的33次的33个模型的起始训练值（应该画最后值）**

**pre\_score, [0.5720520134375654, 0.6921187786285194] [0.6180682745355665, 0.645679841538736]**

**pre, 0.9780907668231612 0.97911227154047**

**rec, 0.2822944896115628 0.16937669376693767**

**SN, 0.2822944896115628 0.16937669376693767**

**SP, 0.9938189845474613 0.996467991169978**

**f1, 0.4381352961794602 0.28879476318829417**

**mcc, 0.39472485205139834 0.29651393484122357**

**roc\_auc， 0.8184598910006762 0.7835121073800878**

**loss: 0.5486 - acc: 0.7127 - val\_loss: 0.5329 - val\_acc: 0.7281**

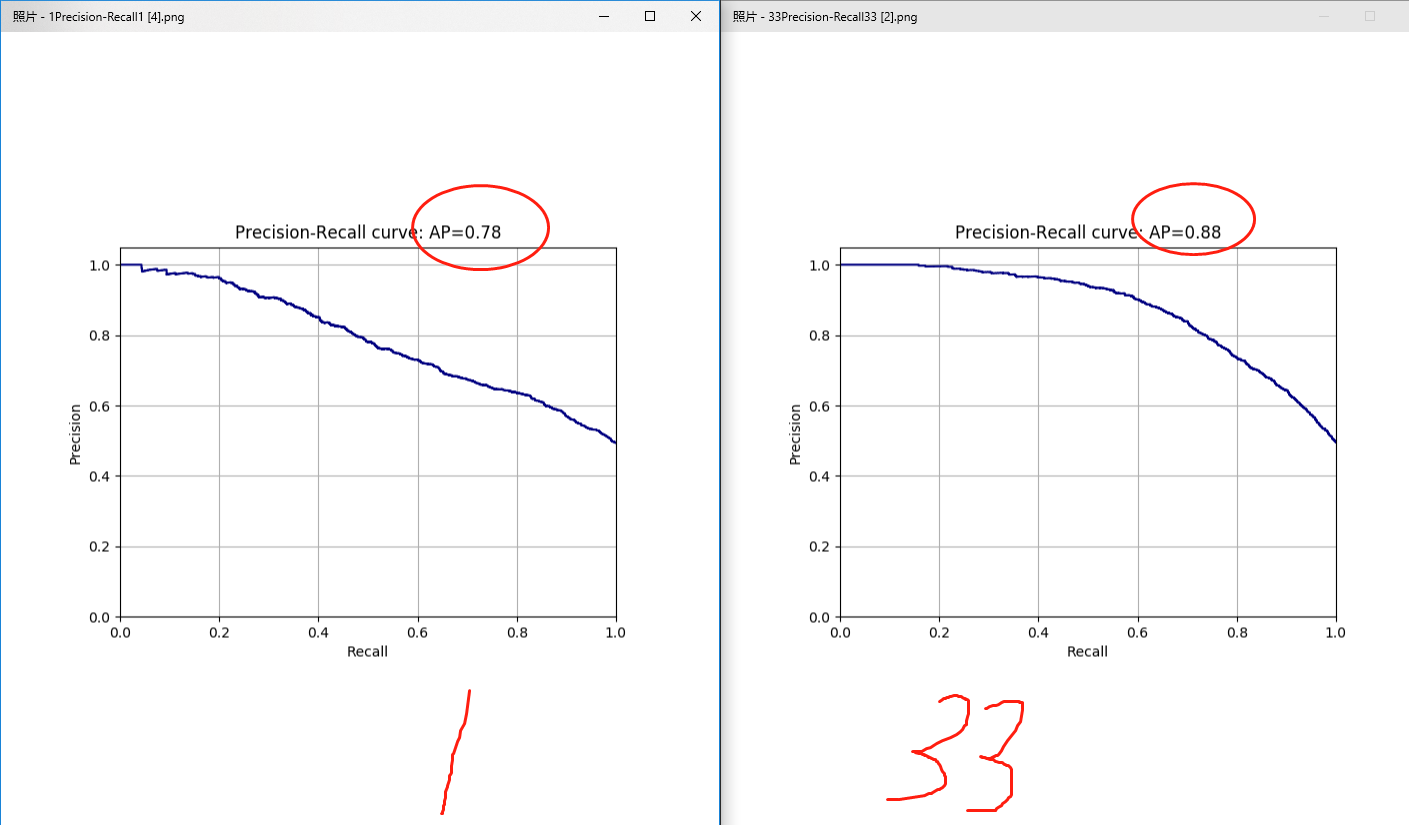
**pre\_score, pre, rec, SN, SP, f1, mcc, roc\_auc**

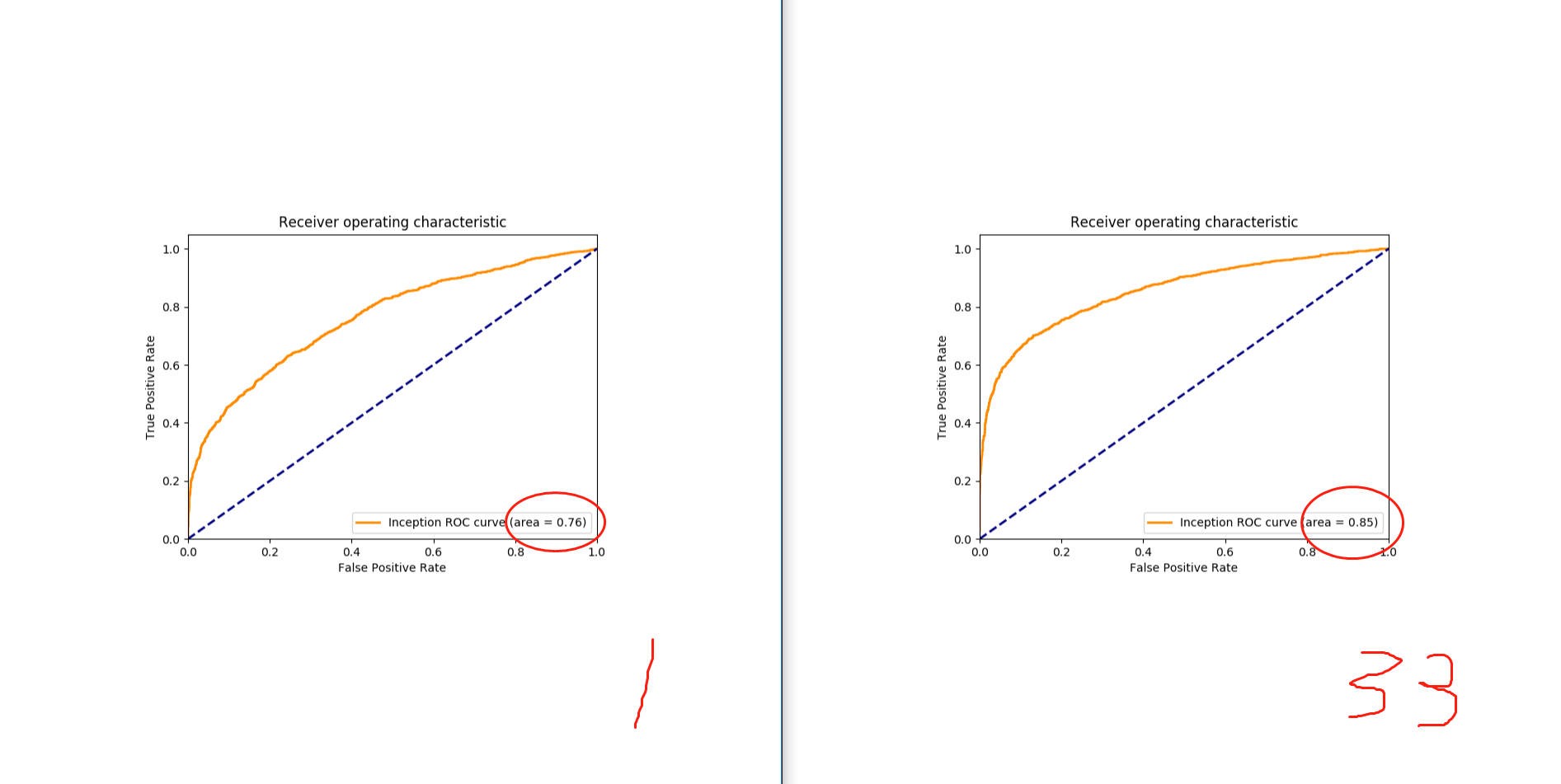
**[0.5329182726260566, 0.7280642987495289] 0.9801980198019802 0.35772357723577236 0.35772357723577236 0.9929359823399558 0.5241561879549966 0.4559431777891813 0.8536756861314015**

**18**

**1月14日8：30记录**

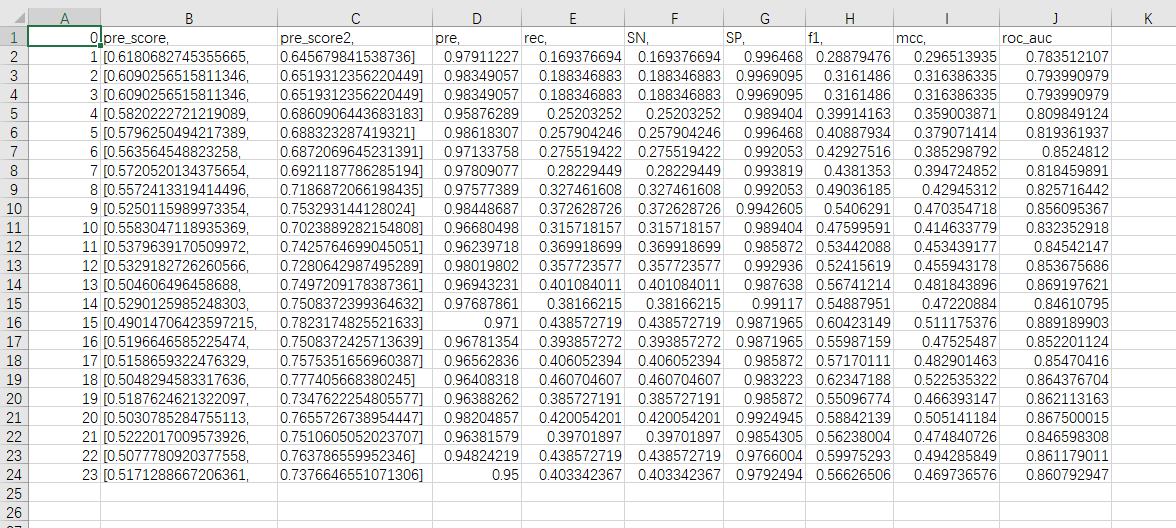
**7号**

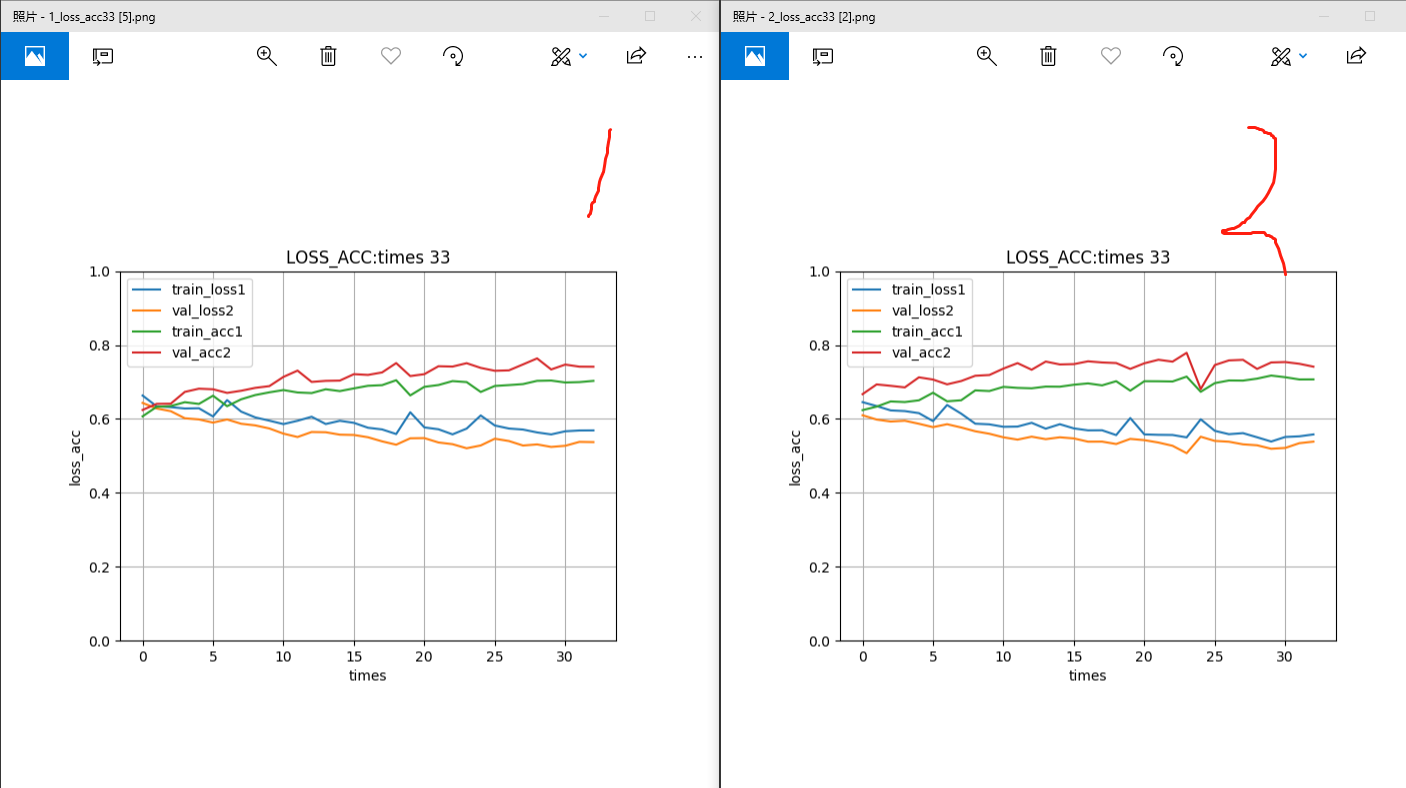


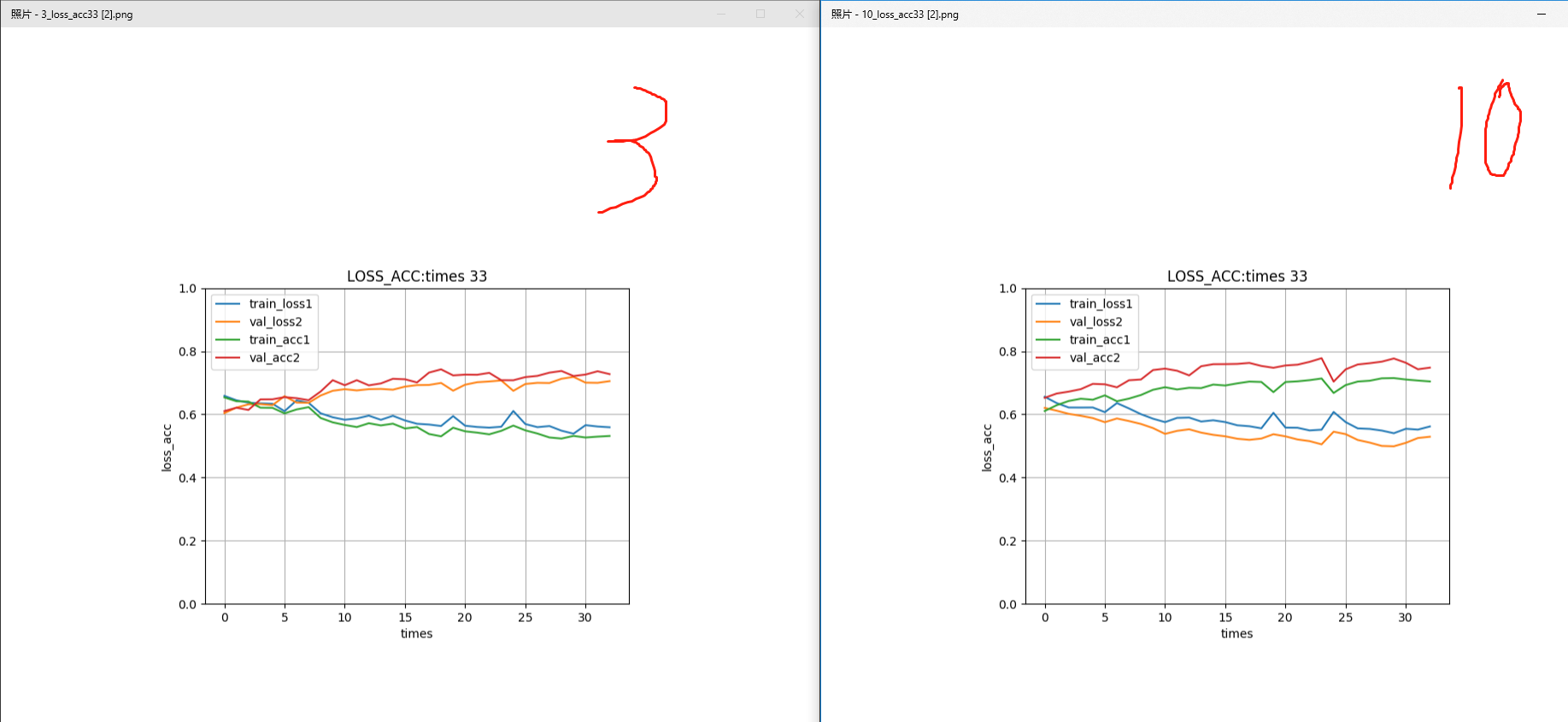


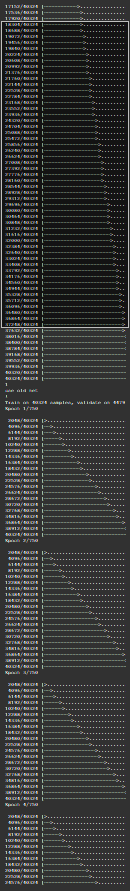
**泳冰说：需要找对比，找其他模型用同一种数据进行比较。**

**从1-33之中的部分数据**









**发现第一次编译网络时，设置默认值batchsize值为128，和传入参数不同。**