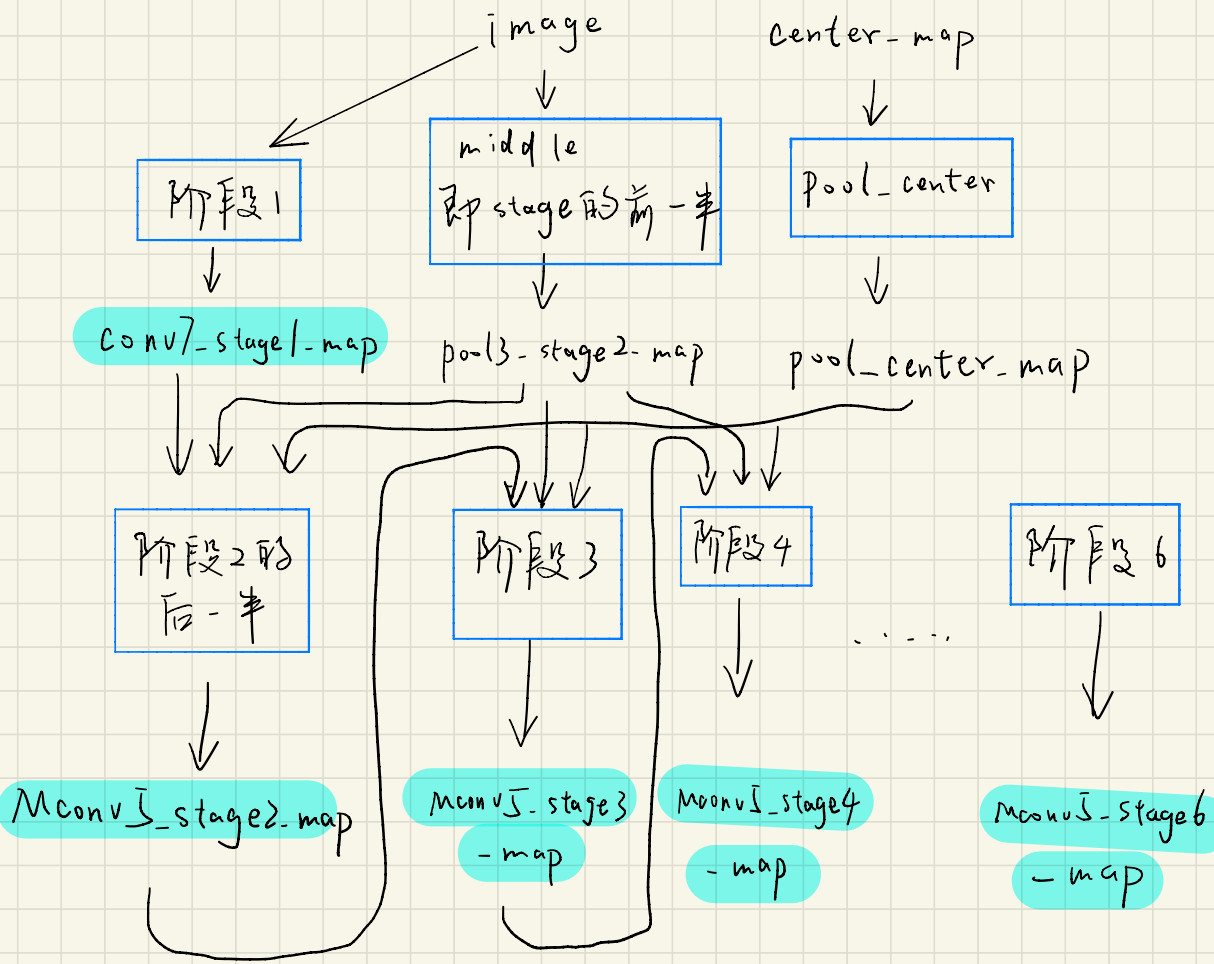
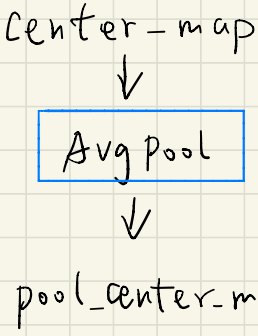
整体结构



1. 被高亮的就是每个阶段的 heatmap
   被传出去用于 intermedia supervise
2. 每个阶段都会用到中心和上一阶段的输出
   但只有第一阶段会用原图，之后都共用第二阶段前一半
   抽取的 46·46·128的 原图特征。

作业：

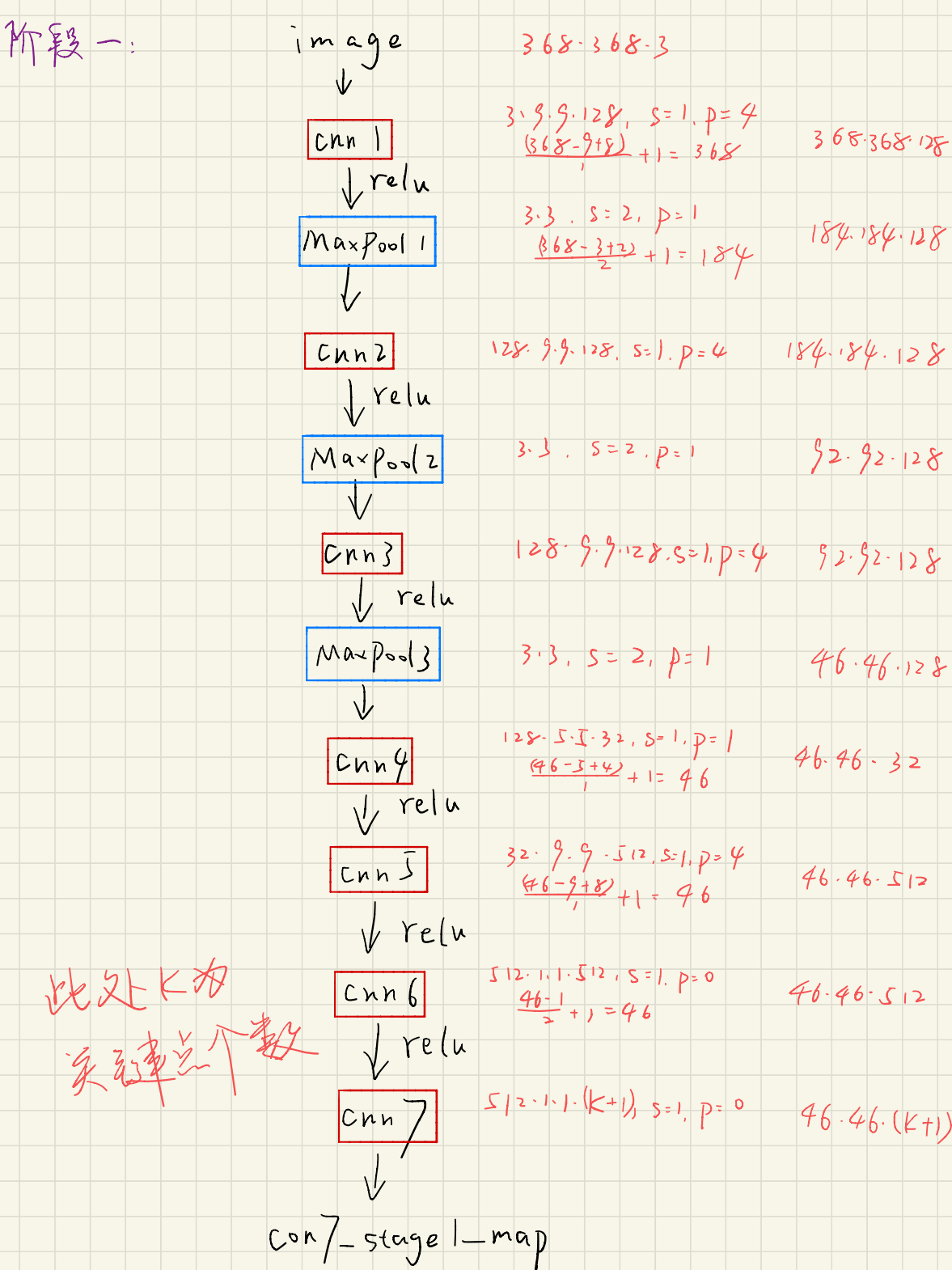


368×368

9×9, s=8, p=1

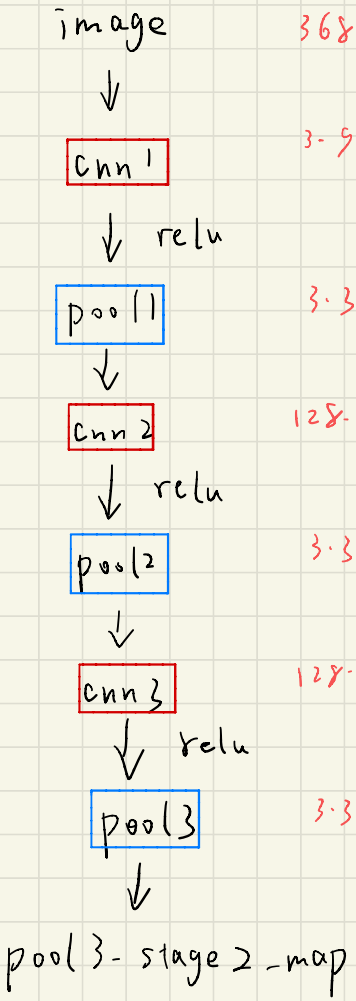
$\frac{(368+2-9)}{8}+1=46$

46·46


阶段一：
image
368·368·3
cnn 1
3·9·9·128, s=1, p=4
(368-9+8)/1 +1 = 368
368·368·128
relu
MaxPool 1
3·3, s=2, p=1
(368-3+2)/2 +1 = 184
184·184·128
cnn2
128·9·9·128, s=1, p=4
184·184·128
relu
MaxPool2
3·3, s=2, p=1
92·92·128
cnn3
128·9·9·128, s=1, p=4
92·92·128
relu
MaxPool3
3·3, s=2, p=1
46·46·128
cnn4
128·5·5·32, s=1, p=1
(46-5+4)/1 +1 = 46
46·46·32
relu
cnn5
32·9·9·512, s=1, p=4
(46-9+8)/1 +1 = 46
46·46·512
relu
cnn6
512·1·1·512, s=1, p=0
46-1/1 +1 = 46
46·46·512
relu
cnn7
512·1·1·(k+1), s=1, p=0
46·46·(k+1)
此处k为关键点个数
con7_stage1_map

中间层
middle

即=阶段前一半



第二阶段后-半和之后所有阶段都类似：