

检测算法

1、现在要构建一个能够识别三个对象并定位位置的算法，这些对象分别是：行人 $c=1$ ，汽车 $c=2$ ，摩托车 $c=3$ 。下图中的标签哪个是正确的？ ($y = [pc, bx, by, bh, bw, c1, c2, c3]$)



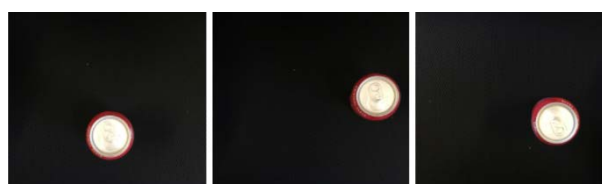
- ☒ A: $y=[1, 0.3, 0.7, 0.3, 0.3, 0, 1, 0]$
- ☐ B: $y=[1, 0.7, 0.5, 0.3, 0.3, 0, 1, 0]$
- ☐ C: $y=[1, 0.3, 0.7, 0.5, 0.5, 0, 1, 0]$
- ☐ D: $y=[1, 0.3, 0.7, 0.5, 0.5, 1, 0, 0]$
- ☐ E: $y=[0, 0.2, 0.4, 0.5, 0.5, 0, 1, 0]$

2、继续上一个问题，下图中 y 的值是多少？注：“?”是指“不关心这个值”，这意味着神经网络的损失函数不会关心神经网络对输出的结果，同上 $y = [pc, bx, by, bh, bw, c1, c2, c3]$



- ☐ A: $y=[1, ?, ?, ?, ?, 0, 0, 0]$
- ☒ B: $y=[0, ?, ?, ?, ?, ?, ?, ?]$
- ☐ C: $y=[?, ?, ?, ?, ?, ?, ?, ?]$
- ☐ D: $y=[0, ?, ?, ?, ?, 0, 0, 0]$
- ☐ E: $y=[1, ?, ?, ?, ?, ?, ?, ?]$

3、你现在任职于自动化工厂中，你的系统会看到一罐饮料从传送带上下来，你想要对其进行拍照，然后确定照片中是否有饮料罐，如果有的话就对其进行包装。饮料罐头是圆的，而包装盒是方的，每一罐饮料的大小是一样的，每个图像中最多只有一罐饮料，现在你有下面的方案可供选择，这里有一些训练集图像示例：



A: Logistic unit (用于分类图像中是否有罐头)

☒ B: Logistic unit, bx 和 by

C: Logistic unit, bx , by , bh (因为 $bw = bh$, 所以只需要一个就行了)

D: Logistic unit, bx , by , bh , bw

4、如果你想要构建一个能够输入人脸图片输出为 N 个标记的神经网络 (假设图像只包含一张脸), 那么你的神经网络有多少个输出节点?

A: N

☒ B: $2N$

C: $3N$

D: N^2

5、当你训练一个视频中描述的对象检测系统时, 里需要一个包含了检测对象的许多图片的训练集, 然而边界框不需要在训练集中提供, 因为算法可以自己学习检测对象, 这个说法对吗?

☒ A: 正确

B: 错误

6、假如你正在应用一个滑动窗口分类器 (非卷积实现), 增加步伐不仅会提高准确性, 也会降低成本。

A: 正确

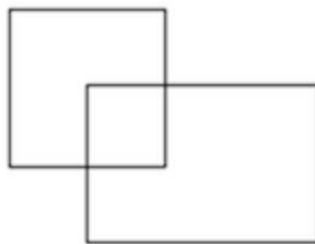
☒ B: 错误

7、在 YOLO 算法训练时候, 只有一个包含对象的中心/中点的一个单元负责检测这个对象。

☒ A: 正确

B: 错误

8、这两个框中 IoU 大小是多少? 左上角的框是 2×2 大小, 右下角的框是 2×3 大小, 重叠部分是 1×1 。



A: $1/6$

☒ B: $1/9$

C: $1/10$

D: 以上都不是

9、假如你使用 YOLO 算法, 使用 19×19 格子来检测 20 个分类, 使用 5 个锚框 (anchor box)。在训练的过程中, 对于每个图像你需要输出卷积后的结果 yy 作为神经网络目标值 (这是最

后一层), yy 可能包括一些“?”或者“不关心的值”。请问最后的输出维度是多少?

- A: $19 \times 19 \times (25 \times 20)$
- B: $19 \times 19 \times (20 \times 25)$
- ☒ C: $19 \times 19 \times (5 \times 25)$
- D: $19 \times 19 \times (5 \times 20)$

人脸识别和神经风格转换

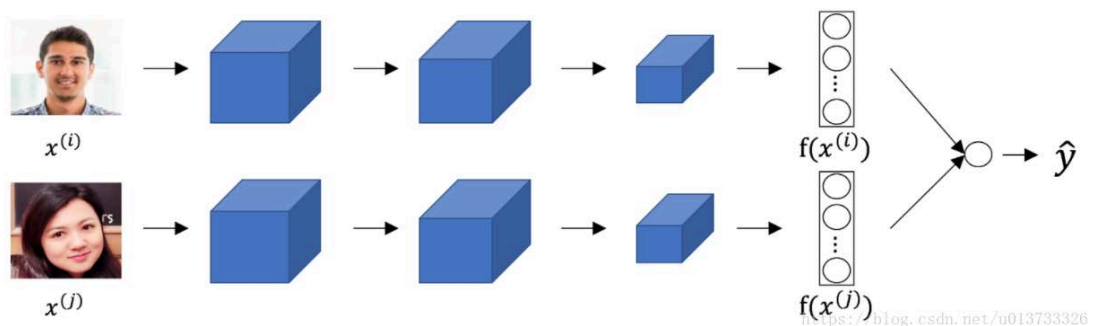
1、面部验证只需要将新图片与 11 个人的面部进行比较, 而面部识别则需要将新图片与 K 个人的面部进行比较。

- A: 正确
- ☒ B: 错误

2、为了训练人脸识别系统的参数, 使用包含了 10 万个不同的人的 10 万张图片的数据集进行训练是合理的。

- A: 正确
- ☒ B: 错误

3、在下图中的孪生卷积网络(Siamese network)结构图中, 上下两个神经网络拥有不同的输入图像, 但是其中的网络参数是完全相同的。



- ☒ A: 正确
- B: 错误

4、你在一个拥有 100 种不同的分类的数据集上训练一个卷积神经网络, 你想要知道是否能够找到一个对猫的图片很敏感的隐藏节点(即在能够强烈激活该节点的图像大多数都是猫的图片的节点), 则该节点更有可能在第 4 层找到该节点而不是在第 1 层找到。

- ☒ A: 正确
- B: 错误

5、神经风格转换被训练为有监督的学习任务, 其中的目标是输入两个图像(x), 并训练一个能够输出一个新的合成图像(y)的网络。

- A: 正确
- ☒ B: 错误

6、在一个卷积网络的深层，每个通道对应一个不同的特征检测器，风格矩阵 $G^{[l]}$ 度量了 l 层中不同的特征探测器的激活（或相关）程度。

☒ A: 正确

B: 错误

7、在神经风格转换中，在优化算法的每次迭代中更新的是什么？

A: 神经网络的参数

☒ B: 生成图像 G 的像素值

C: 正则化参数

D: 内容图像 C 的像素值

8、你现在用拥有的是 3D 的数据，现在构建一个网络层，其输入的卷积是 $32 \times 32 \times 32 \times 16$ （此卷积有 16 个通道），对其使用 32 个 $3 \times 3 \times 3$ 的过滤器（无填充，步伐为 1）进行卷积操作，请问输出的卷积是多少？

☒ A: $30 \times 30 \times 30 \times 32$

B: 不能操作，因为指定的维度不匹配，所以这个卷积步骤是不可能执行的。

C: $30 \times 30 \times 30 \times 16$