

1、在训练神经网络时，损失函数(loss)在最初的几个 epochs 时没有下降，可能的原因是？

- A、学习率(learning rate)太低
- B、正则参数太高
- C、陷入局部最小值
- D、以上都有可能

正确答案是：D

2、下列哪一项属于特征学习算法 (representation learning algorithm) ？

- A、K 近邻算法
- B、随机森林
- C、神经网络
- D、都不属于

正确答案是：C

解析：

神经网络会将数据转化为更适合解决目标问题的形式，我们把这种过程叫做特征学习。

3、阅读以下文字：假设我们拥有一个已完成训练的、用来解决车辆检测问题的深度神经网络模型，训练所用的数据集由汽车和卡车的照片构成，而训练目标是检测出每种车辆的名称（车辆共有 10 种类型）。现在想要使用这个模型来解决另外一个问题，问题数据集中仅包含一种车（福特野马）而目标变为定位车辆在照片中的位置。

- A、除去神经网络中的最后一层，冻结所有层然后重新训练
- B、对神经网络中的最后几层进行微调，同时将最后一层（分类层）更改为回归层
- C、使用新的数据集重新训练模型
- D、所有答案均不对

正确答案是：B

4、假设你有 5 个大小为  $7 \times 7$ 、边界值为 0 的卷积核，同时卷积神经网络第一层的深度为 1。此时如果你向这一层传入一个维度为  $224 \times 224 \times 3$  的数据，那么神经网络下一层所接收到的数据维度是多少？

- A、 $218 \times 218 \times 5$
- B、 $217 \times 217 \times 8$
- C、 $217 \times 217 \times 3$
- D、 $220 \times 220 \times 5$

正确答案是：A

5、假设我们有一个使用 ReLU 激活函数(ReLU activation function)的神经网络，假如我们把 ReLU 激活替换为线性激活，那么这个神经网络能够模拟出同或函数(XNOR function)吗？

- A、可以
- B、不好说
- C、不一定
- D、不能

正确答案是：D

解析：

使用 ReLU 激活函数的神经网络是能够模拟出同或函数的。

但如果 ReLU 激活函数被线性函数所替代之后，神经网络将失去模拟非线性函数的能力。

6、当在卷积神经网络中加入池化层(pooling layer)时，变换的不变性会被保留，是吗？

A、不知道

B、看情况

C、是

D、否

正确答案是：C

解析：

池化算法比如取最大值/取平均值等，都是输入数据旋转后结果不变，所以多层叠加后也有这种不变性。

7、深度学习是当前很热门的机器学习算法，在深度学习中，涉及到大量的矩阵相乘，现在需要计算三个稠密矩阵 A,B,C 的乘积 ABC,假设三个矩阵的尺寸分别为  $m \times n$ ,  $n \times p$ ,  $p \times q$ , 且  $m < n < p < q$ , 以下计算顺序效率最高的是 ( )

A、 (AB)C

B、 AC(B)

C、 A(BC)

D、 所以效率都相同

正确答案是：A

解析：

首先，根据简单的矩阵知识，因为  $A \times B$ ，A 的列数必须和 B 的行数相等。因此，可以排除 B 选项，

然后，再看 A、C 选项。在 A 选项中， $m \times n$  的矩阵 A 和  $n \times p$  的矩阵 B 的乘积，得到  $m \times p$  的矩阵  $A \times B$ ，而  $A \times B$  的每个元素需要  $n$  次乘法和  $n-1$  次加法，忽略加法，共需要  $m \times n \times p$  次乘法运算。同样情况分析  $A \times B$  之后再乘以 C 时的情况，共需要  $m \times p \times q$  次乘法运算。因此，A 选项 (AB)C 需要的乘法次数是  $m \times n \times p + m \times p \times q$ 。同理分析，C 选项 A (BC) 需要的乘法次数是  $n \times p \times q + m \times n \times q$ 。

由于  $m \times n \times p$

8、输入图片大小为  $200 \times 200$ ，依次经过一层卷积 (kernel size  $5 \times 5$ , padding 1, stride 2)，pooling (kernel size  $3 \times 3$ , padding 0, stride 1)，又一层卷积 (kernel size  $3 \times 3$ , padding 1, stride 1) 之后，输出特征图大小为

A、  $95 \times 95$

B、  $96 \times 96$

C、  $97 \times 97$

D、  $98 \times 98$

正确答案是：C

解析：

首先我们应该知道卷积或者池化后大小的计算公式，其中，padding 指的是向外扩展的边缘大小，而 stride 则是步长，即每次移动的长度。

这样一来就容易多了，首先长宽一般大，所以我们只需要计算一个维度即可，这样，经过第一次卷积后的大小为：本题  $(200-5+2*1)/2+1$  为 99.5，取 99

经过第一次池化后的大小为： $(99-3)/1+1$  为 97

经过第二次卷积后的大小为： $(97-3+2*1)/1+1$  为 97

9、考虑某个具体问题，你可能只有少量数据来解决这个问题。不过幸运的是你有一个类似问题已经预先训练好的神经网络。可以用下面哪种方法来利用这个预先训练好的网络？

- A、把除了最后一层外所有的层都冻结，重新训练最后一层
- B、对新数据重新训练整个模型
- C、只对最后几层进行调参(fine tune)
- D、对每一层模型进行评估，选择其中的少数来用

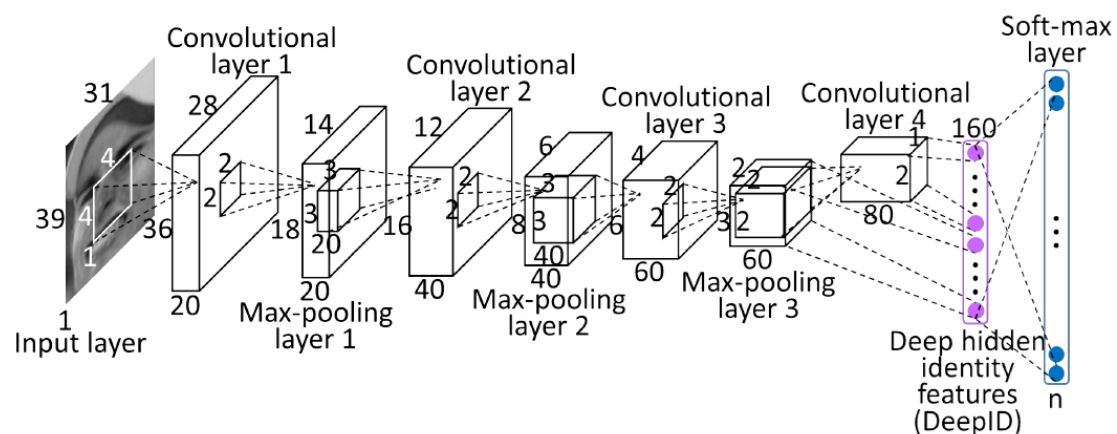
正确答案是：C

解析：

如果有预先训练好的神经网络，就相当于网络各参数有个很靠谱的先验代替随机初始化。若新的少量数据来自于先前训练数据(或者先前训练数据量很好地描述了数据分布，而新数据采样自完全相同的分布)，则冻结前面所有层而重新训练最后一层即可；但一般情况下，新数据分布跟先前训练集分布有所偏差，所以先验网络不足以完全拟合新数据时，可以冻结大部分前层网络，只对最后几层进行训练调参(这也称之为 fine tune)。

设计题：

如何通过深度学习来检测出汽车和卡车等各种车辆，目标是检测出每种车辆的名称（车辆共有 10 种类型）



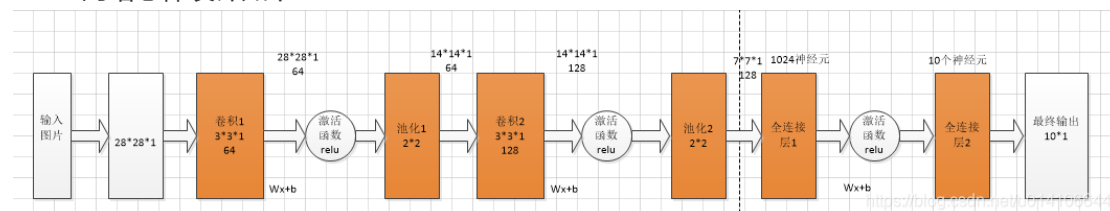
解释如下：

1.input layer: 输入层是一个 31\*39 的灰度图片，通道数是 1，也可以理解输入图片的通道是 1。

2.convolutional layer 1: 第一层卷积层，卷积核大小为 4\*4，输出结果通道为 20，表示有 20 个 4\*4 的卷积核与输入图片进行特征提取。卷积层神经元的大小为 36\*28，这个神经元大小是这么算出来的： $(39 - 4 + 1) * (31 - 4 + 1)$ 。

- 3.Max-pooling layer 1:第一层池化层，池化单元大小为  $2 \times 2$ ，池化不会改变输入数据的通道，所以输出还是 20，池化将神经元的大小降为一半， $18 \times 14$ 。
- 4.convolutional layer 2: 第二层卷积层，卷积核大小为  $3 \times 3$ ，输出结果通道为 40，表示有 40 个  $3 \times 3$  的卷积核与前一级的输出进行了特征提取。卷积层神经元的大小为  $16 \times 12$ ，这个神经元大小是这么算出来的： $(18 - 3 + 1) \times (14 - 3 + 1)$ 。
- 5.Max-pooling layer 2:第二层池化层，池化单元大小为  $2 \times 2$ ，池化不会改变输入数据的通道，所以输出还是 40，池化将神经元的大小降为一半， $8 \times 6$ 。
- 6.convolutional layer 3: 第三层卷积层，卷积核大小为  $3 \times 3$ ，输出结果通道为 60，表示有 60 个  $3 \times 3$  的卷积核与前一级的输出进行了特征提取。卷积层神经元的大小为  $6 \times 4$ ，这个神经元大小是这么算出来的： $(8 - 3 + 1) \times (6 - 3 + 1)$ 。
- 7.Max-pooling layer 3:第三层池化层，池化单元大小为  $2 \times 2$ ，池化不会改变输入数据的通道，所以输出还是 60，池化将神经元的大小降为一半， $3 \times 2$ 。
- 8.convolutional layer 4: 第四层卷积层，卷积核大小为  $2 \times 2$ ，输出结果通道为 80，表示有 80 个  $2 \times 2$  的卷积核与前一级的输出进行了特征提取。卷积层神经元的大小为  $2 \times 1$ ，这个神经元大小是这么算出来的： $(3 - 2 + 1) \times (2 - 2 + 1)$ 。
- 9.接下来，对第四层输出结果，经过一次展平操作， $80 \times (2 \times 1)$  即得到 160 个输出特征。
- 10.经过 softmax 分类，即可得到最终输出预测结果。

CNN 网络总体设计如下：



解释如下：

- 1、输入图片读取图片数据，将图片进行变换为  $28 \times 28 \times 1$ ；
- 2、经过卷积层 1，通过 64 通道的卷积核： $\{3 \times 3\}$ ，进行特征提取，最终得到  $28 \times 28 \times 64$  图片，即这些图片具有 64 个特征图；加以偏移  $b$ ，经过 relu 激活函数，实现非线性映射；
- 3、经过池化层 1，进行图片数据压缩，池化单元大小为 $\{2 \times 2\}$ ，输入图片为  $28 \times 28 \times 64$ ，输出得到  $14 \times 14 \times 64$  图片；
- 4、经过卷积层 2，通过 128 通道的卷积核： $\{3 \times 3\}$ ，进行特征提取，得到  $14 \times 14 \times 128$  图片，即这些图片具有 128 个特征图；加以偏移  $b$ ，经过 relu 激活函数，实现非线性映射；
- 5、经过池化层 2，进行图片数据压缩，池化单元大小为 $\{2 \times 2\}$ ，输入图片为  $14 \times 14 \times 128$ ，输出得到  $7 \times 7 \times 128$  图片；

6、经过全连接层 1，共有 1024 个神经元，实现将输入  $7*7*128$  转换为 1024 维度向量输出特征展平操作；加以偏移  $b$ ，经过激活函数  $\text{relu}$ ，实现非线性映射；

7、经过全连接层 2，共有 10 个神经元，将输入 1024 维度向量输出特征转换为 10 维度向量输出特征，最后经过  $\text{softmax}$  分类，即可得到最终输出预测结果。