

本科毕业论文

题 目：基于深度学习的农业病虫

害检测算法

院 （部）： 计算机科学与技术学院

专 业： 计算机科学与技术

班 级： 计科143

姓 名： 朱忠栋

学 号： 20141111110

指导教师： 李盛恩

完成日期： 2018年5月30日

目 录

**目录的宽度和页眉页脚线的宽度基本冲齐。**

**生成目录的时候保持目录的默认宽度，不要强制设定宽度**

一级标题，黑体，4号。两字间空2格。后面的“结论”、“谢辞”、“附录”中，两字间均空2格

**罗马字符表示的页码手工格式化为**

**Times New Roman,小四号**

黑体，3号，居中

两字中间空2格

摘 要··································································Ⅲ

**ABSTRACT手工格式化为**

**Times New Roman,四号**

ABSTRACT····························································Ⅳ

**正文部分的一级目录手工删除制表符前导符和页码部分**

1 前 言

1.1 空腹桁架钢框架的特点及研究意义········································1

**手工格式化宋体，小4号，1.5倍行距**

1.2 空腹桁架钢框架的研究现状··············································3

1.3 现有研究的不足及本文的研究内容········································5

**所有的制表符前导符都手工格式化宋体，小4号**

2 空腹桁架钢框架有限元建模及验证

2.1 引言··································································8

2.2 弹塑性分析方法简介···················································12

2.3 ANSYS在空腹桁架钢框架弹塑性分析中的应用····························18

2.4 ANSYS分析模型正确性检验············································20

2.4.1 ANSYS分析模型概述················································21

2.4.1.1 ANSYS分析模型概述·············································21

目录最多列至三级标题。下级标题依次比上级标题缩进1格

2.5 小结·································································21

3 空腹桁架钢框架受力性能有限元分析

3.1 引言·································································23

3.2 空腹桁架钢框架与普通钢框架力学性能对比·······························26

**所有的页码都手工格式化为**

**宋体,小四号**

3.3 影响空腹桁架钢框架力学性能的因素·····································29

4 空腹桁架钢框架极限承载力试验研究

4.1 试验目的·····························································30

4.2 模型设计依据·························································32

4.3 试验概况·····························································35

4.4 试验过程描述·························································38

4.5 试验结果·····························································45

目录及中外文摘要的页码使用Ⅰ、Ⅱ、Ⅲ……等符号连续编排，字体Times New Roman，小5号

4.6 小结·································································55

5 结 论································································56

谢 辞··································································57

参考文献·······························································58

附 录··································································59

摘 要

在农作物病例检测中，许多检测技术被应用于此领域，但由于传统机器学习方法检测准确率太低，仍然存在农作物病虫害防治不及时的问题，严重影响农业生产的质量和产量。近些年深度学习的快速发展，基于卷积神经网络的目标检测算法，在检测效率和准确率上有很大的提升。

本论文研究现基于区域的卷积神经网络（RCNN）算法在农业检测上的应用，首先总结我国在农业病虫害检测方面的现状，然后介绍了深度神经网络和RCNN算法的原理，按照RCNN的思想，针对具体的农作物病例，重新设计网络结构，并且讨论了神经网络训练过程中出现的问题，结合卷积神经网络的最新优化技术，对算法进行改进。测试表明在该农作物病例数据集上取得了突破性的成果。

关键词：深度学习；农业病虫害；目标检测；卷积神经网络

**Detection algorithm of agricultural diseases and insect pests based on deep learning**

英文题目：Times New Roman，小3号，居中，加黑，1.5倍行距

Times New Roman，小3号，1.5倍行距

**ABSTRACT**

Vierendeel truss-steel frame has wide technical application that can be applied to the staggered truss structure and transfer structure. It has important theory significance and value of technical application to research the behavior, the ultimate load-carrying capacity and structural destroy mode.

Times New Roman，小4号，1.5倍行距

·········

关键词之间用“；”分开，Times New Roman，小4号，1.5倍行距。最后一词末不加标点

与正文另起一行。Times New Roman，4号，加黑

**Key Words:** Deep learning；Agricultural diseases and insect pests；Object detection；Convolutional Neural Networks

1绪 论

1.1 研究背景

1.1.1 病虫害防治中的问题

中国作为传统的农业大国，农业在我国经济中占有举足轻重的作用，同时也是经济发展的基础。这个基础产业在目前却存在很多问题，例如农产品质量不高，难以保障现代农业的发展等。在农业种植中，病虫害一直是阻碍农作物生长的重要问题，直接导致了农产品的质量和产量的下降。据统计，我国农作物病虫害呈多发重发态势，每年发生面积近70亿亩次，因防控能力不足每年造成粮食损失近500亿斤、经济作物损失350多亿斤。由于我国农业的若干特点，如农民文化素质差，农业领域的专家和科技人员紧缺，和农民普遍利用自己长时间的传统种植经验来防治病虫害，这种基于经验的检测方法具有很大的局限性，造成了滥用农药，防治不准确，不及时的问题，尤其在面对各种新型病虫害时束手无策，直接造成重大的农业损失。过渡使用农药还会使病虫害产生抗药性，进一步加剧了农业病虫害的防治难度[[[1]](#endnote-1)]。

1.1.2 深度学习的复兴

卷积神经网络在上世纪80年代就已经出现，起初用于识别邮政系统中的手写数字，但是碍于几处计算细节和计算机硬件水平，导致神经网络的计算速度非常慢，因而发展进入瓶颈。2012年图像识别大赛的冠军，仅仅做了几处细节的修改，如使用更为线性的ReLU函数作为激活函数等，大体没有发生变化，并且使用GPU加速运算，它出色的表现又引起了深度学习和计算机视觉的复兴，同时证明了深度学习的巨大潜力。之后，关于深度学习的研究开始爆发性的增长，神经网络的深度不断增加。如今，深度学习已经在图像、语音、自然语言处理、CTR预估、大数据特征提取等方面获得广泛的应用。面向图像分类的深度学习研究蓬勃发展，已经在该领域取得了很多重大突破，甚至超越了人类的表现。应用于许多商业化的图像分类问题中，甚至成为了很多新兴科技的基础。在医疗、金融、无人驾驶行业表现出革命性的成果。深度学习作为一种应用技术，在各行业中的应用已经成为目前研究的主要方向。

1.2研究意义

计算机技术正在改变人们的生活，农业想要进一步发展，必然要与计算机相结合。我国也一直看中计算机技术在农业上的应用，提高农业生产科技含量，加快我国农业发展步伐。卷积神经网络在图像分类上展现出来的巨大优势，农作物病虫害是图像分类的问题，深度神经网络也可以适用于病害检测。

结和我国农业病虫害防止困难的现状，基于深度学习的检测算法，利用计算机辅助诊断，可以很好的解决防止困难的问题，提高诊断的准确度和及时性。研究一个用于检测农业病虫害的算法势在必行。

1.3研究现状

物体检测算法在各个领域的应用，鉴于各种原因，对农业应用的不重视还在起步阶段，没有很完备的数据集

1.3.1数据集现状

病虫害往往具有明显的区域性，所以建立一种病例的数据集，可能只适用于一个区域。所以导致病虫害数据集的建立的不完善。然而神经网络需要大量的数据支持，完善的数据集对算法准确度的基础。这也是现阶段检测算法准确度不高的主要原因。

1.3.2目标检测算法

借助计算机视觉结对害虫的识别的应用研究还比较少。传统的检测方法往往需要手动设计特征。对于病虫害图像分析而言，需要对农业病例的症状非常熟悉，来定义病斑的形状、大小、颜色特征。研究者们经常使用梯度直方图（HoG）、颜色直方图。词袋、局部二值特征、SIFT特征、haar特征等方法来抽取图像的特征，然后将这些特征作为支持向量机（SVM）的输入，来训练分类器，用这个模型去预测患病的概率。除此之外特征的定义往往又很强的主观性，误差很大。传统检测方法准确率不高，且鲁棒性很差，这也是计算机辅助诊断没用走向实践的根本原因。

由于计算机技术在农业领域应用探索比较缓慢。利用深度学习技术检测农业病虫害研究还没有人尝试。 深度学习完全以数据作为驱动，可以在数据中学到不同层次的特征，挖掘出超越人工设计的特征，可以更强的表达出数据所包含的内容，没有主观性的影响，在病虫害的检测应用中有着巨大的潜力。

1.4研究内容

本文对基于卷积神经网络的RCNN目标检测算法在农业病虫害识别问题上的作用进行深入探究。针对具体农业病例，收集病例图片，以及对数据集预处理，建立该病例的数据集。探究RCNN算法的原理，按照具体应用场景，经过大量实验，设计网络结构。使用数据集训练网络，评价测试结果，记录准确度变化的过程，不断调节超参数，对网络进行调优。且在训练过程中，根据所发现的问题，借鉴最新神经网络训练优化技术，改造网络结构，进一步提升模型的准确度。

1.5论文的组织结构

本论文共有5章，每一章的内容安排如下。

第一章是绪论，介绍了利用深度学习检测农业病虫害研究的背景，意义，研究现状，并对研究的大体内容作了说明。

第二章是相关知识，本章主要介绍了与RCNN算法的基本知识，包括神经网络，卷积神经网络和目标检测算法。

第三章是算法设计，依据一个具体的农作物病例，分析研究所要解决的主要问题。并且按照RCNN原理，详细描述了网络结构设计过程以及原因。

第四章是测试优化，描述了网络训练的过程和数据集预处理等，并结合最新在神经网络训练中卓有成效的一些技术，改进算法，调节网络结构。最后给出了在病例数据集上检测的准确度，和对结果的评价。

第五章，总结所做的工作，以及未来需要努力的方向。

2神经网络与深度学习

2.1神经网络

2.1.1感知器

神经网络的基础是感知器，神经网络就是由许多感知器组合而成的。这是一种人工神经元模型，模仿神经元做出反应的过程。感知器原理可以被看作依据权重来做出决策的设备。一个感知器模型（如图2.1）接受几个二进制的输入x1,x2,x3…,并产生一个二进制的输出。权重可以理解为相对于输出，输入的重要程度。

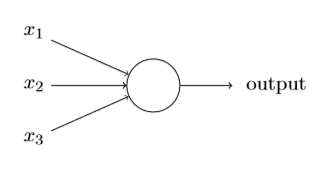


图2.1.1 感知机模型

感知器的代数描述如下，输出由所有输入乘以权重加起来的总和小于或者大于阈值所决定的。

(2.1.1)

多层感知器：第一层的输出作为第二层感知器的输入，第二层中的感知器比第一层中的感知器作更复杂的决策，第n层同理，这便是神经网络的雏形。

2.1.2分类器

分类器根据数据的特征用来决定数据的分类，用于神经网络的分类器主要有KNN，SVM，Softmax。

KNN分类器是K近邻分类器，对于被预测数据所属的种类，由K个最靠近他的数据的种类所决定。它判断数据是否属于某一类别的依据是当前数据与种类的距离，具体到图像分类上，有如下两种方法

L1距离（Manhattan distance） (2.1.2a)

L2距离（Euclidean distance） (2.1.2b)

这两种方法根据所面对的问题灵活选择。

KNN是比较简单的分类器，在训练阶段只是负责记忆，在预测阶段将待测数据与记忆的所有数据进行比较。因此KNN分类器在训练阶段的时间为，而在预测阶段的时间是。这显然是是大问题，我们希望分类器将时间花在训练上，而不是预测上。因为在预测阶段花费时间短我们才能在实践中实时做出决策。

SVM和Softmax分类器属于另一种分类器他们将时间花在训练上而不是预测上。这种方法由两部分组成：一个是评分函数（score function），它是原始图像数据到类别分值的映射。另一个是损失函数（loss function），它是用来量化预测分类标签的得分与真实标签之间一致性的。该方法可转化为一个最优化问题，在最优化过程中，将通过更新评分函数的参数来最小化损失函数值。

2.1.3激活函数

我们想要让神经元学习变为可能，就需要任何权重或偏置的微小变化可以引起输出的微小变化。感知器的缺点是：网络中单个感知器上一个权重或偏置的微小变动有时会引起输出的完全翻转。激活函数可以解决这个问题，可以使神经网络拟合任意给定的函数。如果不使用激活函数，那么即使再多的神经网络和一层没有区别，只是在作矩阵相乘。

经常在神经网络中使用的激活函数如下：

Sigmoid： (2.1.3a)

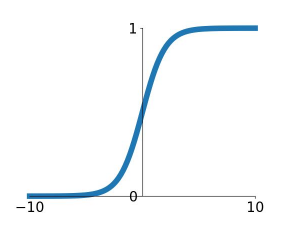


图2.1.3a Sigmoid函数形状

Tanh：tanh(x) (2.1.3b)

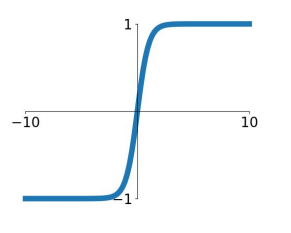


图2.1.3b Tanh函数形状

ReLU： (2.1.3c)

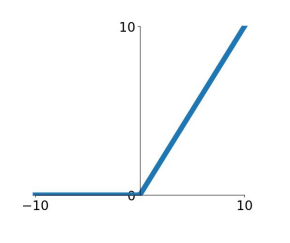


图2.1.3c ReLU函数形状

Leaky ReLU： (2.1.4d)

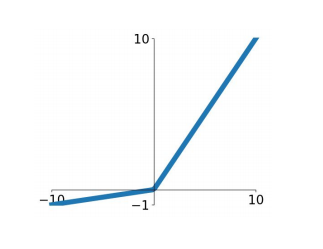


图2.1.3d Leaky ReLU函数形状

ELU： (2.1.3e)

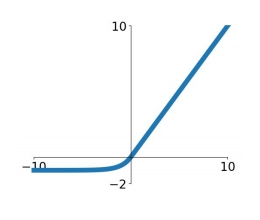


图2.1.3e ELU函数形状

2.1.4损失函数

为了找到合适的权重和偏置，我们需要有一个评价函数，这就是损失函数。

神经网络的损失函数：

(2.1.4a)

这个公式中符号表示的含义为：y(x)表示期望输出，w表示所有网络中权重的集合，b是所有的偏置，n为训练的个数，a表示输入为x时输出的向量。a取决于x、w、b。

我们感兴趣的是能正确分类的图像的数量，为什么不直接最大化这个数量，而是最小化一个类似二次代价的函数来间接评价呢？

因为被正确分类的数量关于权重并不是一个平滑的函数。大多数情况下对权重和偏置的微小变化不会明显的影响被正确分类图像的数量。我们就会陷入一个泥潭，很难改变权重和偏置做出微小的变化来提升性能。

Multiclass SVM Loss：

(2.1.4b)

(2.1.4c)

这是SVM分类器的损失函数，也被称为合页损失函数（Hinge loss）。随着的升高，损失函数不断下降，只到超过某一阈值。这个公式里的1被称为安全边际值（dela）。只要分类器得出的结果中正确分类的结果大于不正确分类结果的，并且超过安全边际值，则认为分类器良好没有损失，否则将计算损失。这里有一个有用的调试技巧，当使用一些很小的值来初始化w，你的分数结果在训练的初期倾向于呈现较小的均与分布的值，所有的得分都近乎为0并且差不多相等，损失函数预计会是C – 1。若果不是程序中就有bug。理解算法原理，使用不同的输入对程序输出结果做出预测是验证程序是否正确的重要方式。

Softmax Loss：

相对于SVM分类器，SVM只关心最后得分相差多少，Softmax给得分赋予新的含义，使用每个类的得分去计算相应的概率。具体损失函数（交叉熵损失函数）如下：

(2.1.5d)

每个概率都介于0和1之间，所有概率和为1.

(2.1.5e)

真实类别的概率再取负值。加负号因为损失函数是度量坏的程度而不是好的程度，加上负号跟自然。在这公式中，因为有指数运算，容易造成运算溢出，所以在实际编程中会在类别得分中，选择一个最大的得分，然后所有的得分都减去这个最大值。

在实际应用中，这两种分类器的表现是相似的，根据测试的结果来选择表现更好的分类器。SVM分类器中，正确分类相较于不正确分类，已经得到了比边界值还要高的分数，那么就认为损失值是0，不再去优化。Softmax分类器中，对于分数是永远不会满意的：正确分类总能得到更高的可能性，错误分类总能得到更低的可能性，损失值总是能够更小。由于Softmax的特性，它可以获得更加精细的分数。而SVM的特性也可以被看作优势，可以提高分类器的鲁棒性。

2.1.5梯度下降

我们有了损失函数来评价神经网络的好坏，那么如何找到损失函数最小值点呢。直接找到是不可能。有两种方法：第一种是一个很自然的想法就是随机搜索，但在高维空间中搜索的范围太大了，这个方法跟本不能在实践中操作。

我们可以将函数想象成一个山谷，假设有一个小球从山坡的斜坡上滚落下来，我们日常经验告诉我们这个球最终滚到谷底。我们可以用这一想法找到最小值。一开始小球随机选择一个起始位置，然后模拟球体滚到谷底的运动。计算函数的导数可以告诉我们山谷中局部的形状，由此我们知道球体怎么滚落。所以另外一种方法就是计算斜率。计算斜率又有两种方法，第一种是有限差分法，根据导数定义，保持一个变量变化，其他参数不变，来计算变化前后的输出差值。这个方法我们每对一个参数求偏导，就要两次在神经网络中的前向传导。在动辄几千万的神经网络面前，时间花费巨大，也是不可行的。但可以使用来检验自己程序中梯度计算是否正确，设置一组具体的参数，改变某个参数的值，来计算变化前后的差值，观察程序输出的斜率是否一致。另一种方法是利用微积分的公式，直接推到梯度。

我们只要告诉球体“现在向下”，沿梯度的相反的方向，是函数在该点下降速度最快的方向，因此我们每次按照梯度相反的方向计算小球移动的位置。这个过程被称为梯度下降。更新规则可以写为：

(2.1.5a)

公示中的，被称为学习速率，他决定小球一步要走多远，这是一个超参数，神经网络训练中首先要经过大量测试，来确定这个值，他要保证损失函数可以平滑的降到最低点，太大会导致小球越过最小值点并且来回震荡，太小又会导致神经网络训练时间的加长，很难快速的到达最小值点。反复使用这个公式，来计算下一次移动位置，直到C到一个最小值。

总结一下，梯度下降算法工作方式就是重复计算梯度，然后沿着相反的反向移动，沿着山谷滚落。

2.1.6反向传播

梯度下降算法中最重要的就是计算梯度，我们已经知道可以用微积分公式推到梯度的计算公式，但是在正向推导中，会发现很明显的问题，就是好多节点的梯度值被重复利用，造成了极大的计算浪费，解决方法就是从神经网络的最后一个节点开始计算，反向推到，前一个节点只需知道上游传过来的上游有梯度，与自己本身节点的梯度相乘，然后继续向第一层几点的方向传播，这样反向遍历一次神经网络就可以将所有节点的梯度计算出来。具体实例如图2.1.6a所示。

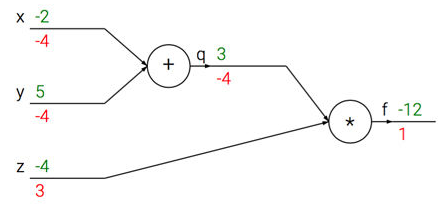


图2.1.6a 神经网络前向传导与后向传播模型

反向传播是链式法则的递归调用，反向传播是一个优美的局部过程。在整个计算线路图中，每个门单元都会得到一些输入并立即计算两个东西（如图2.1.6b）：这个门的输出值和其输出值关于输入值的局部梯度。门单元完成这两件事是完全独立的，它不需要知道计算线路中的其他细节。然而，一旦前向传播完毕，在反向传播的过程中，门单元门将最终获得整个网络的最终输出值在自己的输出值上的梯度。链式法则指出，门单元应该将回传的梯度乘以它对其的输入的局部梯度，从而得到整个网络的输出对该门单元的每个输入值的梯度。

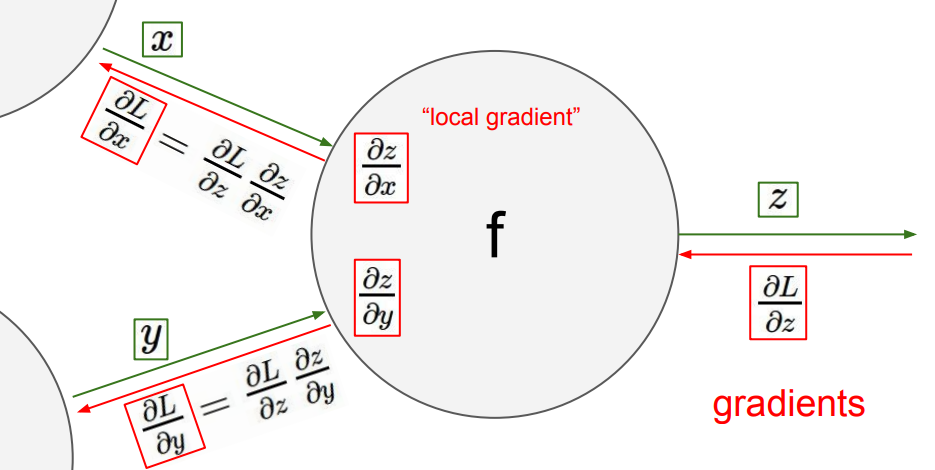


图2.1.6b 后向传播计算过程图

2.1.7正则化

正则化是帮助减轻神经网络过拟合的技术。假设有一个数据集和一个权重集能够正确的分类每个数据。问题在于这个权重集可能并不唯一。我们希望模块能够以某种方式选择跟简单的权重，以此来消除模糊性，迫使线性回归朝着地接多项式演进。我们可以在损失函数上添加一个正则惩罚项，最常用的是正则惩罚范式是L2范式：

(2.1.7a)

加入正则化的损失函数为：

(2.1.7b)

直觉上看，规范化的效果是让网络倾向于学习小一点的权重，因为我们希望网络输出的结果不是由某一参数决定，而是由整体的结果决定的，正则化可以抑制大的参数，是参数更加的平均化，提高神经网络的鲁棒性。它可以被理解为一种寻找小的权重和最小原始代价函数的折中。公式2.1.7b中的，是一个超参数，用来控制前后两部分的重要性，越小就越偏向最小化原始代价函数，反之，倾向于小的权重。

小的权重在某种程度上来说意味着更低的复杂性，也就是对数据给出了一种简单却更加强大的解释。不去规范化偏置的原因是有一个大的偏置并不会像权重那样使神经网络对输入变得太多敏感。允许大的偏置能够让网络更加的灵活。大的偏置可以让神经元更加容易饱和。

2.2卷积神经网络

卷积神经网络是目前在图像分类领域最流行的神经网络。它由神经网络发展而来，与神经网络相似，也是由神经元构成，具有学习能力的权重和偏置。整个网络依旧是一个可导的评分函数，在网络的最后一层往往是一个全连接层，也有一个和神经网络一样的损失函数。卷积神经网络的结构基于一个假设，即输入数据是图像，基于该假设，我们想网络结构中添加一些特有的性质，使得前向传播函数实现起来更高效。它在特征采集上的结构上变了，大幅减少了网络中参数的数量。卷积神经网络的结构如图2.2

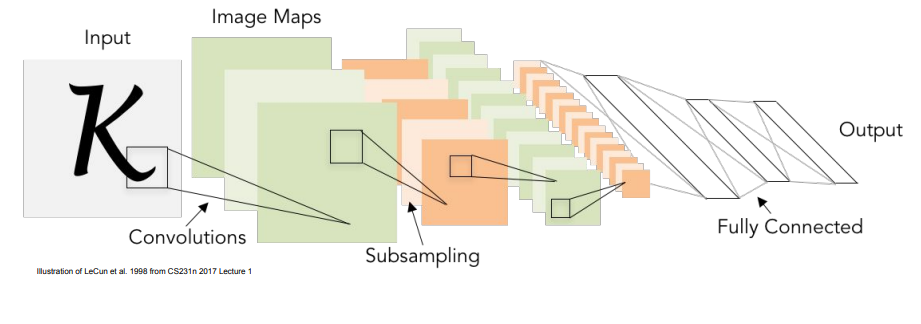


图2.2 卷积神经网络实例结构模型

2.2.1卷积层

全连接神经网络有一个缺陷，没有考虑空间结构，在完全相同的基础上对待相距很远和相聚很近的输入像素。神经元有局部连接性，不关注输入图片的全部，关注图像空间的一个局部区域，。卷积层依据这种思想，可以很好的保留空间结构，很好的适应图像的平移不变性。卷积层采用了两种基本概念：局部连接和共享权重。

局部连接就是让每一个神经元与输入数据的一个局部区域连接，该连接的空间大小叫做神经元的感受野，如图2.2.1中，感受野为5\*5大小。感受野的尺寸是一个超参数。在深度方向上，连接的大小总是和输入量的深度相等。卷积层的输出大小由3个超参数控制：深度，步长，零填充。输出数据大小的计算公式为：

(2.2.1a)

共享权重是对于每个神经元使用相同的权重和偏置这大大减少了神经网络的参数。这意味着每个神经元检测相同的特征。卷积核按步长扫描输入数据，输出一张激活表。有几个卷积核就会有输出几张激活表。

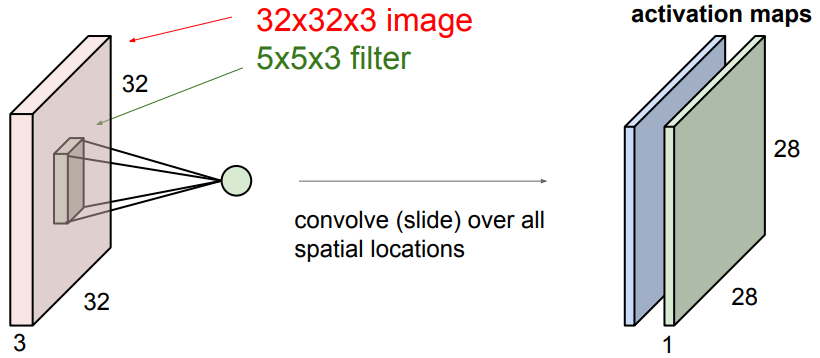


图2.2.1 卷积层模型

2.2.2池化层

池化层通常跟在卷积层之后使用。他要做的是简化从卷积层输出的信息。池化层操作类似于卷积层操作，也包含池化范围，步长，零填充等，不过操作不再是卷积而是在池化范围按照池化规则对激活值进行操作，例如最大值池化（如图2.2.2）是在池化范围内选出最大的激活值，平均值池化是在所有在池化范围内的激活值求平均值。对于图像处理，最大值池化要好于平均值池化，最大值池化可以保留更明显的激活特征。池化操作一般不在深度方向上池化处理，而是在平面上。所以池化层输出的深度与输入是一样的，池化操作的步长一般是不重叠的。总结一下，池化层就是尽力用一个数值来表示整个区域，目的为了降采样处理，经过池化层将减少75%的参数。

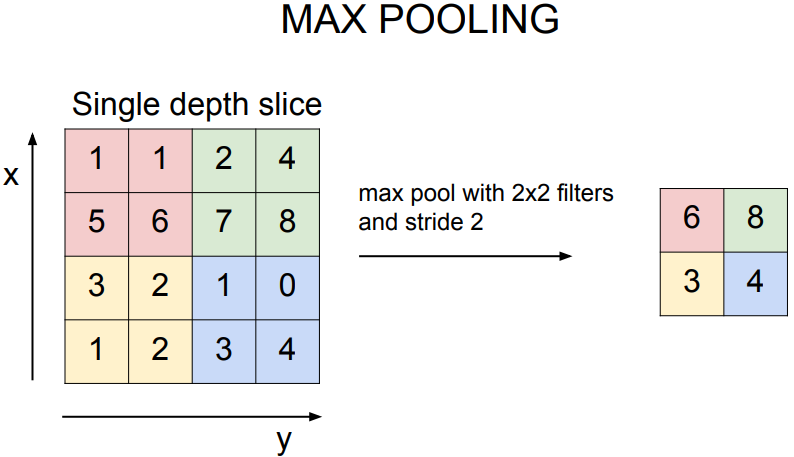


图2.2.2 最大值池化操作

2.2.3全连接层

全连接层通常在卷积神经网络的最后，在图2.2的最后3层就是全连接层，它用来将最后的输出映射到线性可分的空间。这一层是传统的神经网络，起到分类器的作用确切的说，这一层将输入数据的每一个神经元连接到每一个输出神经元，将卷积层提取的特征汇聚起来，得到相应类别的分数。全连接层将卷积层产生的特征图映射成一个固定长度(一般为输入图像数据集中的图像类别数)的特征向量。这个特征向量包含了输入图像所有特征的组合信息，虽然丢失了图像的位置信息，但是该向量将图像中含有最具有特点的图像特征保留了下来以此完成图像分类任务。

2.3 基于卷积神经网络的目标检测算法

目标检测与识别（如图2.3）作为计算机视觉领域的基石，也越来越受到重视。在实际生活中应用也越来越广泛，例如目标跟踪，视频监控，信息安全，自动驾驶，图像检索，医学图像分析，网络数据挖掘，无人机导航，遥感图像分析，国防系统等。目标检测算法的任务是在指定的一张图片中找出要识别的目标。所以不仅仅是要识别出图像中的物体是什么，还要把物体的范围标识出来。由于给定的图像没有所包含物体的数量不确定，任务难点在于待测区域的提取与识别。

任务具体的过程有：

1. 从场景中选出待测区域。
2. 识别待测区域的目标。
3. 精调分类模型，有效候选框的位置精修。

如今目标检测与识别的研究方法主要由两大类：基于传统图像处理和机器学习算法的目标检测与识别方法；基于深度学习的目标检测与识别方法。

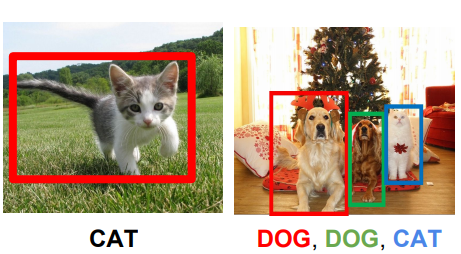


图2.3 目标检测任务

2.3.1 传统目标检测算法

传统的目标检测与识别方法主要可以表示为：候选区域确定、目标特征提取、目标识别、目标定位。首先按照大小固定的窗口在图像水平和水平方向平移以遍历图像，然后提取特征，这里所用到的特征都是人为设计的，例如SIFT (尺度不变特征变换匹配算法Scale Invariant Feature Transform), HOG(方向梯度直方图特征Histogram of Oriented Gradient), SURF( 加速稳健特征Speeded Up Robust Features),等。通过这些特征对目标进行识别，用到SVM等分类器，然后再结合相应的策略对目标进行定位。但传统算法存在的问题是：滑动窗口虽然简单，但产生大量无效窗口，严重影响检测的速度。其次特征主观性大，鲁棒性低，导致准确率不高。

2.3.2 RCNN

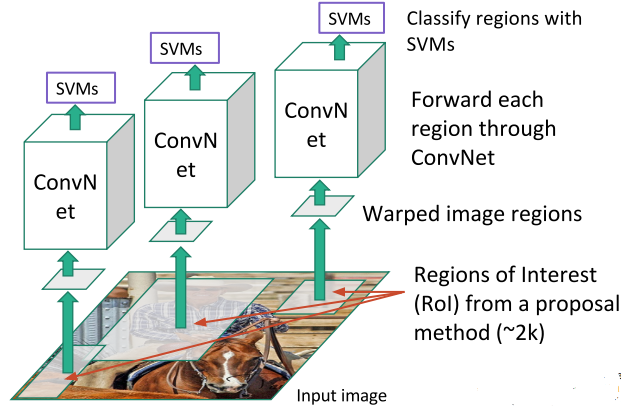


图2.3.2 RCNN结构模型

RCNN(Regions with CNN features)是将CNN方法应用到目标检测问题上的一个里程碑，是利用深度学习进行目标检测的开山之作。这个算法借助卷积神经网络强大的特征提取和分类的性能，通过使用Region Proposal方法实现将目标检测任务转化为分类问题。

算法可分为三步（如图2.3.2）：

1. 候选区域的选择

选择搜索算法，通过算法利用图像中的色彩，梯度，边缘等信息，选出有可能存在目标的区域，Selective Search方法从一张图像生成约2千个候选区域。候选区域生成和后续步骤相对独立，实践中可以利用任意算法进行。

1. CNN特征提取

首先预处理候选区域，因为卷积神经网络接受的区域大小固定，所以候选区域归一化为统一的大小，后续分类中使用AlexNet模型，这里大小固定为227\*227。然后将归一化后的候选区域送入卷积神经网络得到固定长度的特征。

1. 分类与边界回归

类别判断，对于每一类，使用一个SVM二分类器进行判别，将上一步输出的特征向量输入SVM分类器，边界回归（bouning-box regression）得到精确的目标区域。

目标检测的标注数据很少，如何用少量的标注数据训练高质量的模型，可以使用迁移学习。迁移学习是使用已经训练好并且表现出良好的网络模型，冻结网络结构中的特征提取阶段的参数，根据具体任务可以替换卷积神经网络最后的分类器，只训练特征分类阶段的参数。RCNN中先使用ILSVRC2012数据集训练一个1000类的分类器模型，然后使用该模型来初始化我们的CNN模型参数，使用我们的训练和验证数据集合进行微调。

因为CNN是有监督的学习，我们需要对候选区域的类别进行标注。我们要先标记出目标的真实边框和类别，再计算候选区域与真实边框的交集和并集，在计算交集和并集的比例，如果这个比例大于0.5，则将这个区域标记为正样本，否则就是负样本。在SVM训练时，这个比例小于0.3的区域为负样本。其他的区域忽略，真实边框作为正样本。这样做是因为，CNN训练时做了比较宽松的标注，候选区域只包含物体的一部分，那么我也把它标注为正样本，可以防止CNN过拟合。而SVM训练的时候，SVM适用于少量样本训练，对于训练样本数据比较严格，只有当候选区域将整个样本包含进去才把他标记为正样本。

RCNN在PASCAL VOC2007数据集上的检测结果由历史上的32.3%提升到66%，证明了深度学习在目标识别上的巨大潜力。

3算法分析与设计

3.1 问题描述（对所解决问题的完整描述和定义）

3.1.1 采用模型

3.1.2 问题定义

3.1.3 问题的解释

谢 辞

本文是在指导教师张九光教授的悉心指导下完成的，从论文选题、课题调研、试验指导、理论分析到论文撰写，无不倾注了老师的心血和汗水。向所有曾经关心和帮助过我的老师、同学和朋友致以诚挚的谢意！

参考文献

[1] 李加成.我国农业病虫害防治现状及建议分析[J].种子科技,2017,35(06):106+108.

附 录

一、公式推演

黑体，小3号，居中，1.5倍行距，两字间空1格，若无附录的内容，将该标题删除

二、符号说明

……

宋体， 5号， 1.5倍行距

1. [↑](#endnote-ref-1)