1. **梯度消失问题是否可以通过增加学习率来缓解？**

* 梯度消失问题主要是在深层神经网络中梯度在反向传播过程中逐渐变小，导致网络的参数更新缓慢甚至停滞。增加学习率在一定程度上可以缓解梯度消失问题，增加学习率实际上就是放大了梯度，让参数更新更明显，但是这并不是一个“有效”或者说“根本”的方法
* 增加学习率可能会让会让训练不稳定，梯度也许会爆炸使得参数更新过大根本收敛不了，也有可能形成局部最优解问题，跳过了全局最优解。
* 如果要解决梯度消失问题，我们可以使用残差网络，跳跃链接，或者使用ReLU,Leaky ReLU等激活函数等方法来缓解这个问题。

1. **学习参数数量**

* 通过题目可以了解到输入是一张100×100的RGB图像，因此输入维度100×100×3=30000。全连接层有100个神经元，每个神经元都与输入的30,000个元素相连，因此权重参数的数量为30000×100=3000000
* 此外，每个神经元还有一个偏置参数，因此偏置参数的数量为100。总的可学习参数数量为：3000100

1. **卷积层学习参数数量**

* 对于输入为256×256的RGB图像，输入通道数为3。使用100个大小为3×3的卷积核，每个卷积核的参数数量为3×3×3=27（每个卷积核处理3个输入通道）。100个卷积核的总参数数量为27×100=2700。 每个卷积核还有一个偏置参数，因此偏置参数的数量为100。总学习参数数量为2800
* 如果输入是灰度图，输入通道数为1。每个卷积核的参数数量为3×3×1=9。100个卷积核的总参数数量为9×100=900。偏置参数的数量仍为100。总的可学习参数数量为1000

1. **等宽卷积**

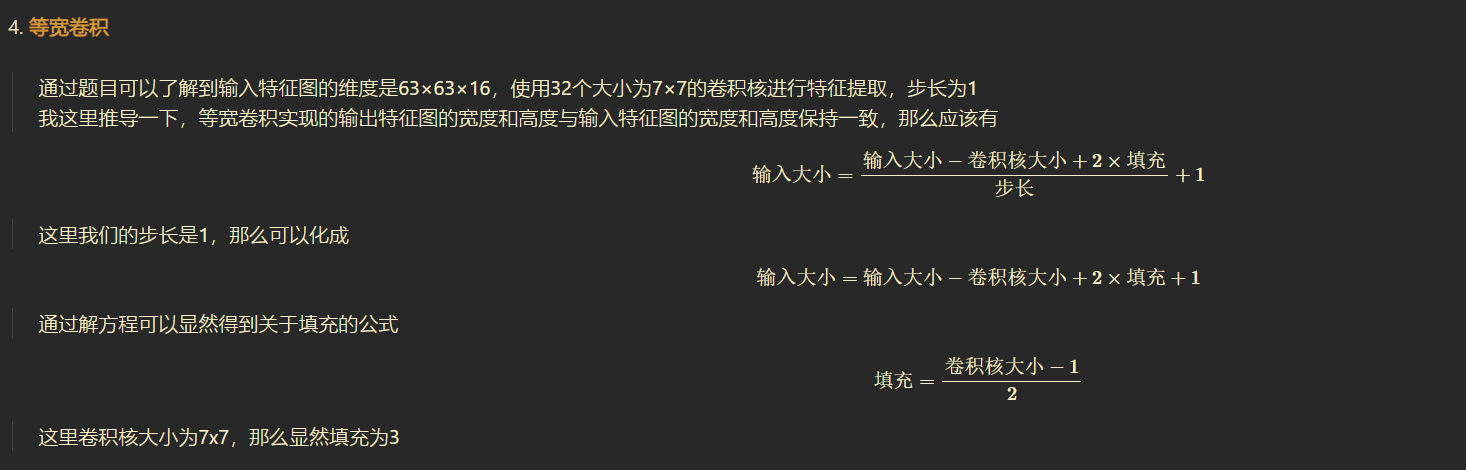
* 通过题目可以了解到输入特征图的维度是63×63×16，使用32个大小为7×7的卷积核进行特征提取，步长为1 我这里推导一下，等宽卷积实现的输出特征图的宽度和高度与输入特征图的宽度和高度保持一致，那么应该有 $$

= + 1

= - + 2 + 1

= $$ 这里卷积核大小为7x7，那么显然填充为3

（这里pandoc转换支持不是很好，我贴一下具体图片）



1. **前馈神经网络** X1, X2为网络输入，Y为网络输出，w1, w2, w3, w4, w5, 以及w6为神经元权重，所有神经元无偏置参数，无激活函数。

* **第一问** 对于第一问，显然的结果是，对于没有激活函数和偏置的网络，那么就是一种线性组合函数，线性计算就可以得到。
* 对于原网络，两个隐藏层输出和输出层输出应该是
* 展开后X1,X2合并同类项即可得到新的，整理后可以得到Y关于这两个量的系数，作为新的权重即可，新的权重如下
* **第二问** 对于第二问，先分析一下输入输出关系

| X1 | X2 | Y |
| --- | --- | --- |
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

这里我们隐藏层和输出层输出应该是

这里直接给出一组可能的解： W1 = 10 , W2 = -10 , W3=10 , W4=−10 , W5=10 , W6=−10 模拟一下对于非全0输入的模拟 X1=0,X2=1

X1=1,X2=0

X1=1,X2=1

X1=X2=0，对于全0输入，因为使用sigmoid激活函数，隐藏层输出都是0，那么计算权重后结果Y也是0

综合一下结果确实能够实现XOR计算