

2024 春季学期-智能计算系统-作业二

(要求：独立完成，4 月 3 日前在 BB 平台提交)

1. (10 points) (课本 P100 页习题 3.1) 计算 AlexNet 网络中的神经元数目及可训练的参数数目。

AlexNet 网络结构图如图 1所示。AlexNet 原文采用两块 GPU 进行网络训练，除了卷积层 1 和 3 以外，其他卷积层均只使用各自通道内的神经元进行计算。

[卷积层 1]

输入神经元数量，即输入图像大小为 $224 \times 224 \times 3 = 150528$ 。(注：部分框架的预训练模型使用 $227 \times 227 \times 3$ 大小的图像作为输入。)卷积计算的输出神经元数量为 $55 \times 55 \times 96 = 290400$ 。最大池化输出神经元的数量为 $27 \times 27 \times 48 \times 2 = 69984$ 。

只有卷积计算包含可训练参数，其中权重数量为 $11 \times 11 \times 3 = 34848$ ，偏置数量为 96。

[卷积层 2]

输入神经元数量即为上一层的输出神经元数量。卷积计算的输出神经元数量为 $27 \times 27 \times 128 \times 2 = 186624$ 。最大池化输出神经元的数量为 $13 \times 13 \times 128 \times 2 = 43264$ 。

只有卷积计算包含可训练参数，其中权重数量为 $5 \times 5 \times 48 \times 2 = 307200$ ，偏置数量为 $128 \times 2 = 256$ 。

[卷积层 3]

输入神经元数量即为上一层的输出神经元数量。卷积计算的输出神经元数量为 $13 \times 13 \times 192 \times 2 = 64896$ 。

权重数量为 $3 \times 3 \times 256 = 884736$ ，偏置数量为 384。

[卷积层 4]

输入神经元数量即为上一层的输出神经元数量。卷积计算的输出神经元数量为 $13 \times 13 \times 192 \times 2 = 64896$ 。

权重数量为 $3 \times 3 \times 192 \times 2 = 663552$ ，偏置数量为 $192 \times 2 = 384$ 。

[卷积层 5]

输入神经元数量即为上一层的输出神经元数量。卷积计算的输出神经元数量为 $13 \times 13 \times 128 \times 2 = 43264$ 。最大池化输出神经元的数量为 $6 \times 6 \times 128 \times 2 = 9216$ 。

只有卷积计算包含可训练参数，其中权重数量为 $3 \times 3 \times 192 \times 2 = 442368$ ，偏置数量为 $128 \times 2 = 256$ 。

[全连接层 6]

输入神经元数量即为上一层的输出神经元数量。全连接计算的输出神经元数量为 4096。

权重数量为 $9216 \times 4096 = 37748736$ ，偏置数量为 4096。

[全连接层 7]

输入神经元数量即为上一层的输出神经元数量。全连接计算的输出神经元数量为 4096。

权重数量为 $4096 \times 4096 = 16777216$ ，偏置数量为 4096。

[全连接层 8]

输入神经元数量即为上一层的输出神经元数量。全连接计算的输出神经元数量为 1000。

权重数量为 $4096 \times 1000 = 4096000$ ，偏置数量为 1000。

[总计]

神经元数量为 $932264 \approx 932.26K$ 。(注：池化层神经元也可不单独计算。)

可训练参数为 $60965224 \approx 60.97M$ 。

2. (10 points) (课本 P100 页习题 3.2) 计算 AlexNet 网络中完成一次正向过程所需要的乘法数量和加法数量。

这里只计算卷积层和全连接层的运算量，忽略激活函数、池化层等的运算量。

[卷积层 1]

卷积计算乘法量 $11 \times 11 \times 3 \times 55 \times 55 \times 96 = 105415200$

卷积计算加法量与乘法量相同 105415200。

[卷积层 2]

卷积计算乘法量 $5 \times 5 \times 48 \times 27 \times 27 \times 128 \times 2 = 223948800$

卷积计算加法量与乘法量相同 223948800。

[卷积层 3]

卷积计算乘法量 $3 \times 3 \times 256 \times 13 \times 13 \times 384 = 149520384$

卷积计算加法量与乘法量相同 149520384。

[卷积层 4]

卷积计算乘法量 $3 \times 3 \times 192 \times 13 \times 13 \times 192 \times 2 = 112140288$

卷积计算加法量与乘法量相同 112140288。

[卷积层 5]

卷积计算乘法量 $3 \times 3 \times 192 \times 13 \times 13 \times 128 \times 2 = 74760192$

卷积计算加法量与乘法量相同 74760192。

[全连接层 6]

全连接计算乘法量 $6 \times 6 \times 128 \times 2 \times 4096 = 37748736$

卷积计算加法量与乘法量相同 37748736。

[全连接层 7]

全连接计算乘法量 $4096 \times 4096 = 16777216$

卷积计算加法量与乘法量相同 16777216。

[全连接层 8]

全连接计算乘法量 $4096 \times 1000 = 4096000$

卷积计算加法量与乘法量相同 4096000。

[总计]

乘法量为 $724406816 \approx 724.41M$ 。

加法量为 $724406816 \approx 724.41M$ 。

3. (10 points) (课本 P100 页习题 3.3) 简述错误率与 IoU、mAP 的关系。

IoU 用于判断边界框定位的准确性。其定义为预测的边界框和真实边界框之间，交叉部分的面积和总面积的比值。IoU 越大，定位越准确，相应的检测错误率越低。

mAP 则是针对不同的置信度 (IoU) 和召回率，对精度求出的平均值。mAP 越高，模型在不同置信度 (IoU) 阈值下的检测结果更准确，错误率更低。

4. (10 points) (课本 P100 页习题 3.6) 简述 R-CNN、Fast R-CNN 和 Faster R-CNN 的区别。

R-CNN 系列模型都属于二阶段目标检测算法。

R-CNN 使用 selective search 的方法生成候选区域，使用深度神经网络进行特征提取，使用 SVM 对每个候选区域做分类，使用线性回归修正边界框的位置和大小。在 R-CNN 系列三种算法中的 mAP 最低，计算速度最慢。

Fast R-CNN 同样使用 selective search 的方法生成候选区域，但是后续的特征提取、物体分类和候选框位置和大小修正都由深度学习完成。整体性能和计算速度相比 R-CNN 有显著的提升。

Faster R-CNN 则使用深度学习的方法完成所有目标检测步骤，将候选区域生成、特征提取、分类、位置调整统一到一个深度网络框架之内，大大提高了运行速度，mAP 性能也获得了提升。

5. (10 points) (课本 P100 页习题 3.7) 简述 LSTM 标准模型中三个门的主要作用，并给出计算公式。

遗忘门：记住前一时刻单元状态的多少内容。

$$\mathbf{f}^{(t)} = \sigma(U^f \mathbf{x}^{(t)} + W^f \mathbf{h}^{(t-1)} + \mathbf{b}^f) \quad (1)$$

其中 \mathbf{x} 为输入， \mathbf{h} 为隐藏状态， U 和 W 为权重参数， \mathbf{b} 为偏置， $\sigma()$ 为激活函数。

输入门：写入多少输入到当前单元。

$$\mathbf{i}^{(t)} = \sigma(U^i \mathbf{x}^{(t)} + W^i \mathbf{h}^{(t-1)} + \mathbf{b}^i) \quad (2)$$

输出门：输出多少当前单元状态。

$$\mathbf{g}^{(t)} = \sigma(U^o \mathbf{x}^{(t)} + W^o \mathbf{h}^{(t-1)} + \mathbf{b}^o) \quad (3)$$

6. (10 points) (课本 P100 页习题 3.8) 简述 GAN 的训练过程。

判别网络采用常规的训练方法。训练数据包括小批量的真实样本 x 和小批量的噪声 z ，真实样本的标记为 1，生成网络输入噪声后生成的假样本的标记为 0。

利用这些数据训练出判别网络之后，再将判别网络用到生成网络训练过程中。生成网络训练时，它根据输入训练数据，输出假样本，并将其作为判别网络的输入。判别网络可能会被欺骗，如果被欺骗，则输出 1，否则输出 0。接着做反向传播更新生成网络的参数，判别网络的参数不变。

然后再继续交替训练判别网络和生成网络，经过多次迭代，生成网络就可以生成出非常逼真的假样本。

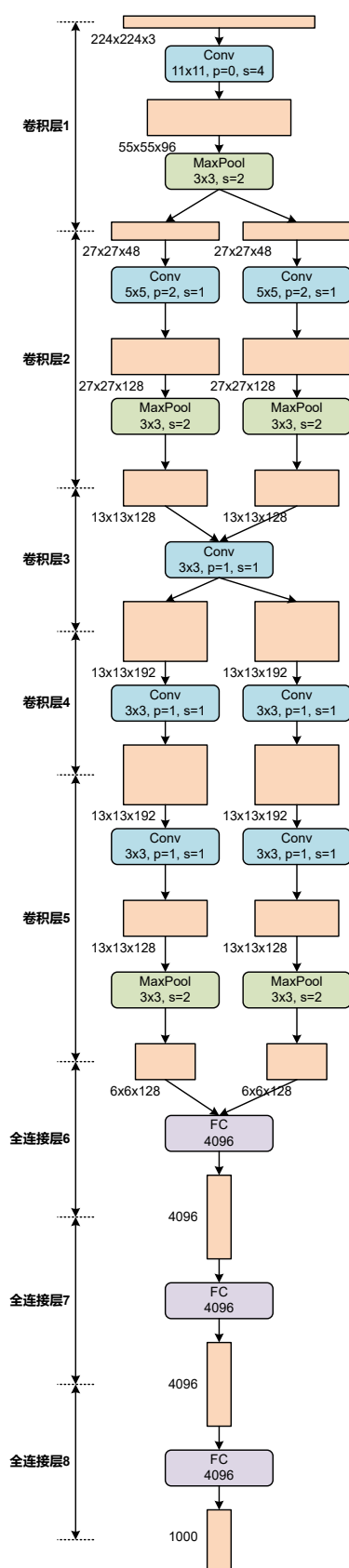


图 1: AlexNet 架构示意图
Page 5