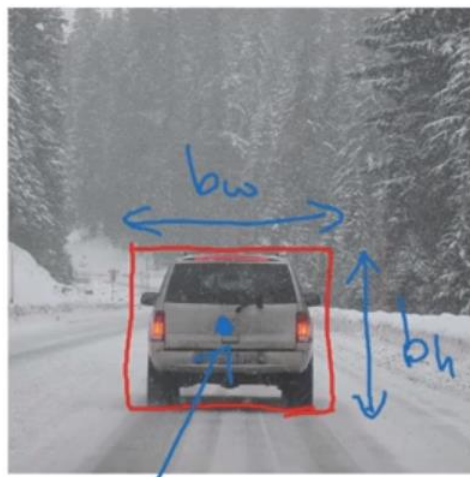


1、图像定位

之前所学神经网络只是单纯的将图片对象进行分类，但是如果要识别出图片中对象的具体位置，怎么办，该理论的实际意义可以运用在比如无人驾驶上面。接下来，将引入一些新的输出标签表达方式。



在这副图中，bx, by 表示对象中心点的坐标，bw, bh 分别表示对象的长和宽。输出标签，如下表示：

$$y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

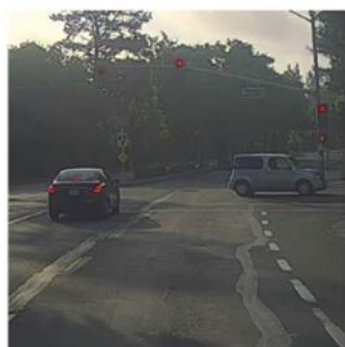
p_c 代表有无对象出现，为 1 表示有，否则其余参数都无意义， c_1 、 c_2 、 c_3 表示对象属于哪一类。



图像特征定位也可用于表情识别。在以前的算法中，如果要将人脸表情识别出来，那么将遍历图片中大量无关的特征，如果使用特征定位，那么该算法较之前将会更省时省力。

2、滑动窗口检测

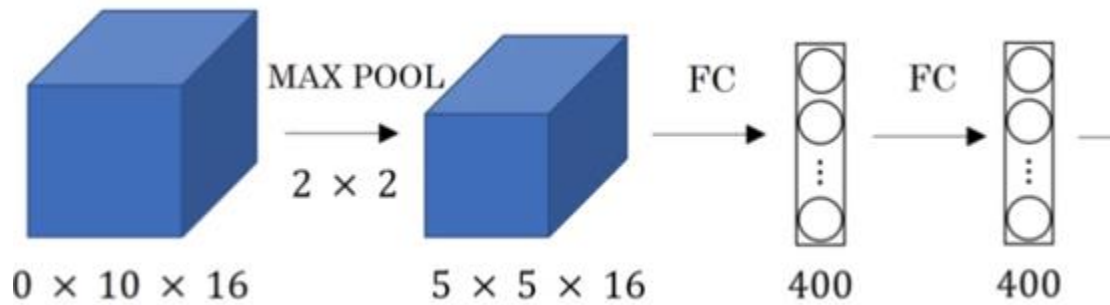
如何获取特征对象的具体位置呢，首先用较小的窗口进行整体遍历，相当于将原



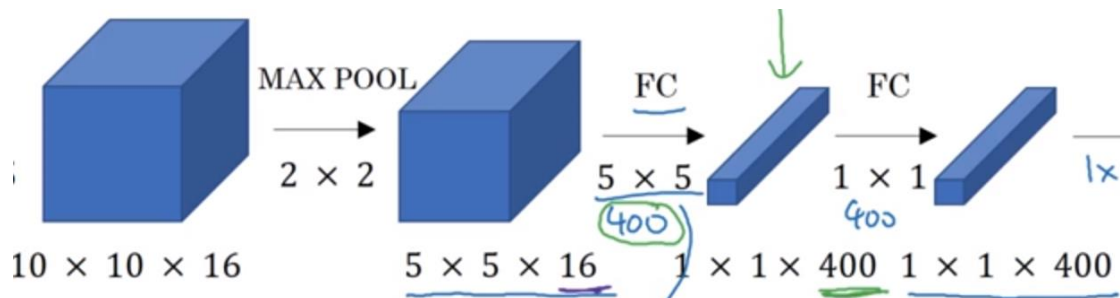
图分割成了很多较小的图片，再输入卷积神经网络进行预测，这样就能输出上一节讲到的标签。

2、卷积代替全连接

在讲如何输出图像定位之前，先来看看如何用卷积层代替全连接层。

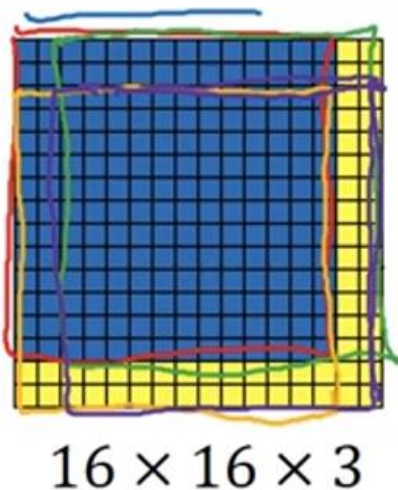


上图中，使用 FC 层将特征向量化，便于得到最后的预测分类结果。而另外一张图

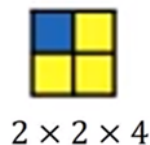


中，取而代之的是卷积层，数学意义上它们是一样的表达意义。只不过，它更加适合滑动窗口目标检测。

3、滑动窗口对象同时检测



这样一幅图中，使用滑动窗口检测器，在图上分成了4块区域，然后进行卷积池化再卷积操作，得到了



最终4个分类结果，只不过在2维平面上，不再是 1×1 的，而是 2×2 ，这源于我们将原始图片分成了4个区域，如果是其他数量，则结果在2D面上就要有对应的数量。

使用这种方法，既可以检测出对象位置，同时又提高了算法效率。