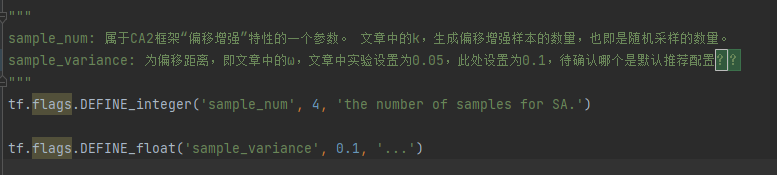
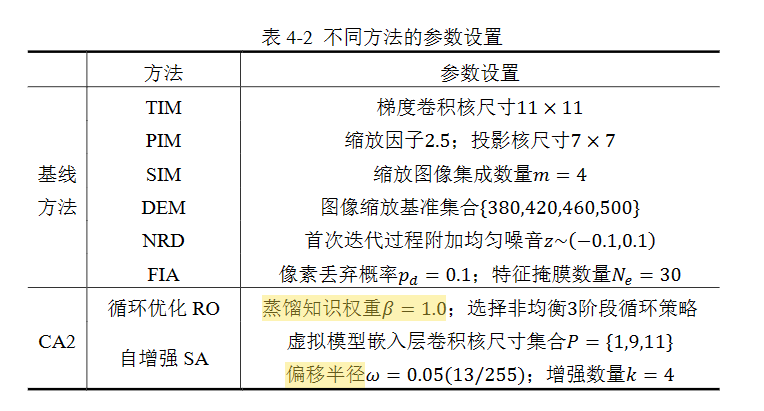
## 【OK】问题1

自增强策略中有一个参数ω（偏移距离），文章实验章节推荐配置为0.0.5，但是CA2 Github项目代码中配置为0.1。

哪个是最终的推荐配？

|  |
| --- |
| **黄博：**  **因为论文里归一化是[0,1]，所以参数0.05**  **代码里图像归一化是[-1,1]，所以是0.1，跑代码用0.1** |

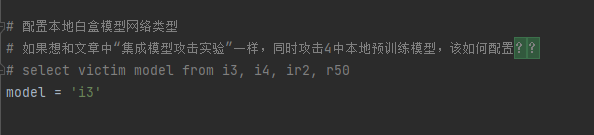


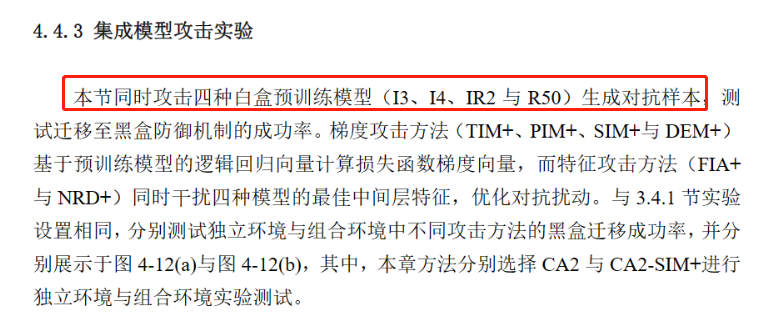


## 【OK】问题2

如果想和文章中“集成模型攻击实验”一样，同时攻击4中本地预训练模型，该如何配置？

|  |
| --- |
| **黄博：**  **同时读取4个模型，然后计算4个模型的loss值，求平均值后反向传播计算梯度。**  **参考[注]这个代码**  **https://github.com/dongyp13/Translation-Invariant-Attacks/blob/master/attack\_iter.py** |





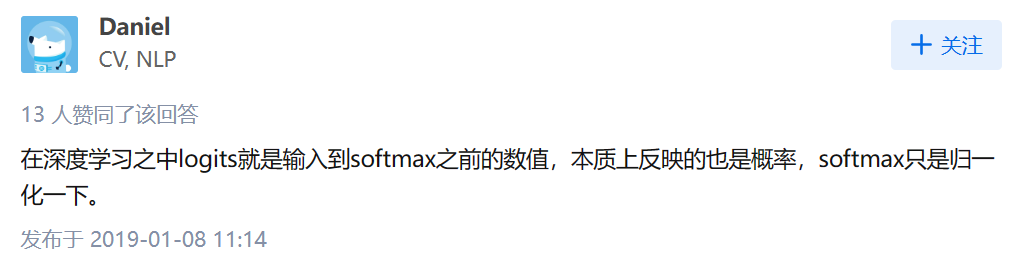
[注]:

补充参考的github项目相关代码段截图



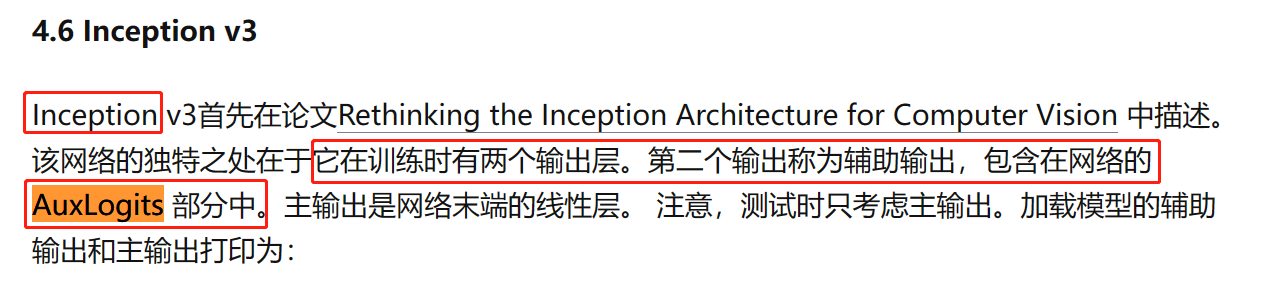
**进一步补充理解 logits & auxlogits**

*logits: <https://www.zhihu.com/question/60751553>*





*auxlogits: <https://zhuanlan.zhihu.com/p/349947340>*



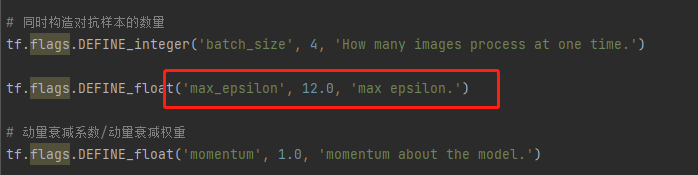
**根据以上信息，可以基本解释清楚，为什么对于Inception类模型在计算loss时，存在两个`logits`输出用于计算loss。**

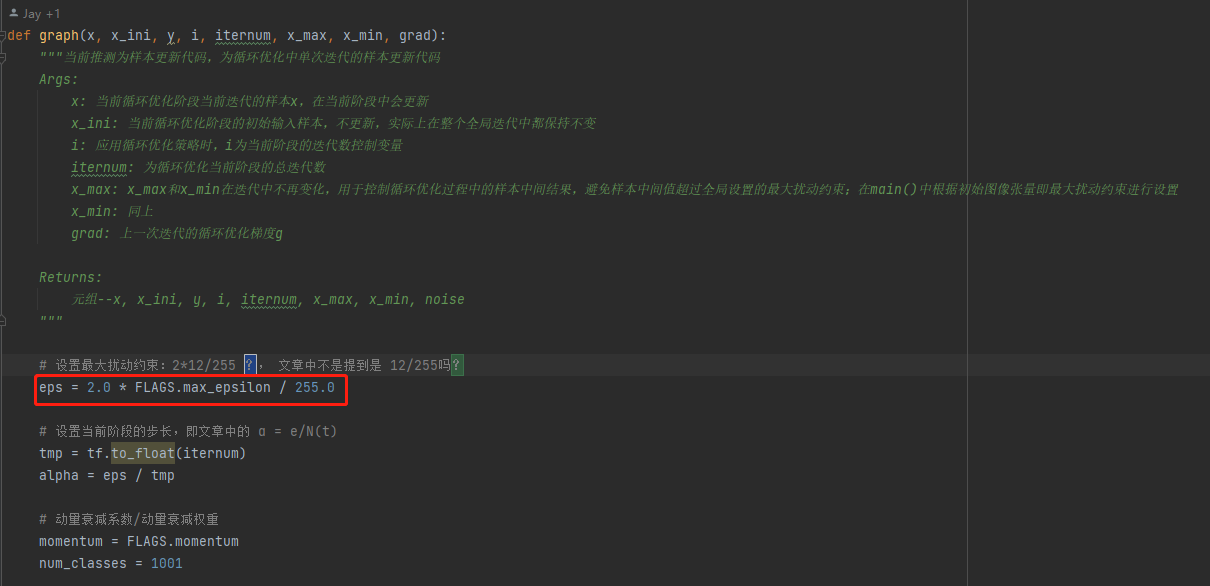
## 【OK】问题3

最大扰动幅度约束，文章中设置的为12/255，但是Github项目中graph()和main()的实现都把eps设置为24/255。

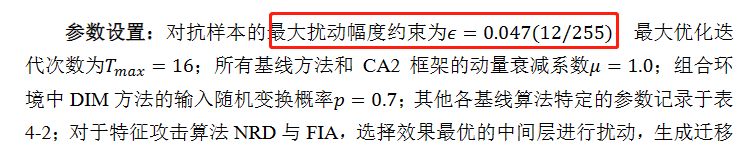
最终推荐设置是？

|  |
| --- |
| **黄博：**  **同问题1** |





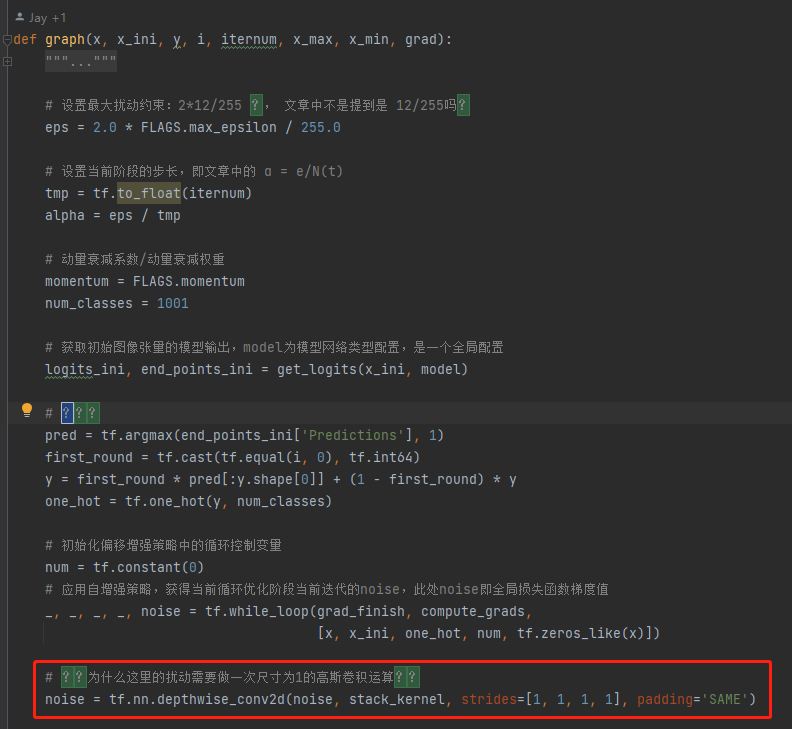




## 【OK】问题4

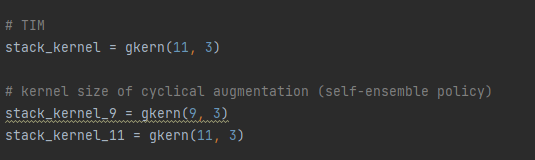
graph()的实现中，采用自增强策略计算得到损失函数梯度后，为什么需要进行一次卷积核尺寸为1的高斯卷积运算？ *//卷积核尺寸为1，计算后数值应该没有变化吧？*

|  |
| --- |
| **黄博：**  **如果加上TIM方法，这里的stack\_kernel尺寸不是1\*1，如果不加，就是1\*1** |

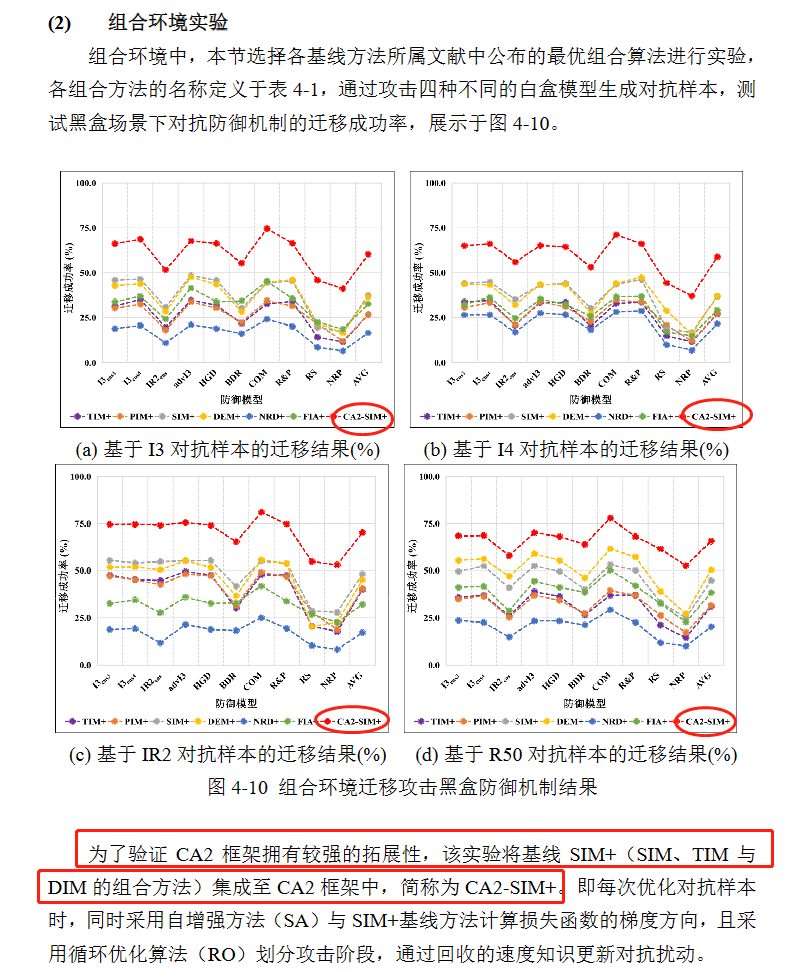


补充：

在CA2 github项目中的CA2-SIM.py。



CA2-SIM+版本，实则为集成了SIM、DIM、TIM三个当下前沿算法的究极小宇宙之YYDS版本。

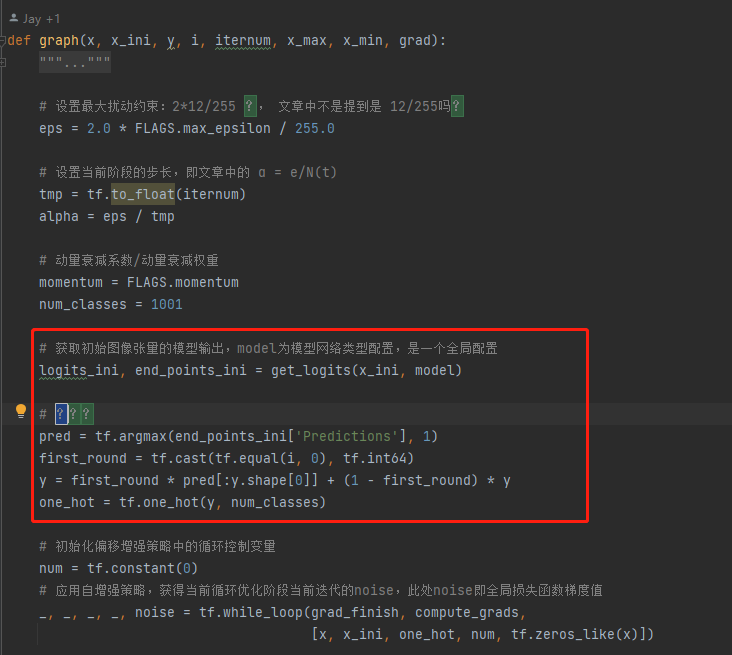


## 【OK】问题5

还是关于graph()的实现，红框内的这段代码目的是为了计算one\_hot，没有搞明白。后面作为入参传递给compute\_grads()计算梯度。

one\_hot的含义和作用是什么？

|  |
| --- |
| **黄博：**  **One\_hot是把标签变成向量，比如class有10个类**  **标签为2，则对应one-hot=[0,0,1,0,0,0,0,0,0,0]** |



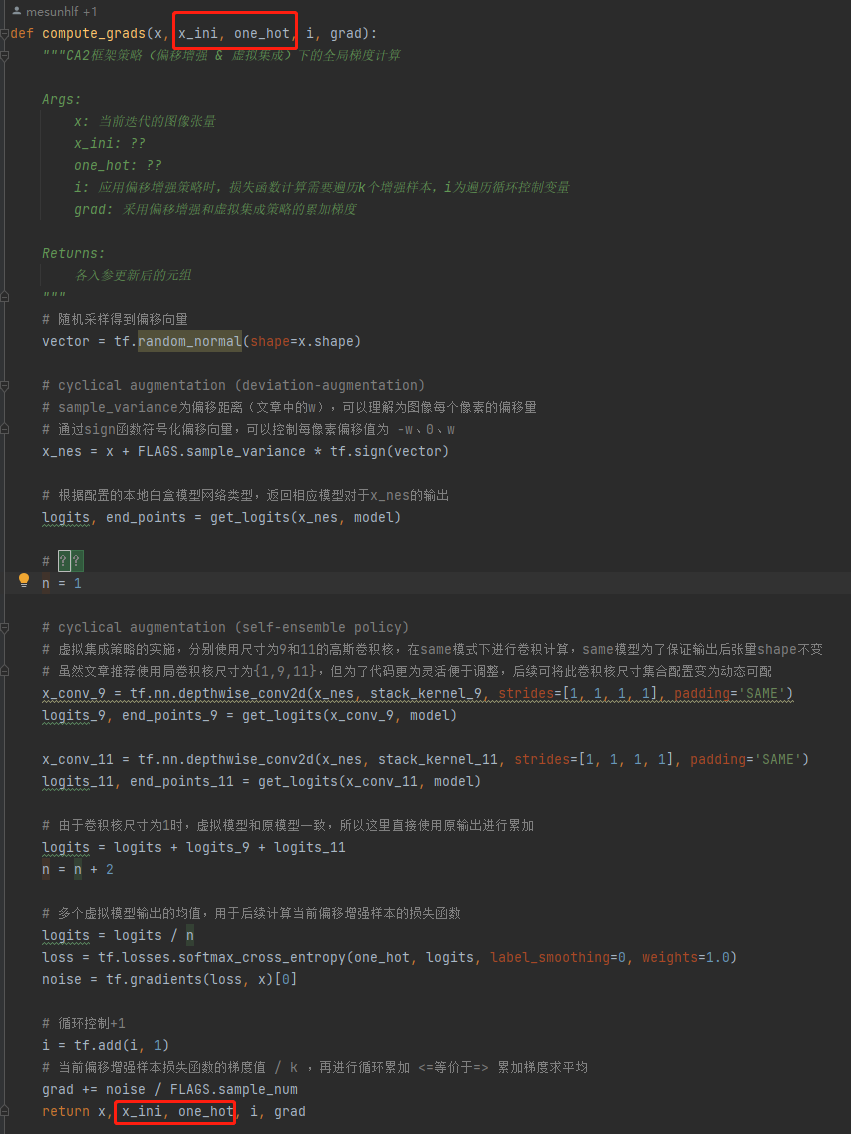
## 【OK】问题6

compute\_grads()的实现，入参x\_ini 和 one\_hot没搞明白作用。

x\_ini好像在return前没有变化，在自增强策略梯度计算的迭代中，起到的作用是？

one\_hot的疑问同上一个问题。

|  |
| --- |
| **黄博：**  **x\_ini保持不变，x会迭代变换，在第一次迭代时，x\_ini=x，所以用x\_ini记录原始数据的标签** |



## 【OK】问题7

还是关于compute\_grads()的实现的问题：

文章中关于虚拟模型集成后的损失函数，定义为各个虚拟模型损失函数的累加和；

但是代码中是将各个虚拟模型的输出(logits)加和求平均后再计算交叉熵损失函数；

不太确认这里实现是否和文章定义是一致的？

|  |
| --- |
| **黄博：**  **这里是论文写的不严谨，应该是这种写法更准确，迁移性结果上应该有一点差别** |

补充：

