

基于深度卷积神经网络的物体识别算法

王华溢

新疆大学 新疆 乌鲁木齐 830008

摘要: 在计算机物体识别中,深度卷积神经网络因其独特的作用机制,能够凭借其场景数据库,将单通道的信息进行转换,形成三通道信息,之后通过训练集内彩色图片及三通道深度图片,通过微调神经网络模型,提取图片特征,经三层卷积层和一个全列阶层,串联两种模态特征,将提取的算法在场景理解物体中应用,可以对彩色、深度图片的特征进行训练分析,从而能提升物体分类的准确性。

关键词: 深度卷积神经网络; 物体识别; 模型

物体识别可以说是计算机视觉领域内的基础性研究,从上世纪六十年代起,计算机教授就开始探索这一技术,经过几十年的发展,这一技术正在逐步完善之中^[1]。卷积神经网络在物体识别任务中能够发挥重要作用,近年来有较多研究在各种物体识别分析中^[2,3],开始探索这一神经网络的应用。

一、物体识别及其过程

物体识别目标,是对图片中是否存在一个物体展开分析,首先需要对图片特征进行提取,经过分类器训练后,才可实现。可以说,近年来计算机的视觉研究如火突入,而在这一领域中,物体识别是重要的基石,对计算机其他视觉任务的发展有不可估量的推动作用。

计算机视觉领域内,物体识别典型系统为:预处理-提取特征-选择特征-建模-匹配-定位。预处理是计算机视觉算法的首要环节,是在不对图像承载本质信息改变的情况下,让每一张图片的颜色分布、尺寸及明暗程度等表面特性保持一致,为后续处理提供便利。提取特征是为了自模式样本内,提取对这一模式有代表性的特质,这是重要的步骤,同时又是难以控制的一步。图像特征的提取,是从一幅画像内提取出和其他图像不同的根本属性,从而对不同图像进行区分[4]。提取特征后,需要对特征进行选择,之后集合给定的特征,将相同点提取出来,建模,获取训练结果后匹配算法,经算法对物体进行识别,之后便可进行定位。

二、深度卷积神经网络

目前,在物体识别中,卷积神经网络逐渐获得重视,且其在诸多标准数据集上,获取了很大成果。这一网络有权值共享的特征,在图1中,如果大矩形框表示一张图片,将这张图片输入其中,而图片中的不同像素值,表示输入的对应该神经元,卷积神经网络中的神经元,和输入图片中的神经元,数量相符,有权重连接。输入区域中若是3*3,那么在卷积中,先对卷积模板进行分析,得到9个权重参数,连接卷积层及输入的神经元,并且根据步长,将上述卷积模板于图片上进行滑动,当其到达某一位置后,根据输入神经元和

权重参数,展开计算,主要计算乘积的和。假设各个权重参数是{x1,x2,...,x9},以f为激活函数,那么输出的y如下:

$$Y=f(b+x1W1+x2W2+... , x9W9)$$

其中,b:输出神经元和上层输出是1的神经元,通过权重b进行连接。卷积操作,是通过卷积模板,提取输入图片中每个3*3区域内的特征,如图1内的3个矩形,可以看做是某一个卷积模板,存在权值共享。这是因为,对图片某一位置进行某一特征的提取工作时,我们同时也可对这一图片中其他部位的相同特征进行提取,这是因为图片内的物体,可在某一图片中的各个位置出现。对一个图片,根据卷积的模板,对其展开卷积,即可得到特征图一张,特征图中,不同像素是和卷积层神经元保持对应的。若这样的卷积模板数量很多,而每一个卷积模板,均可对图片展开一次相同的卷积操作,有几个模板,获得特征图的数量就是几个。经过上述操作,能够促使模型中需要进行训练的参数量大大降低,节约训练模型的时间,同时还可将待提取特征维数增加。这一操作可以将原始图片作为输入图片进而为物体的识别提供技术支持。

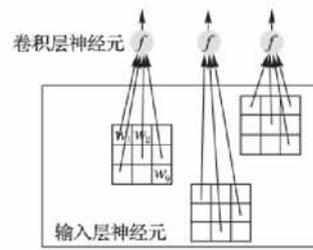


图1: 卷积操作

三、物体识别算法

在识别图像的过程中,要想应用卷积神经网络,我们还需要解决一个技术难题,那就是设计适当的卷积神经网络结构。要达到这一目的,就需要构建相应模型,而目前模型构建的方法主要有两种,可通过从头设计、现有模型微调两种方法来实现。第一步,以彩色图像为基准,展开对比实验,观察这两种方法在彩色图片测试中,哪个的分类更为精准,

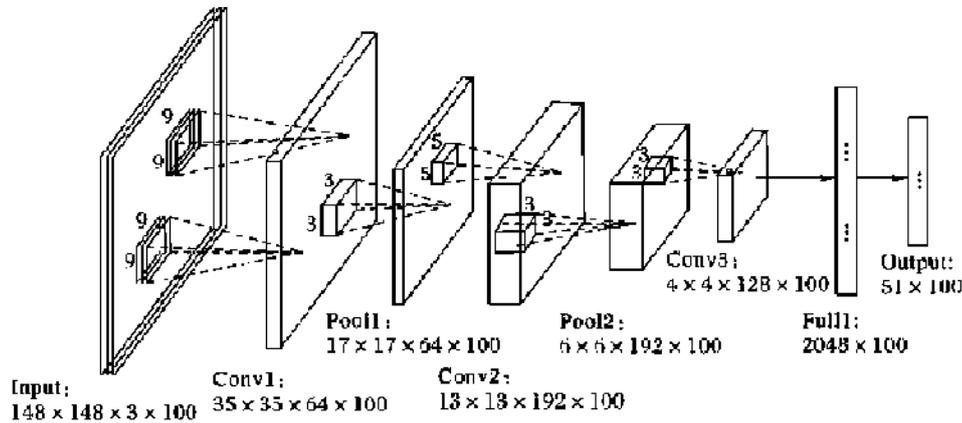


图 2：深度卷积神经网络结构

将精度更高的方法，作为训练模型的最最终方案。通过在测试集、训练集上展开多次实验，笔者发现，对于 RGB-D 物体数据，深度卷积神经网络结构的选择，以图 2 为佳。经对比实验，发现微调经典 AlexNet 网络，其效果要优于笔者自行设计的结构，之所以得到这一结果，可能是因为研究中所用数据库，从体量看略显不足，无法对多层的神经网络展开拟合。总体来看，本研究创建的模型，包含卷积层 3 个，全连接层 1 个。

1. 第一个卷积层之前的操作

1.1 图像预处理

借助深度卷积神经网络，我们可以对图像特征进行提取，而不用处理原始的底层图片，这提供了极大便利，因为我们可以将原始图片向网络模型内直接输入。计算机的计算难以完全通过增加训练数据来实现大规模计算，需要给训练数据添加先验知识，这就是我们说的数据增强处理方法。因此，在将数据输入，并且到达第一个卷积层时，需要我们对图像预处理操作，将图像分辨率适当调整，使之成为 $148 \times 148 \times 3$ ，再进行划分，分别是训练集、测试集，测试集中，数据是从 51 类物体中获取图片，最后形成，而其他图片则用来作为训练集的素材。我们对这一些素材求取均值，进一步计算，通过模型的训练及测试，将均值去除，数据之间的相似性大幅下降，这样，整体的训练速度会大大提高。

1.2 卷积操作

根据前面的论述可知，对所有图片卷积操作之后，将函数予以激活，可以得到 64 个卷积图，特征图与卷积模板保持对应的状态，特征图符合下公式：

$$h=w-(x-k)/s-1$$

h 为特征图高度， w 为宽度， x 为输入图像的宽度或高度， k 为卷积模板高度或宽度， s 为步长。在这一过程中， x 值为 148， s 为 4， k 为 9。

2. 第二卷积层

第一个卷积层结束后，需要向第二个卷积层过度，这一过程中，需要通过以下步骤实施：

2.1 选择激活函数

激活函数是为了引入非线性因素，这是因为线性函数在表达上存在欠缺，而激活函数能够避免其存在的问题。激活函数和神经元呈现对应关系，如本研究选取的函数如下：

$$f = \max(0, b + \sum_{i=1}^n x_i w_i)$$

其中 n 为输入神经元个数。

2.2 最大池化

卷积神经网络中，所提取特征，不会发生平移，不会出现旋转，输入数据之后，不会对这一网络造成直接影响。若提取的特征需要保持一定变换不变性，则需通过采样法，将卷积后形成特征图的分辨率降低，也就是应用 pooling 技术。笔者在尝试中，在激活层之后，加入了 pooling 层，这样提取特征图后，可以按照步长 2 的方式，依次选取 3×3 区域内最大的数值，病区可获得各区域均值，获得特征图。

2.3 局部相应归一

Pooling 层后面还有归一化层，这就是局部相应的归一化。将这一技术加入其中，可对隐藏层输出的较大激励予以抑制，促使模型泛化能力提高。其思想为：对于局部输入区域，实施归一化处理，方式为对相邻特征图内激励实施归一化^[5]。

2.4 自全连接层至输出层

从第二到第三个卷积层，从第三卷积层到全连接层，其技术是类似于前文所述的，都是从卷积层开始，到激活层、池化层、归一化层，之后再次进入到卷积层，进入归一化层，最终进入全连接层。

卷积层每张特征图，均代表了图像特征之一，随着层数的增高，其特征的抽象程度也就越高。全连接层是为了将底层所有特征图特征综合起来，通过 SVM 分类。本研究中，神经元共 2048 个，各神经元均与上一层全部神经元连接。在全连接层输出归一后，加入了 Dropout 技术，防止训练模型发生过拟合^[6]。测试过程中，需要全连接层上全部神经元激励，均和比例系数相乘，也就是从不同模型预测平均值来计算。

2.5 到反向传播的过程

模型输出层以 Softmax 分类器展开分析, 可对结果进行提前判断, 输出不同结果的概率情况。RGB-D 物体数据库中, 所含物体共 51 种, 这一分类器输出的数量也为 51 个, 且输出的数值是和图片中各类别的预测结果呈现对应关系的。

在反向传播中, 需要明确传播的算法^[7,8]。神经网络内的参数学习, 一般是采取随机梯度下降的方法, 主要是学习最优的权重参数, 尽可能的减少损失函数值, 而参数更新, 是对函数权重参数偏导数计算的重要方法。

四、结束语

传统人工提取图片特征的方法费时费力, 难度较大, 微调深度卷积神经网络学习法, 能够对图片特征分层、深入提取, 且可将深度数据进行转换, 实现 RGB-D 物体数据可中物体识别功能, 且精度较高。值得一提的是, 这一方法也有其局限性, 主要表现在其对模型图片的分辨率、通道数等都有要求, 至少需要这两方都是固定的。相信随着科技的不断进步, 在深度卷积神经网络基础上构建物体识别算法的过程会不断完善。

参考文献:

[1] 陈胜娣, 魏维, 何冰倩, 等. 基于改进的深度卷积神经网络的人体动作识别方法 [J]. 计算机应用研究, 2019, 36(03): 945-949.

[2] 赵德安, 吴任迪, 刘晓洋, 等. 基于 YOLO 深度卷积神经网络的复杂背景下机器人采摘苹果定位 [J]. 农业工程学报, 2019, 35(03): 172-181.

[3] 张顺, 龚怡宏, 王进军. 深度卷积神经网络的发展及其在计算机视觉领域的应用 [J]. 计算机学报, 2019, 042(003): 453-482.

[4] 孟庆成, 高朋瑞, 郭兰伟, 等. 基于多阶段三维深度卷积神经网络的计算机辅助诊断系统在肺癌筛查中的应用 [J]. 中华放射学杂志, 2020, 54(06): 552-556.

[5] 程嘉晖. 基于深度卷积神经网络的飞行器图像识别算法研究 [D]. 杭州, 浙江大学, 2017.

[6] 程嘉晖. 基于深度卷积神经网络的飞行器图像识别算法研究 [D]. 杭州, 浙江大学, 2017.

[7] 孙平安, 祁俊, 谭秋月. 利用卷积神经网络改进迭代深度学习算法的图像识别方法研究 [J]. 计算机应用研究, 2019, 036(007): 2223-2227.

[8] 周天怡. 基于深度卷积神经网络的目标检测算法研究 [D]. 哈尔滨, 哈尔滨工业大学, 2019.

作者简介: 王华溢 (1999-10), 男, 汉族, 河南焦作, 本科在读, 新疆大学, 研究方向: 计算机视觉