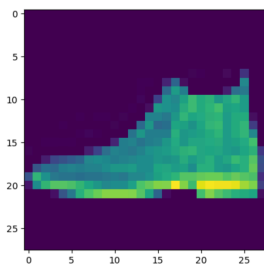


测试

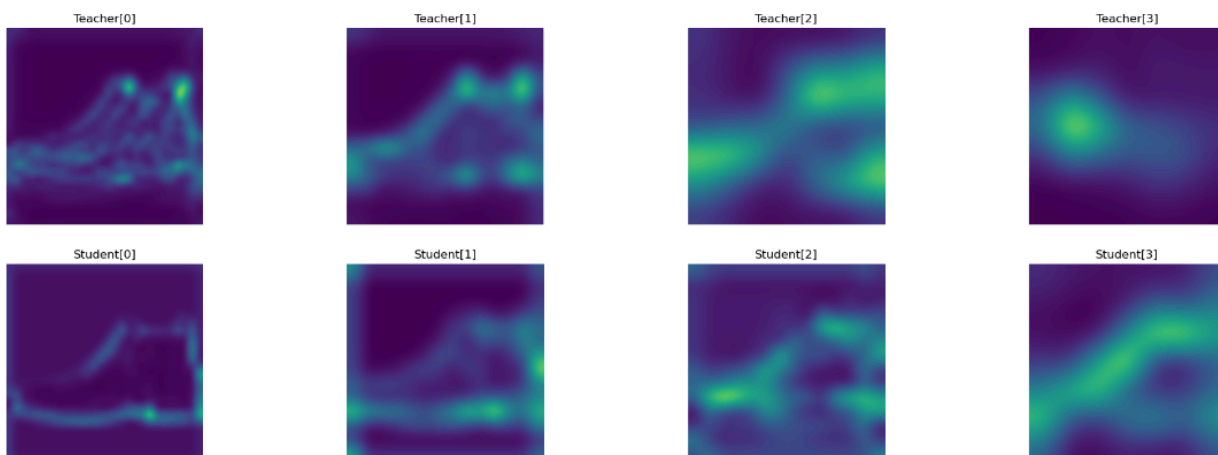
测试CNN注意力

下面是我自己的模型上，使用用 $p=2$ 之后 $\text{mean}()$ 处理：
在Fashion-MNIST

- 原图



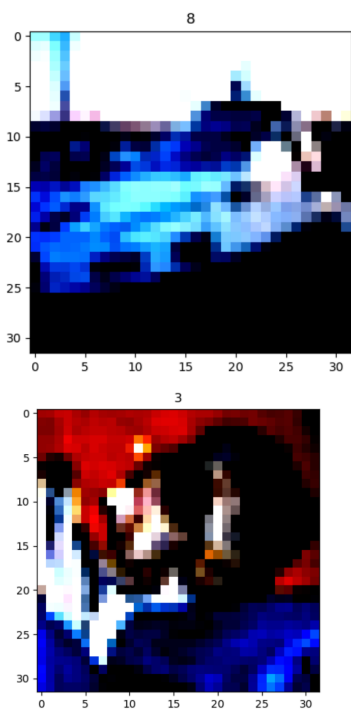
- 教师和学生的处理



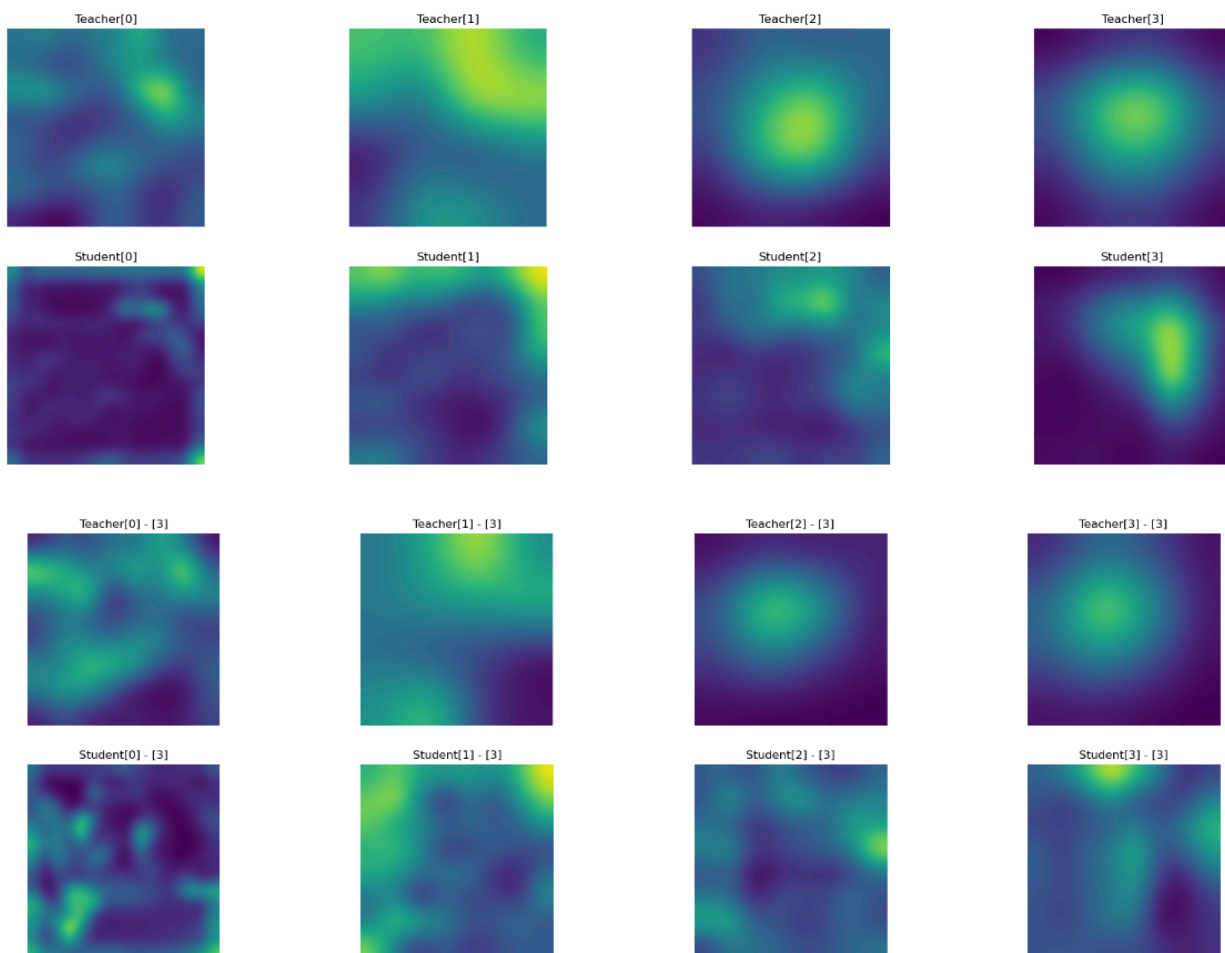
其中教师网络在最后已经收敛到一个点，实际上已经做出了决策
学生网络由于还停留在中间隐藏，还处于关注的状态

在Cifar10

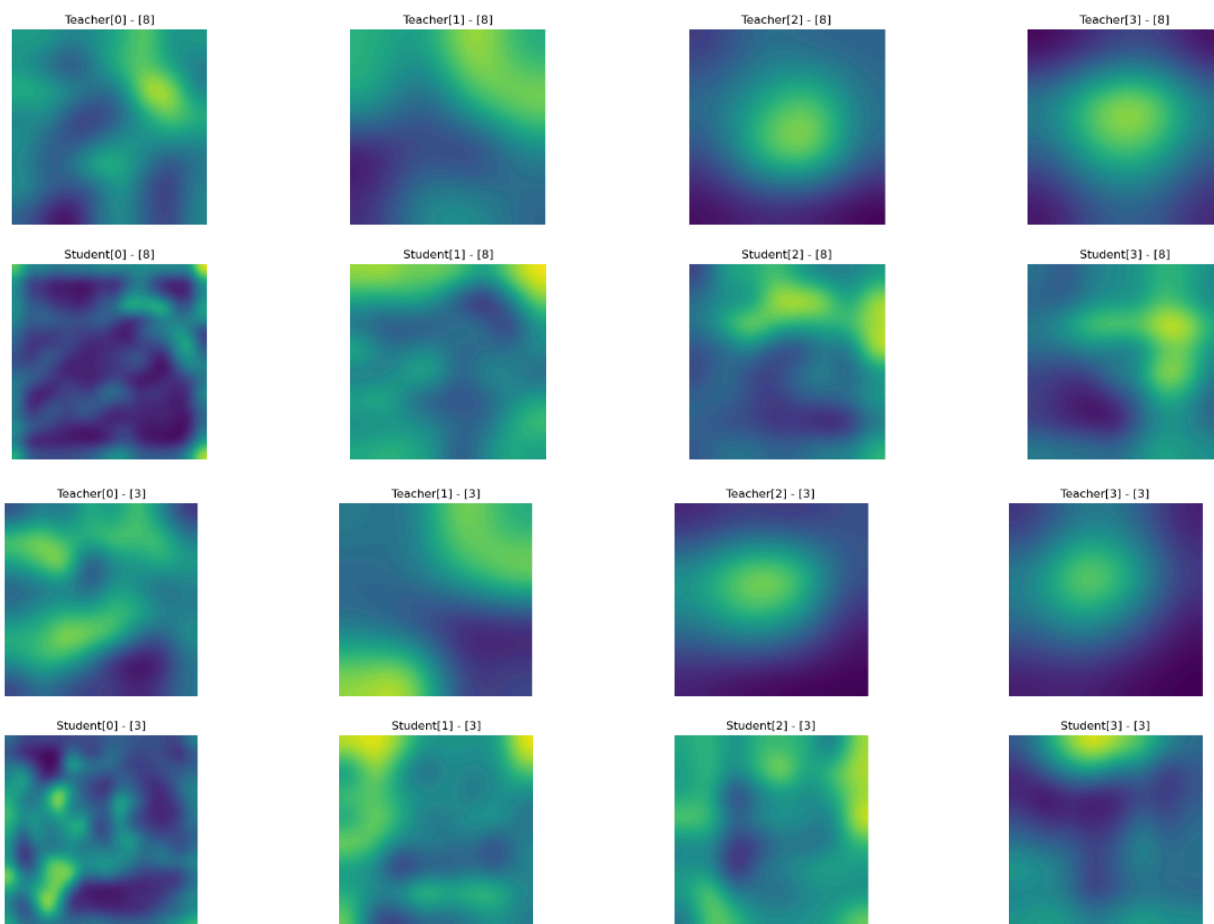
- 原图



- 教师和学生的pow(2).sum(dim=1)处理



- 教师和学生的sum处理



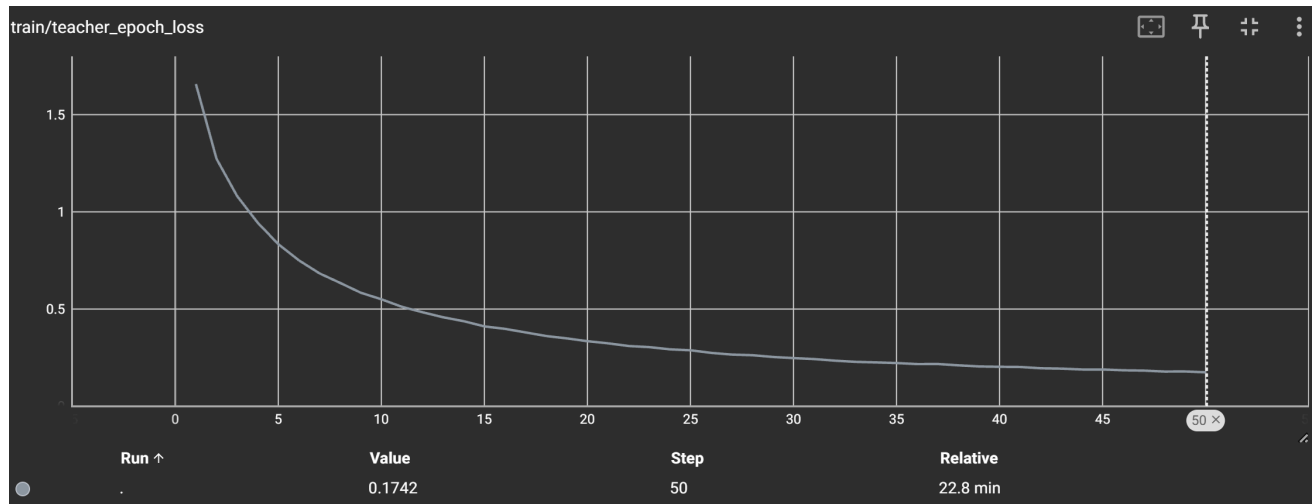
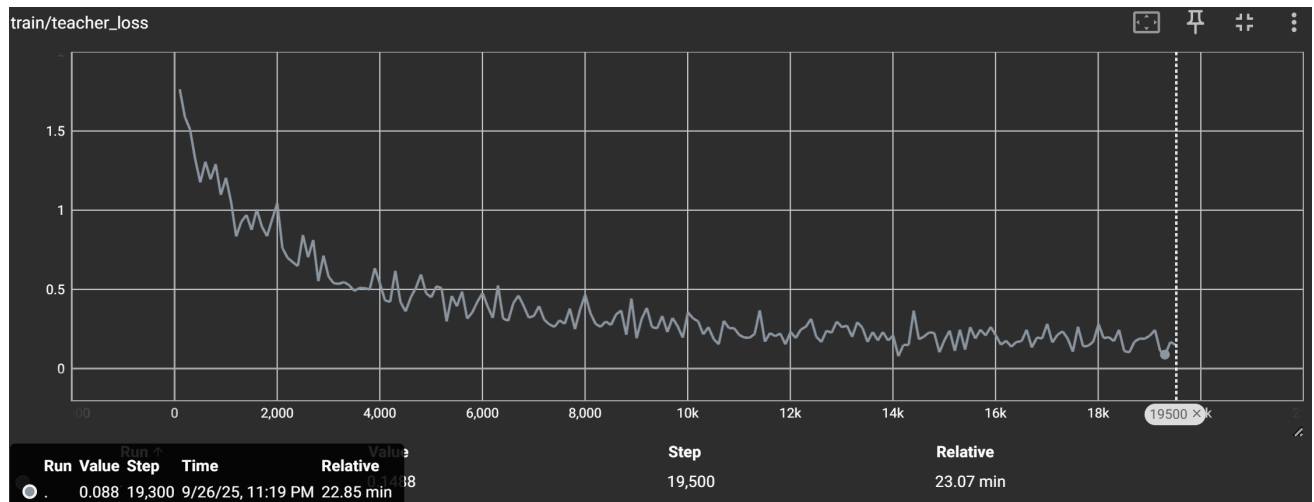
我们发现，到后面，实际上已经脱离了注意力的范畴了
 一方面，注意力关注于边界，并不是物体
 一方面，后面的图过于抽象，不能使用注意力来抽象理解了

项目

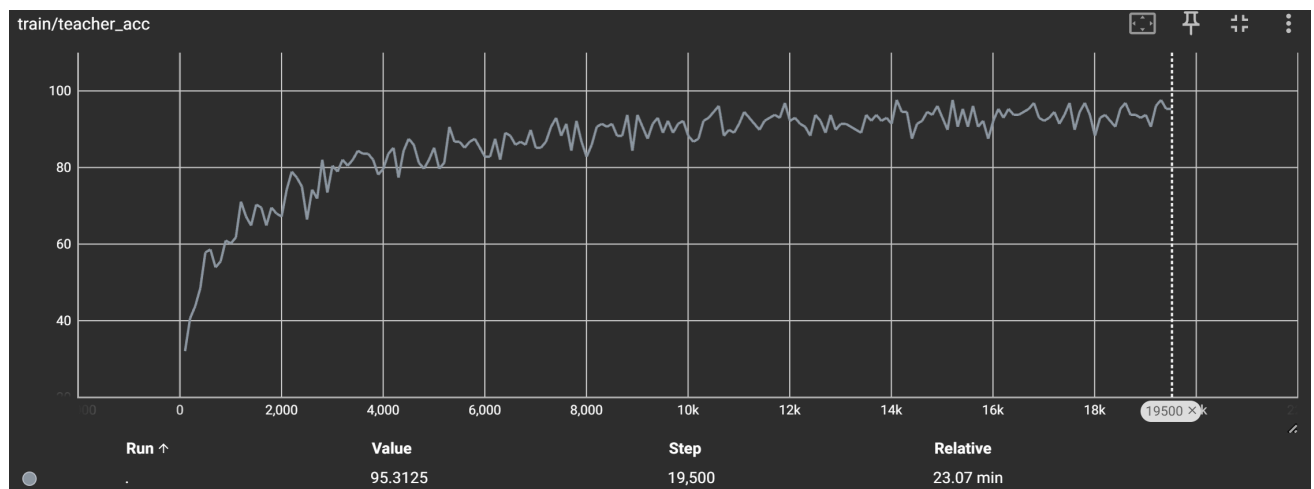
模型	教师	基线	学生 (AT+KD)	学生 (KD)
参数量	2243546	691674	691674	691674
训练损失	0.174	0.23	1.75	
训练准确	94.02%	90.14%	90.15%	
AT损失	-	-	1.17	
验证准确	89.91%	88.40%	88.19%	

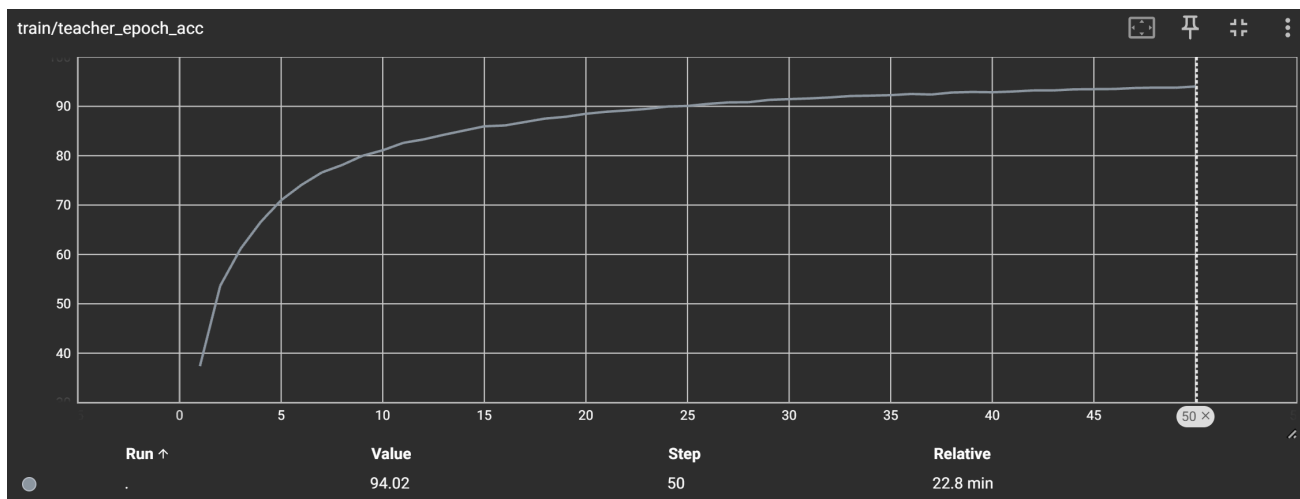
教师

- 损失



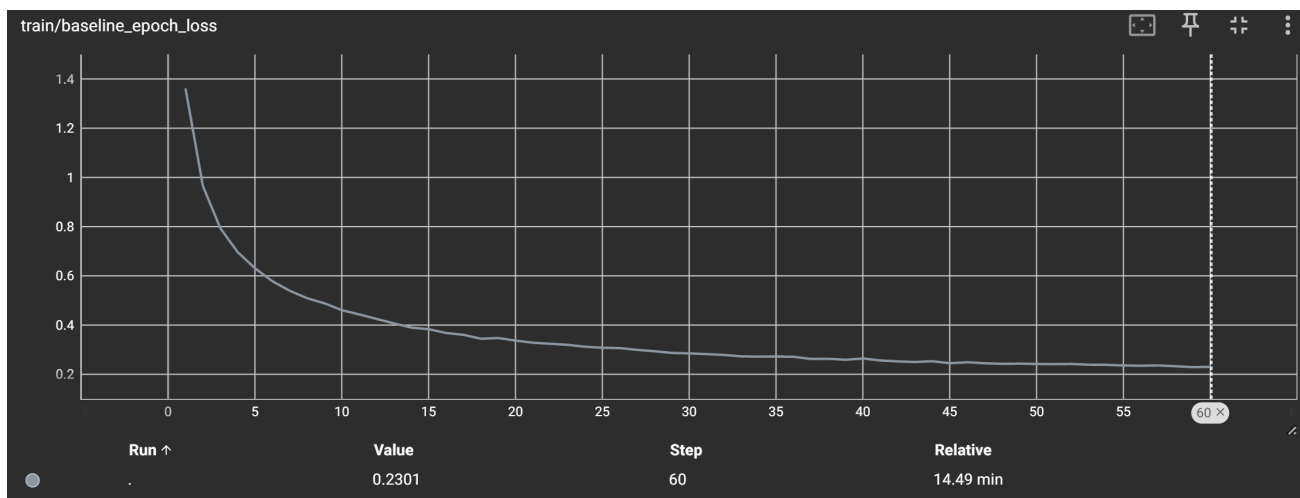
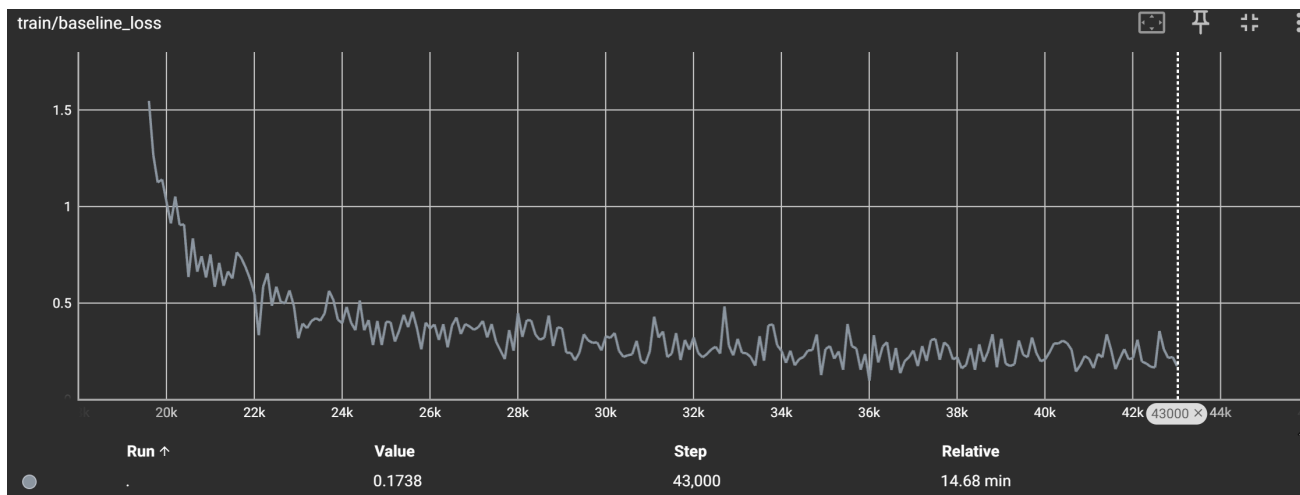
- 准确率



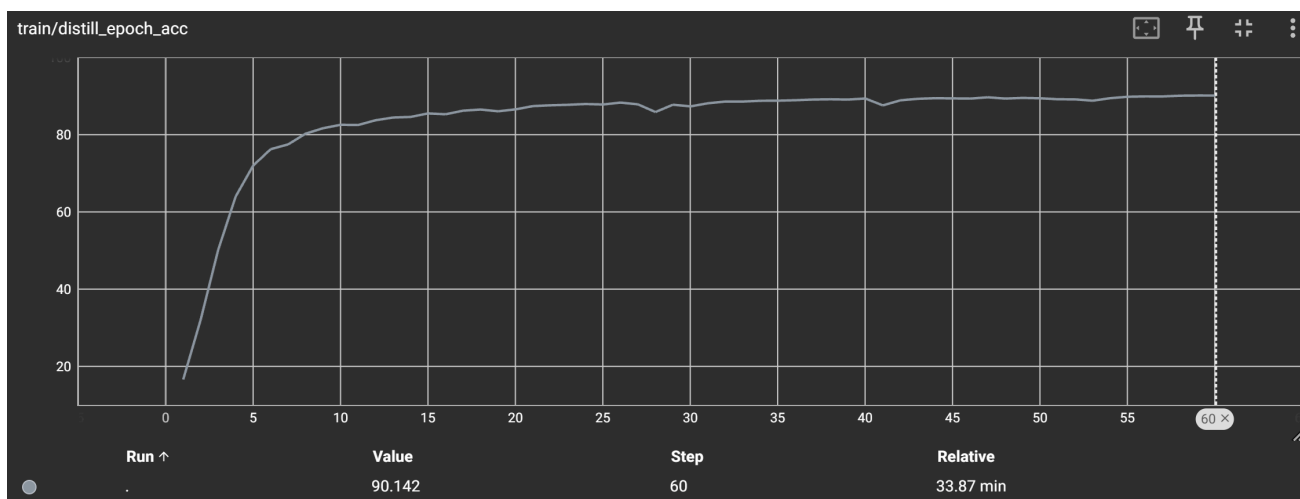
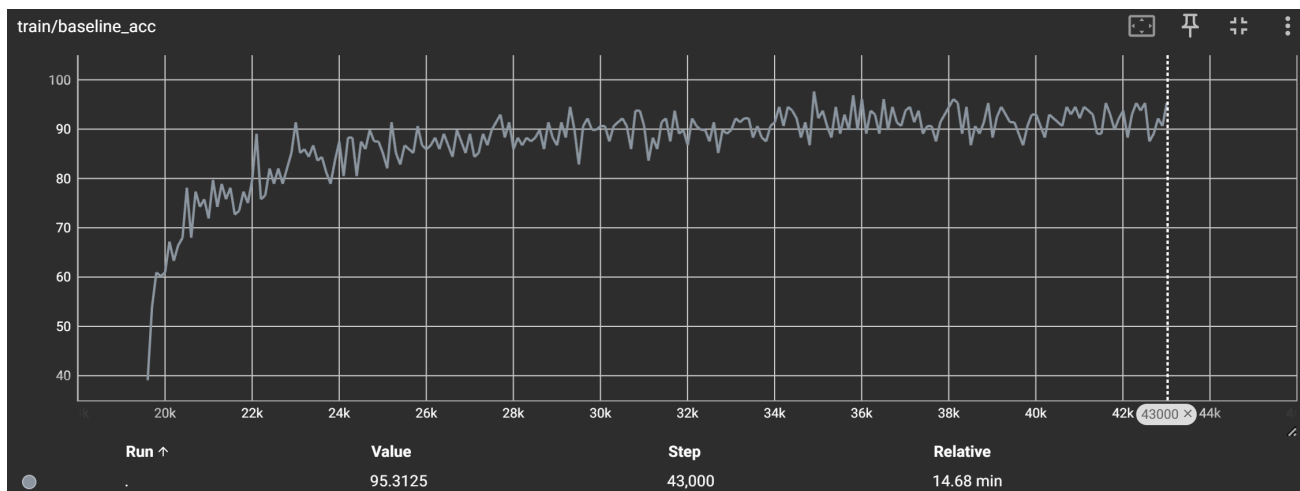


基线

- 损失

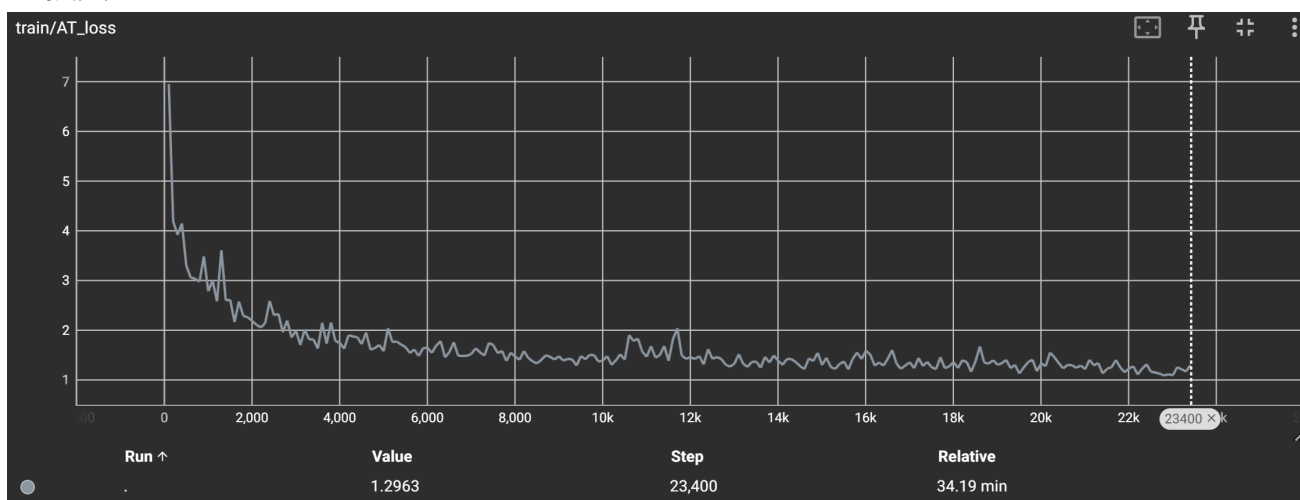


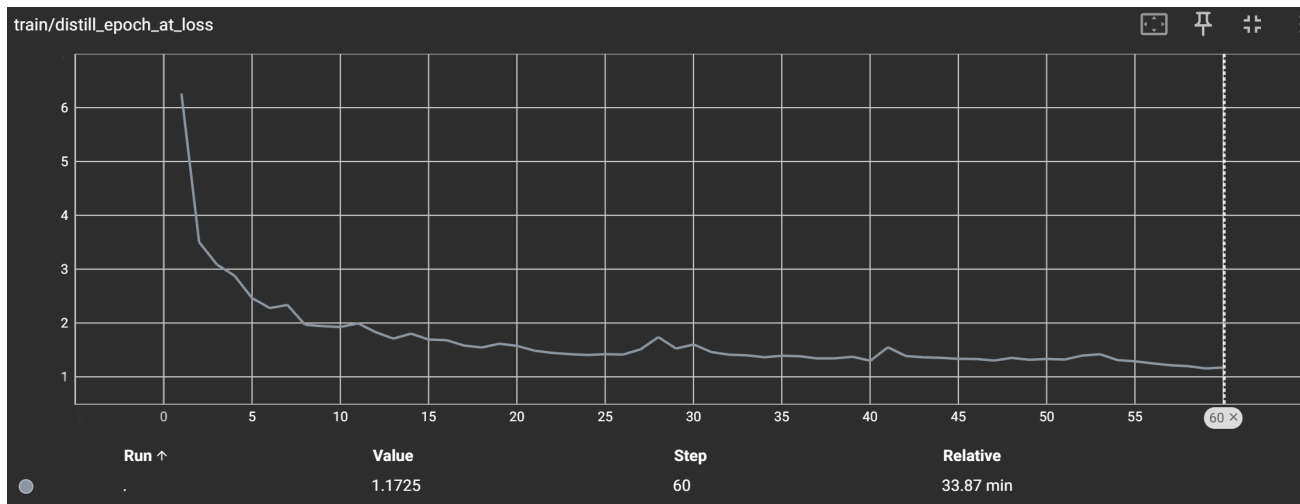
- 准确率



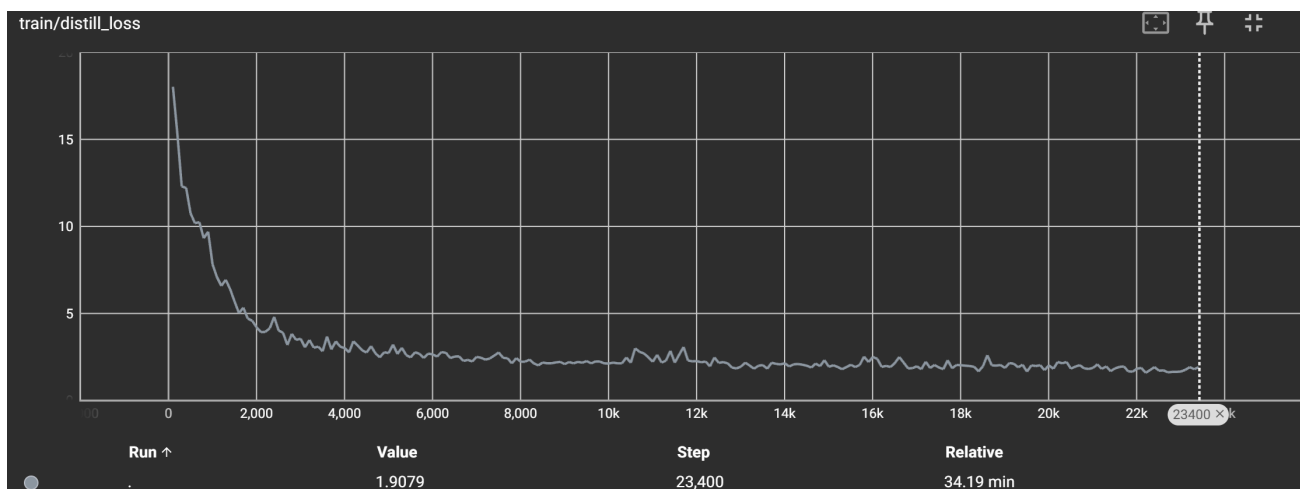
学生

- AT损失

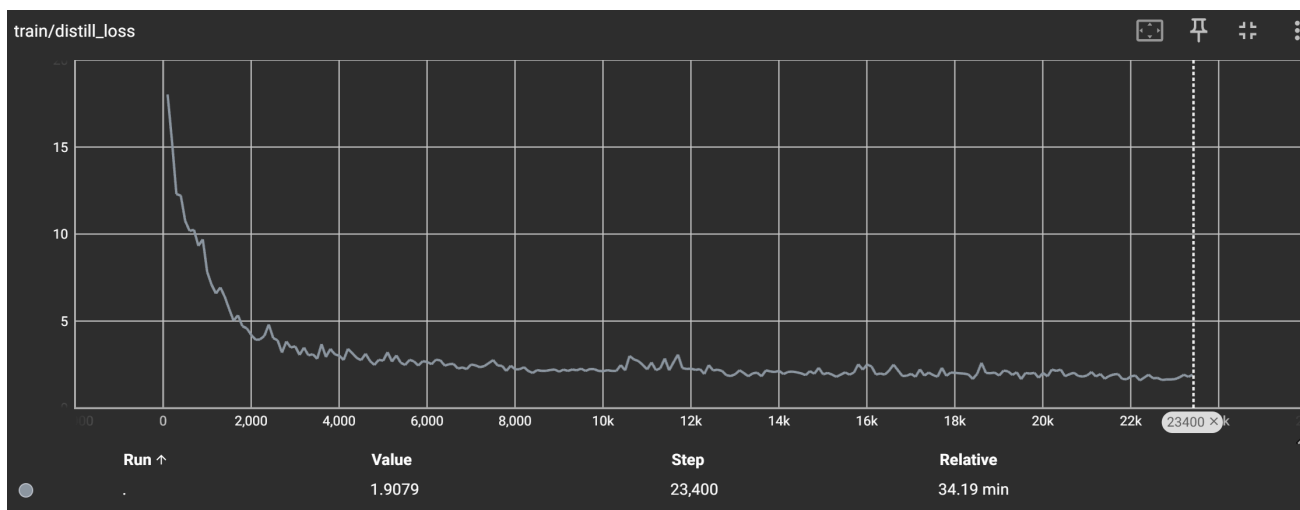


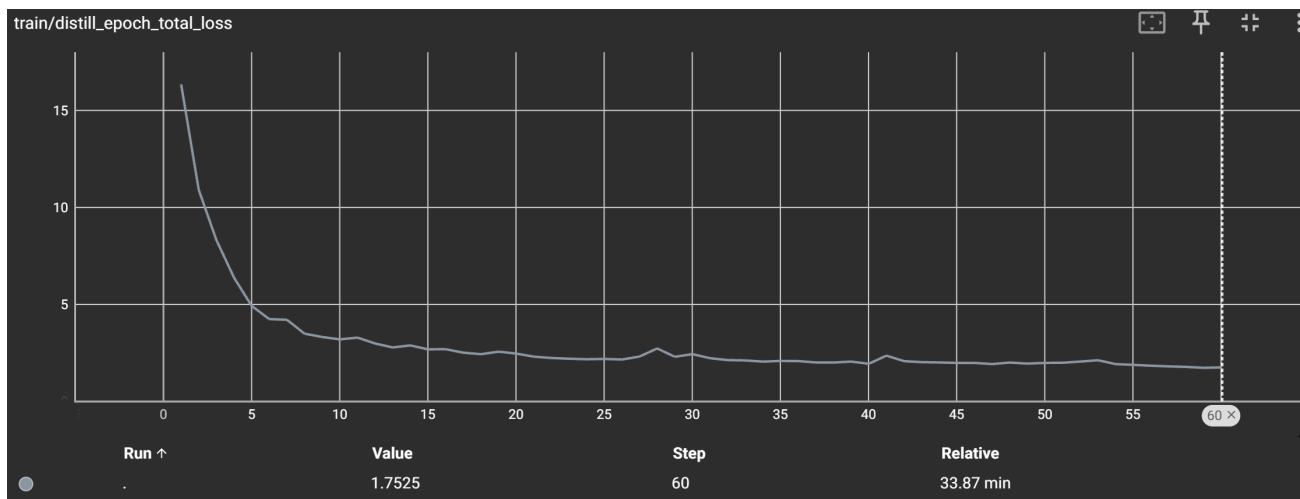


- 蒸馏损失

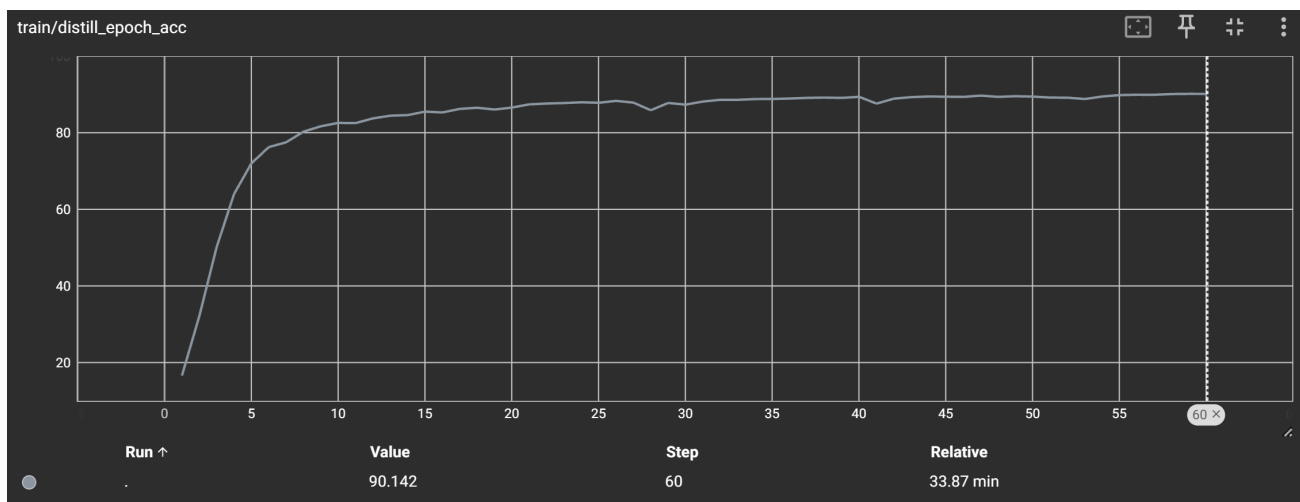
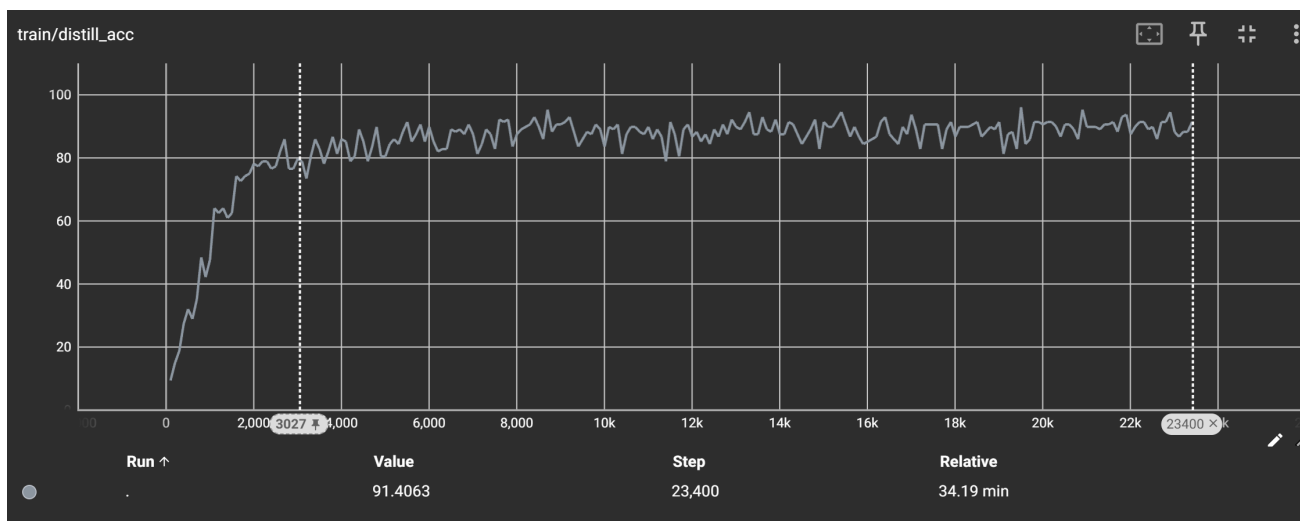


- 总损失





• 准确率



验证

