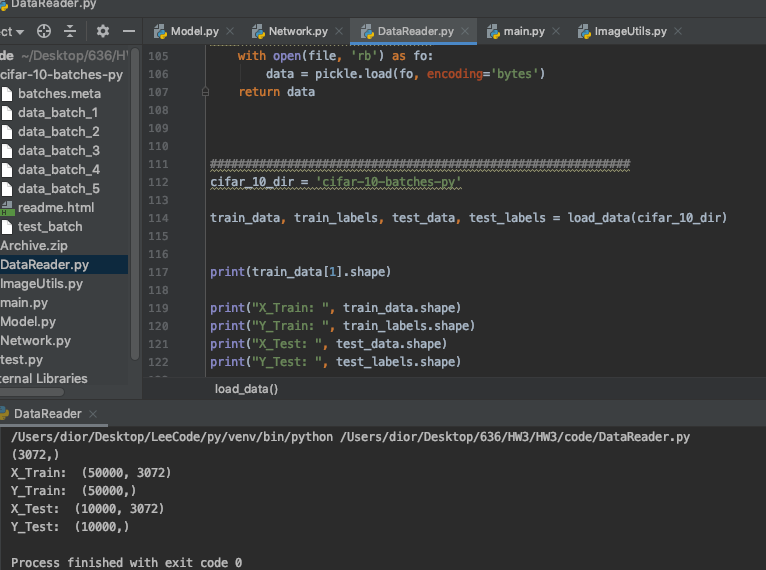
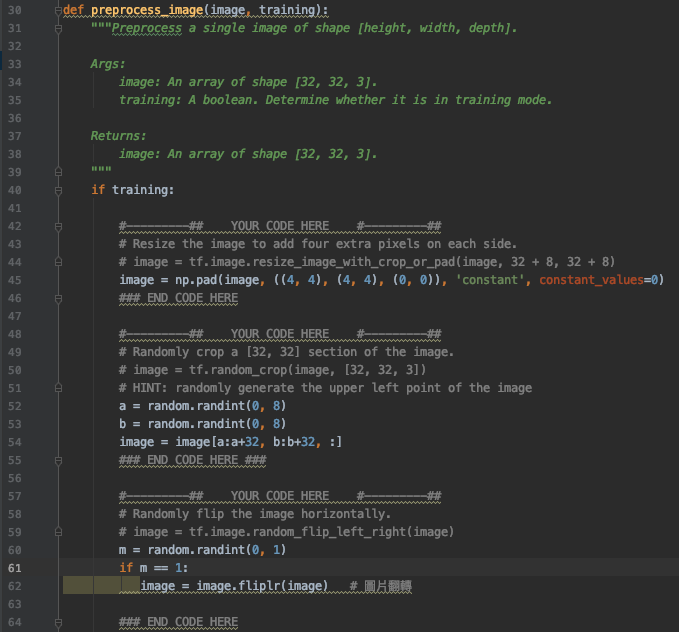
UIN: 128007734 | Zebo Xiong

1 (a)

We use the unpickle function to extract the data

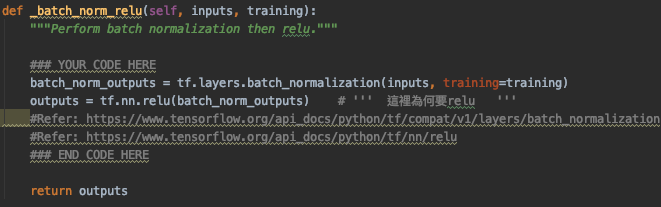


1(b)

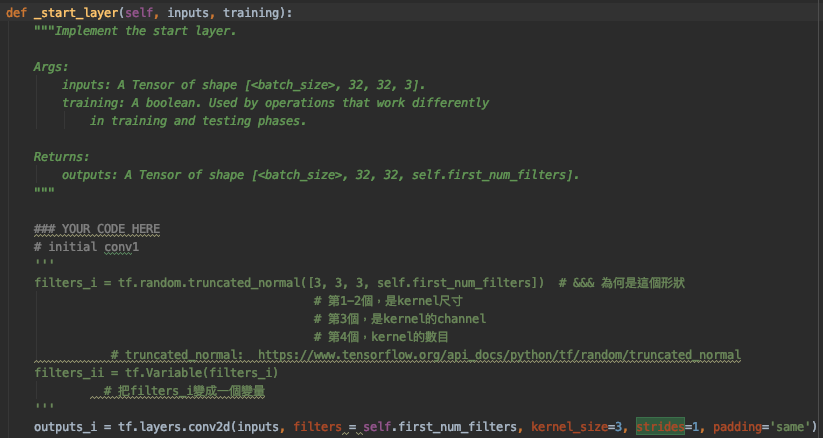


1(c)

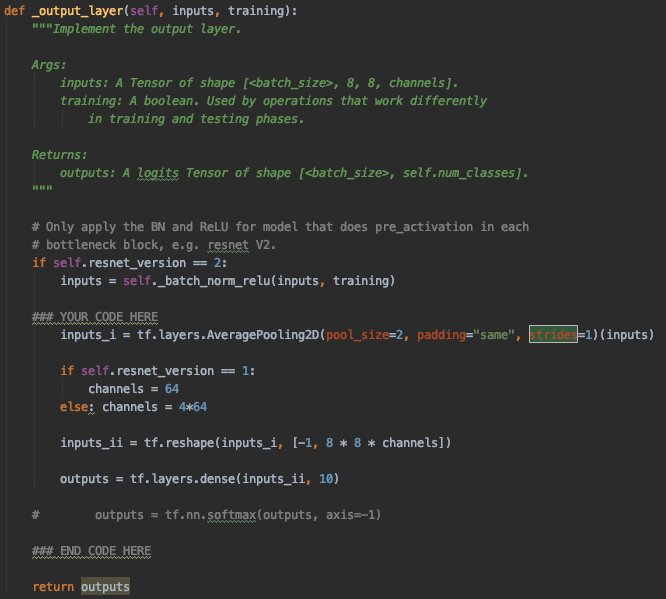
* Batch norm relu



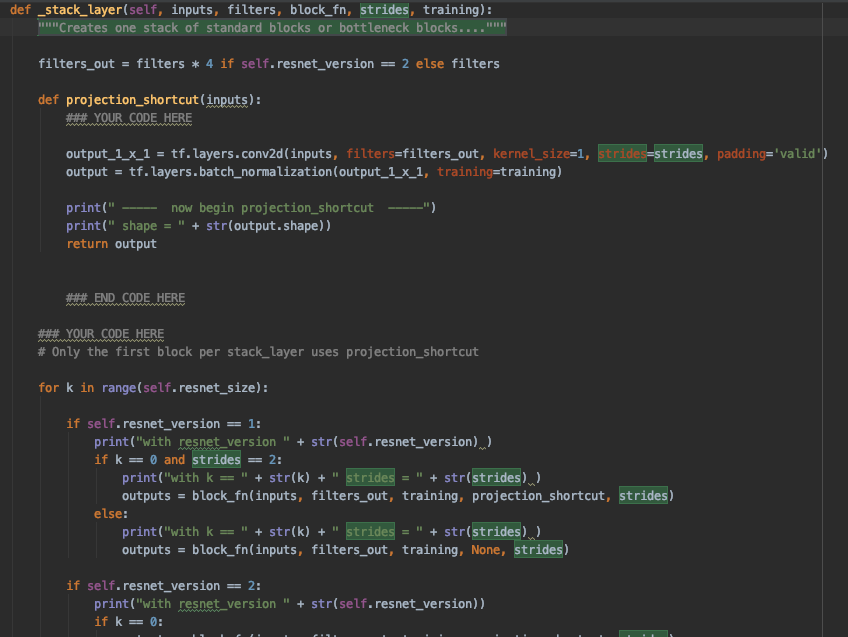
* Start layer 🡪 which use to start the network



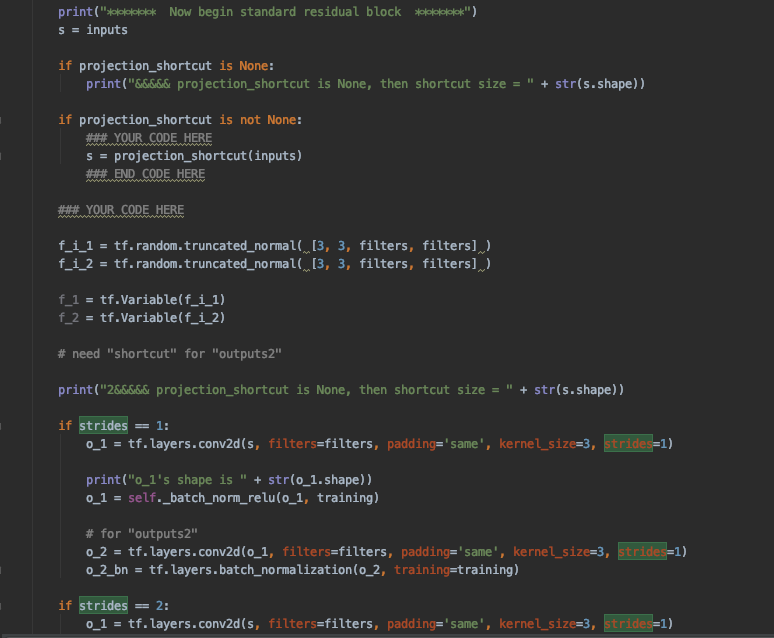
* Output layer 🡪 which finish the network



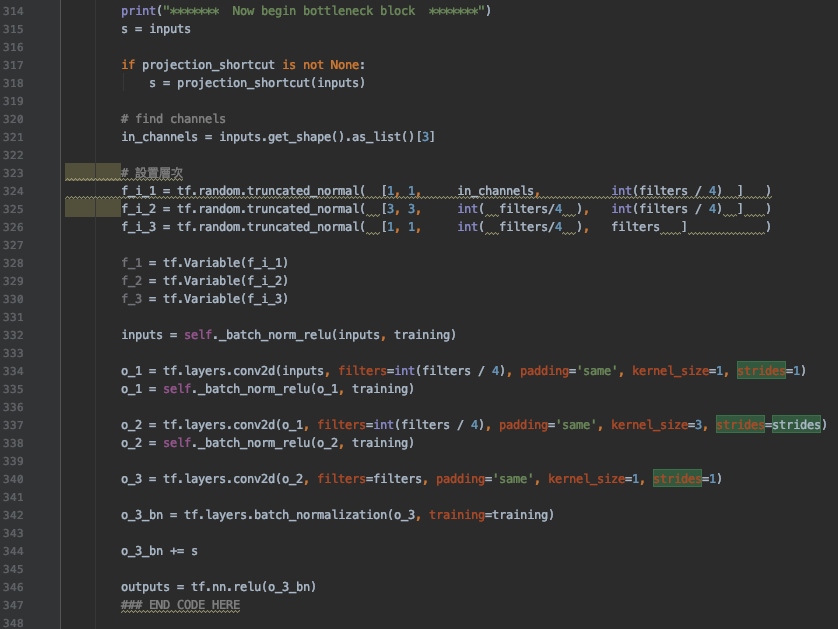
* Stack layer 🡪 which content the several components



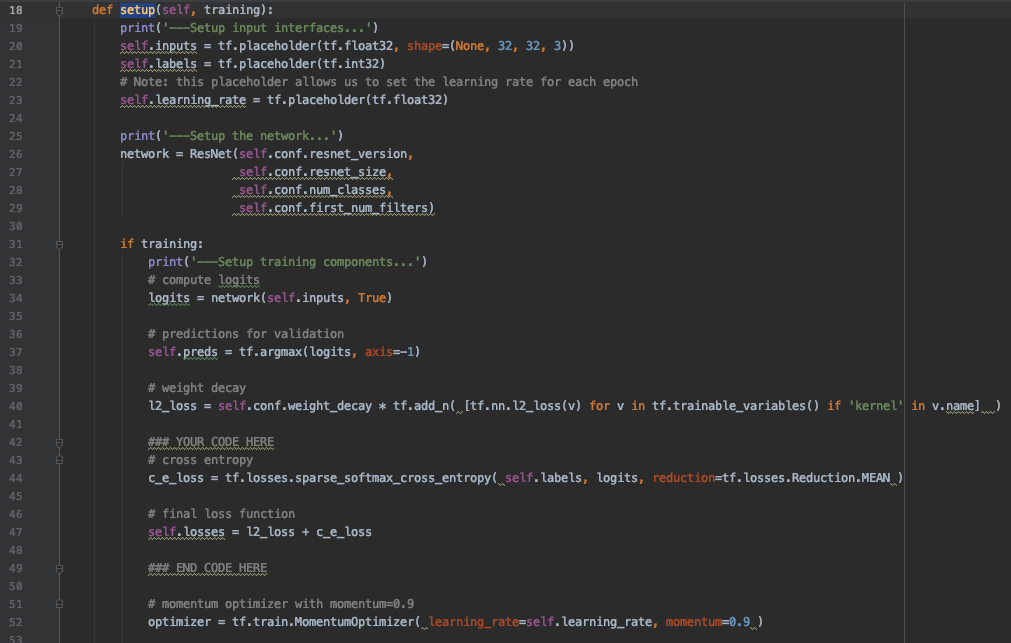
* standard residual block 🡪 this is what we need to construct



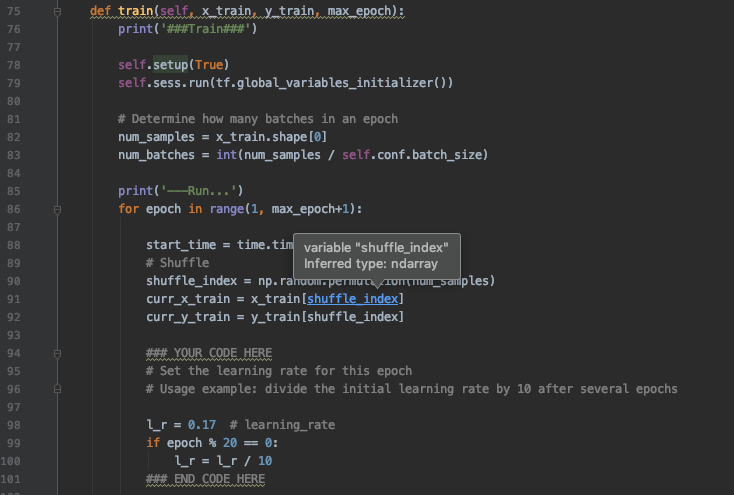
* bottleneck block 🡪 this is also what we need to construct



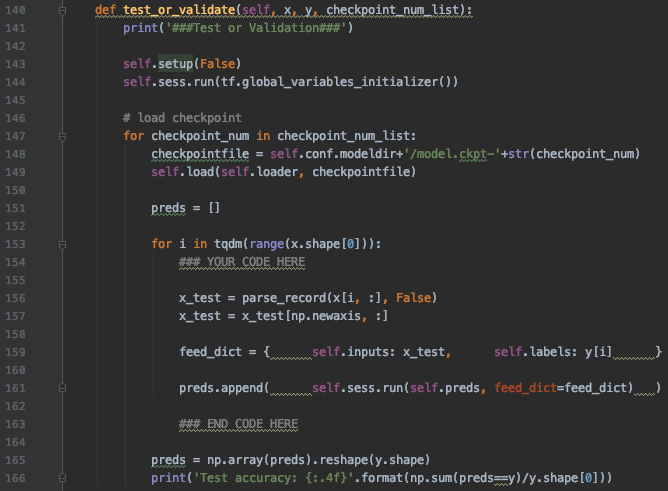
1(d)

Setup 

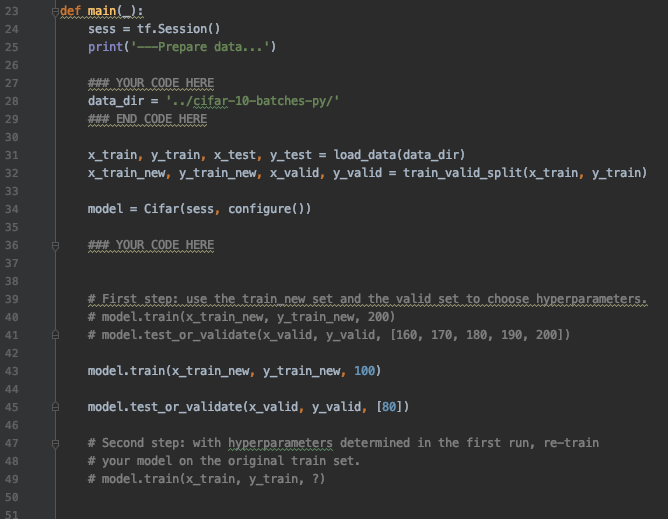
train

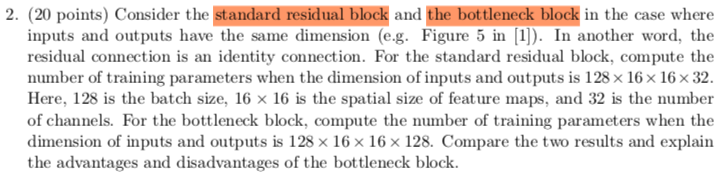


* test\_or\_validate



1(e)

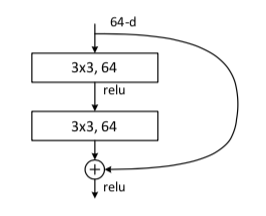


2. 

1. standard residual block

32\*3\*3\*32 + 32\*3\*3\*32 = **18432**

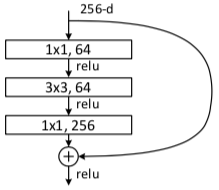
Reference graph ONLY



We only consider 32 🡪 the channel number

1. the bottleneck block

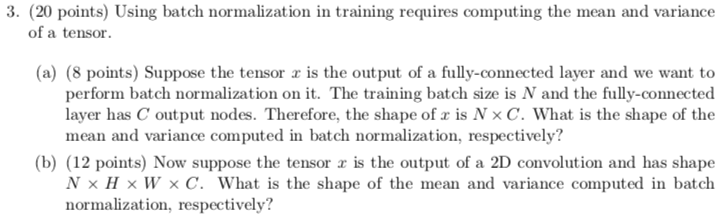
Reference graph ONLY



128\*1\*1\*32 + 32\*3\*3\*32 + 32\*1\*1\*128 = **17408**

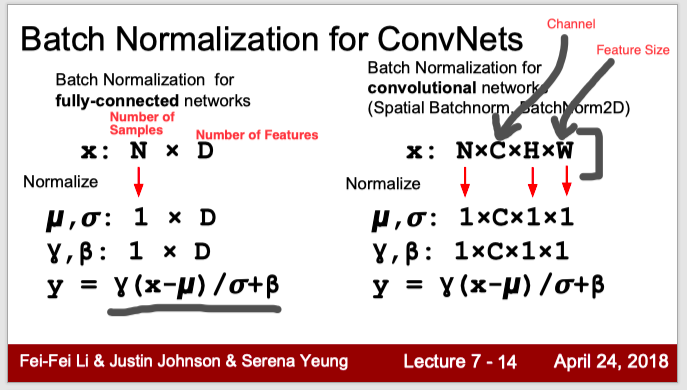
🡪 the advantage for bottleneck block is: less computation complexity

3.



1. 1\*C for both mean and variance
2. 1\*1\*1\*C for both mean and variance

Reference:



4.

