

收录于合集

#深度学习 478 #卷积神经网络 336 #目标检测 404 #计算机视觉 157 #图像识别 230

1 目标检测算法

基于卷积神经网络的目标检测及识别任务主要包含两个部分，即目标检测任务和目标识别任务，下面详细介绍这两个任务。

目标检测任务。目标检测任务常常需要预测实例在图像中的位置，通常使用四个坐标来确定实例在图像中的位置，分别是左上、左下、右上和右下四个坐标。预测出准确的实例边界框对目标检测任务来说至关重要。目前主流的目标检测任务普遍采用了基于锚框（Anchor）的边界框预测方法。实现方法为预先在每个位置上定义一组大小不同的边界框，然后对预测出来的边界框和真实的边界框之间进行微调，直到预测出来的边界框非常接近真实边界框为止。如图1所示，图右半部分表示特征图中任意一个像素点，该点对应了 9 个 Anchor，表示为图的左半部分，这 9 个 Anchor 的大小和形状均不同。这 9 个 Anchor 对应到输入图像中用来预测实例的边界框，实例较大的可能最先由蓝色框预测出来，实例较小的可能最先由绿色框预测出来，实例中等的可能最先有红色框预测出来，如图 1 所示，其中只表示了特征图上的一个像素点所对应的 Anchor，而实际中特征图的每个像素点上都会对应一组 Anchor。

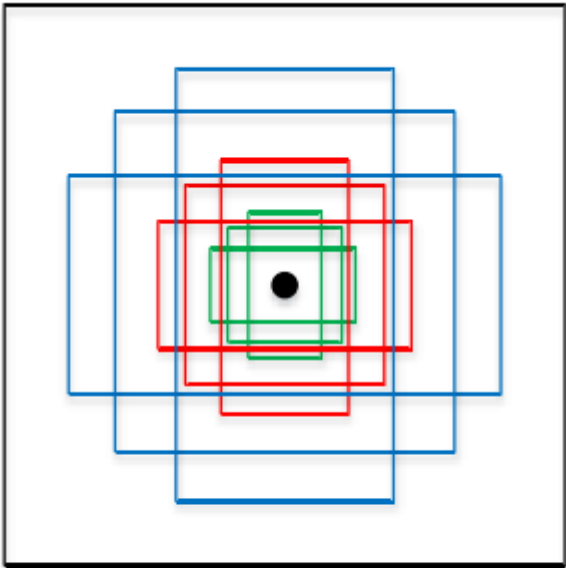


图 1 Anchor 生成图

目标识别任务。目标识别任务是预测实例的类别，目标识别任务的输出是一个C 维的向量，C 表目标类别的总数。在大多数识别任务中都会使用 Softmax 函数来预测类别，该函数的作用是将输出值限制在 0 至 1 之间，并且所有的输出值和为1。如下公式表示了其计算过程。

$$y_i = \frac{e^{x_i}}{\sum_{i=1}^C e^{x_i}}, \quad \forall i \in 1, 2, \dots, C$$

其中 $x=[x_1,x_2,\dots,x_C]-R(C)$  为输入向量， $y=[y_1,y_2,\dots,y_C]-R(C)$  为输出的概率向量，即预测的类别结果， $y_i$  表示第  $i$  个类别的概率，最大概率值代表了分类的结果。

## 2 语义分割算法

语义分割任务被定义为预测图像中每个像素点的类别。2015 年，Jonathan Long 等人提出 Fully Convolutional Networks for Semantic Segmentation（FCN），FCN 网络主要提出以下结构。

### (1) 全卷积网络

卷积神经网络（CNN）在输出预测结果前会添加一定数量的全连接层。不同大小的输入图像得到的特征图尺寸是不同的，全连接层无法将不同尺度的输入图像映射为固定大小的向量。FCN 将网络中的全连接层替换为卷积层，得到全卷积网络，解决了输入图像多尺寸问题。

### (2) 反卷积

语义分割网络要求输出和输入的尺寸相同，而卷积操作往往需要减小输入图像的尺寸，因此在进行卷积操作减小图像后，还要进行一定的操作将图像大小恢复。恢复图像大小的方法称之为上采样操作，FCN 使用反卷积对图像进行上采样，如图2 所示。



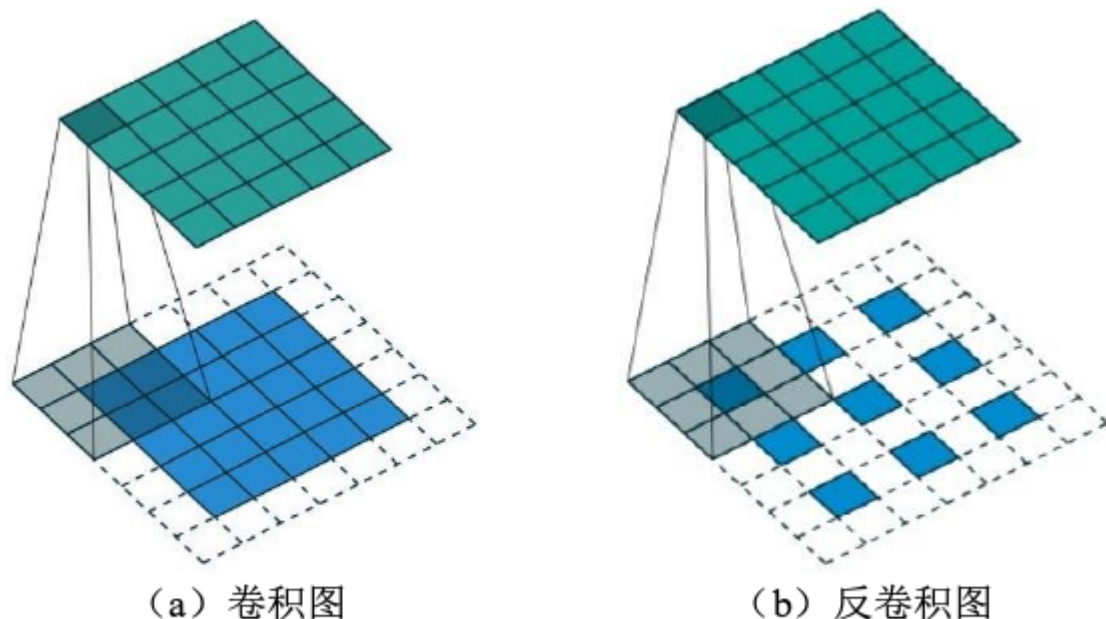


图 2 卷积和反卷积图

其中图 2 (a) 表示卷积的过程，图 2 (b) 表示反卷积的过程。蓝色的部分表示输入特征图，绿色部分表示输出的特征图，卷积核大小  $3 \times 3$ 。以特征图的左上角为例，对于卷积操作来说，先对输入特征图进行长宽扩展，表示为虚线部分，然后使用卷积核对输入特征图的左上角的位置进行卷积计算，得到的结果为绿色左上角的特征值。遍历输入特征的所有位置即可得到输出特征图。对于反卷积来说，其目的在于将特征图的尺寸放大，因此先对输入特征图进行扩散处理，在每个特征点之间进行填充操作，如图 2 (b) 虚线表示的部分。然后对填充后的特征图进行正常的卷积操作得到特征图反卷积计算结果。公式如下表示反卷积后特征图长宽的变化。其中  $H_{out}$  和  $W_{out}$  表示反卷积后特征图的高度和宽度， $H_{in}$  和  $W_{in}$  表示输入特征图的高度和宽度， $S$  表示步长， $F$  表示卷积核的大小，这里认为卷积核的长宽相等， $P$  表示填充数量。

$$H_{out} = S(H_{in} - 1) + F - 2P$$

$$W_{out} = S(W_{in} - 1) + F - 2P$$

使用反卷积的优点在于和正常的卷积操作相似，反卷积操作也是可以学习的。反卷积的正向传播可以看做卷积的反向传播过程。另外，记卷积矩阵为  $A$ ，对输入特征图进行卷积操作后，其值左乘  $A^T$  矩阵后得到的结果与输入特征图的尺寸相同。因此，反卷积操作具有图像放大的功能，而且能够根据局部信息来恢复图像细节信息。



微信公众号对应的群号:



微信公众号二维码



微信公众号：人工智能感知信息处理算法研究院

知乎主页：<https://www.zhihu.com/people/zhuimeng2080>

收录于合集 #深度学习 478

上一篇

深度学习两阶段实例分割算法-Mask R-CNN算法

下一篇

基于深度学习行人属性识别算法研究现状

People who liked this content also liked

基于 U-Net 网络改进

人工智能感知信息处理算法研究院