

## 摘要

我国是一个农业大国，有多种农作物在大规模种植，虽然历史悠久，但还是存在部分问题制约着农业的发展。其中农作物的病虫害就是限制农业发展的一大因素。例如，番茄和马铃薯的产量都会受到早疫病、晚疫病等数十种病害的影响，还有玉米也会受到灰斑病和叶枯病等病害的影响。及时采取措施对病株进行针对性治疗，便可以在一定程度上减少农作物产量的损失。目前，我国大多数农民对农作物病虫害的判别方法只停留在传统的实地目测上，通过农作物的外观表面形态等进行判断，往往多依赖于个人经验。所以如何快速、准确地判断农作物病虫害，已经成为智慧农业必须考虑的问题。为了解决上述问题，本文设计了多尺寸卷积核及最大值池化的轻量级卷积神经网络，一定程度减少了参数量，提高了识别正确率。为了进一步提高识别正确率，选取 ImageNet 进行迁移学习，提出了对卷积层进行批量归一化，加快了收敛速度，减轻了计算负担，并且通过训练最终能达到较高的准确率。本文通过对卷积神经网络进行迁移学习，设计并实现了能够快速、准确地判断农作物病虫害的模型，可以较好地解决农民在种植农作物的过程中遇到病虫害诊断准确率导致农作物产量减少等问题。

关键词：病害识别 深度学习 卷积神经网络 迁移学习

## 1 前言

### 1.1 研究目的和意义

长期以来，农业生产在现代社会中都占据着重要地位，农作物的产量以及品质直接影响人们的生活。因此，在农业生产中田间农作物病虫害的有效识别和防治显得尤为重要。准确识别病虫害的类型是监视种群动态，揭示灾难规律并科学指导预防和控制的基础(杨红珍，张建伟，李湘涛，2008；Neethirajan S, Karunakaran C, Jayas D S, 2007；张红涛，毛罕平，2008)。目前对植物病虫害的检测和鉴定主要依靠少数植物保护专家和农业技术人员来完成，通常的方法是使用人眼根据有害生物的形态特征，比如颜色、大小、形状等进行识别，然而人的视觉系统也往往存在各种不足，如主观性、局限性、缺乏持久性等(方如明，蔡健荣，许俐，2008)，这都将大大降低病虫害识别的准确度。鉴于多种病虫害，每位植物保护专家或农业技术人员只能使用其专业知识识别部分病虫害；对于一些不常见的病虫害，往往需要充分了解其危害状况，还要查阅相关资料，甚至要结合显微镜才能进

行准确的判断，这通常会延迟最佳的防治时间。此外，具备专业的农作物病虫害诊断知识的植保研究人员人数有限，所以往往无法满足广大农业生产者的实际需求（刘立波，2010）。因此，如何能够快速准确的诊断病虫害并及时做出相应的措施变得尤为重要。 1.2 国内外研究现状

数字图像处理技术在 20 世纪 50 年代出现，在 60 年代初成为正式的学科。在过去的 20 年中，随着数字图像处理技

术越来越专业和各種识别技术的日渐成熟，数字图像处理技术在农作物营养成分的诊断和识别，农产品质量检验等方面进行了广泛研究，并且在农业领域中显示出极大的发展潜力。但是图像识别技术在农作物病虫害方面的研究起步相对较晚，文献较少（王娜，2009）。

1.2.1 国外研究现状国外对田间农作物病虫害的识别诊断研究开始于 20 世纪 80 年代。

Yuataka SASAKI 研究了黄瓜炭疽病的自动识别技术。针对不同的光谱反射特性和各种光学滤波对病虫害识别的影响，并且采用遗传算法，从分光反射特性和形状特征的角度出发，建立了识别参数，对黄瓜炭疽病进行了识别。

因为他们没有充分利用病害的颜色及纹理等特征信息，所以识别精度不高（Yutaka SASAKI, Tsuguo Okamoto, 1999）。

Mohammad Sammany 等利用遗传算法优化神经网络的参数和结构来识别农作物病虫害图像；同时使用支持向量机和神经网络两种方法来识别植物病害。后来，他们使用粗糙集来减少神经网络分类器的输入特征向量，从而高分类效率

（El-Helly M., El-Beltagy S., and Rafea A., 2004）。

深度学习是一种自动识别植物病害的有前途的方法。利用卷积网络的参数共享，稀疏交互和变量表示，可以有效地减少网络参数，节省模型存储需求，提高网络的统计效率。近年来，利用深度学习方法在植物物种识别和害虫识别方面取得了一些进展。Geetharamani 提出了一个 9 层的深度卷积神经网络的植物病害识别模型，该模型在开源数据集 PlantVillage 上对 39 个不同植物不同病害进行识别，最后该模型达到了 96.46% 的分类精度（Geetharamani G, Pandian A, 2019）。Hang 利用 VGG16 模型，Inception 模块，SE 模块，和全池平均化相结合的模型，对樱桃，苹果，玉米共 10 种病

害进行识别, 准确率达到 91.7%, 同时大大减少了训练时间和模型参数 (Hang J, Zhang D, Chen P, 2019 )。

### 1.2.2 国内研究现状

虽然国内研究农作物病虫害识别起步相对于国外较晚, 但是有一些高校和研究院长期以来都有在开展农作物病害图像识别技术的研究。

张长利等在研究自动识别番茄成熟度时使用了遗传算法训练的多层前馈神经网络, 在不同成熟度的 50 个番茄样品上进行了测试, 从结果可以得出, 人工神经网络与遗传算法相结合具有很大的潜力, 在农产品质量自动识别中具有广阔的应用前景 (张长利, 房俊龙, 2001)。

郑一力等人采用 InceptionV3、AlexNet 模型, 基于迁移学习的方法对植物叶片进行识别, 并最终在测试集上测得 95.31%和 95.40%的准确率 (郑一力, 张露, 2018)。

廖经纬等人针对农作物叶片病虫害的识别问题, 设计了一种卷积神经网络来进行识别研究。首先, 将农作物叶片的病害部位与图片的背景进行分离, 然后采用 AlexNet 网络进行识别。最终识别准确率达到 98.44% (廖经纬, 蔡英, 王语晨, 2018)。

刘永波等对 10 类常见的玉米病害进行了研究, 在训练过程中使用深度学习网络对健康的玉米图像特征进行学习, 并且采用了迁移学习的方法对玉米的病害图像特征进行学习, 其准确率达到 90%以上 (刘永波, 雷波, 曹艳, 2018 )。

谭云兰等利用深度卷积神经网络模型对自然场景中采集的 8 种水稻的病害图像进行分类识别。所采用的图像是在多株水稻的场景下, 使用了图像扩增技术对水稻病害数据进行扩充, 并利用微调方法对网络中的参数进行调整, 最终取得了较高的识别精度 (谭云兰, 欧阳春娟, 2019)。

由此可见, 近年来国内外相关研究人员都对农作物病虫害识别进行了广泛的研究, 通过对农作物病虫害图像的预处理, 然后提取病虫害各种特征参数, 最后可通过使用不同的分类器来实现智能识别的目的。

## 2 深度学习技术研究

### 2.1 深度学习阐述

深度学习(DL, Deep Learning)是机器学习(ML, Machine Learning)领域中一个新的研究方向,这个概念在上世

纪中叶被相关研究人员提出,由于当时的研究环境有限,人们不能很容易的获得大量相关数据,而且当时计算机的性能不够强大,不能对大量训练数据进行运算处理,导致深度学习没有获得较多的资源支持,发展过程变得缓慢。进入 21 世纪,人们可以通过互联网或者其他路径快速地获得大量的数据资源,同时计算机技术发展飞快,处理器性能越来越成熟,深度学习得以被广泛应用在各个领域,比如在搜索技术,数据挖掘,机器学习和个性化技术,以及其他很多相关领域都取得了重大成果。

## 2.1 神经网络

### 2.2.1 感知器

深度学习中的感知器以生物神经元为原型,感知器互相连接构成神经网络(Zhang Lei, Du Zidong, Li Ling, 2020)。图 1 为生物神经元的工作原理,图 2 为感知器的工作原理。分析图 1 和图 2 得知,感知器的  $x_i$  与生物神经元的树突相对应,感知器对输入的加权求和计算与生物神经元的输入向细胞体发送脉冲相对应,感知器的激活函数与生物神经元的脉冲相对应。感知器的工作原理为:

$$y = f(\sum_{i=1}^n w_i x_i + b) \quad (1)$$

式中,  $f$  —— 激活函数;  $b$  —— 感知器的偏置参数;

$x_i$  —— 感知器的输入;  $y$  —— 感知器的输出。

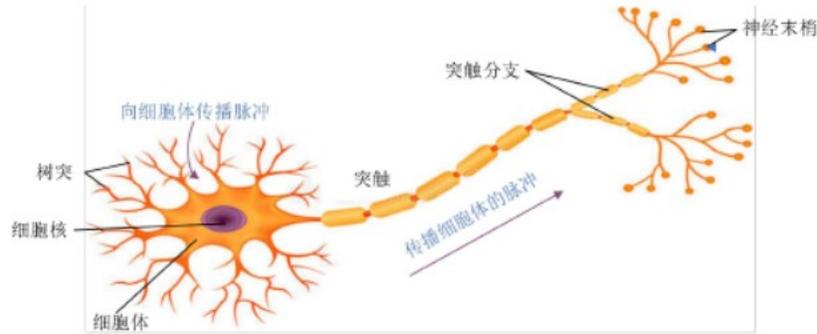


图 1 生物神经元工作原理图

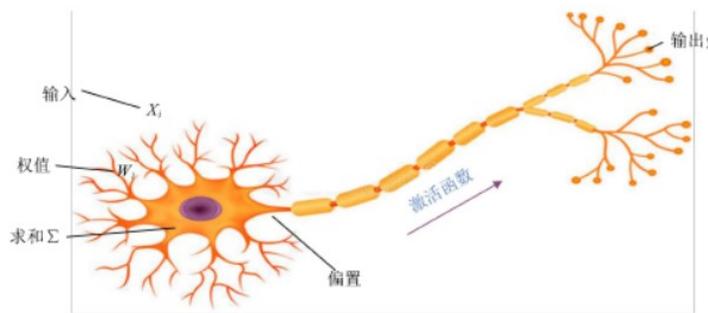


图 2 感知器工作原理图

### 2.2.2 激活函数

激活函数对于人工神经网络模型学习以及理解一些非常复杂且非线性的函数来说非常重要。激活函数向神经元引入了非线性因素，因此神经网络可以逼近任何非线性函数，从而神经网络就可以使用在许多非线性的模型中。激活操作包括连续但不全部可微和平滑非线性两种函数。常用 Sigmoid、ReLU 和 Tanh 三种激活函数，其中 ReLU 为连续但不全部可微函数，Sigmoid 和 Tanh 为平滑非线性函数。

Sigmoid 激活函数将输出控制在 0-1 的界限内，输入趋近于正无穷，输出就趋近于 1；输入趋近于负无穷，输出就趋近于 0，很容易达到饱和，会出现梯度消失等问题，导致训练结果不佳。同时函数表达式存在幂运算，使得计算求解耗时多。

Tanh 激活函数与 Sigmoid 激活函数类似，不同的是将输出控制在 -1-1 的界限内，输入趋近正无穷，输出就趋近 1；输入趋近负无穷，输出就趋近 -1。使用 Tanh 激活函数同样存在幂运算与梯度消失，依然没有改善因为饱和性导致训练结果不佳的问题，训练时间也较长。

ReLU 激活函数虽然在大于 0 的范围和小于等于 0 的范围分别呈线性，但整体呈非线性，因此 ReLU 函数收敛更快，不容易达到饱和。由于 ReLU 激活函数可以将负值输出进行归零屏蔽处理，其计算过程只需要与实数 0 进行一次比较，

因此运算速度相比前面两种函数有很大提升。ReLU 激活函数的计算公式如式 (2)，函数图像如图 3 所示。

$$f(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (2)$$

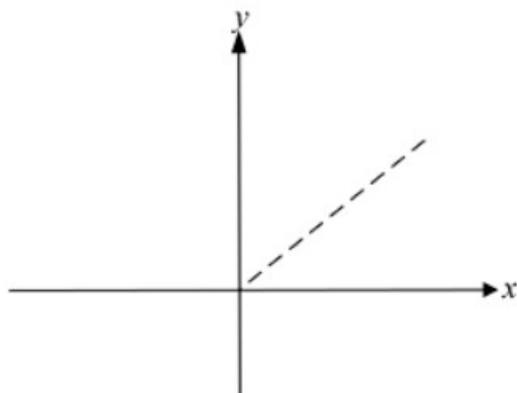


图 3 ReLU 函数

### 2.2.3 输入层、输出层及隐藏层

网络结构主要包括输入层、输出层以及隐藏层。输入层为整个网络的第一层，输出层为网络的最后一层，隐藏层可以是一层或者多层，在输入层和输出层之间，如果隐藏层有两层或者更多层，便将该网络叫做“深度”神经网络。隐藏层和输出层通常包含感知器，从而实现非线性。如图 4 为只有一个隐藏层的网络模型图，通过分析隐藏层的感知器，看出其包括输入与加权运算结果  $z$ ，对  $z$  进行激活得到  $y$  两步操作。

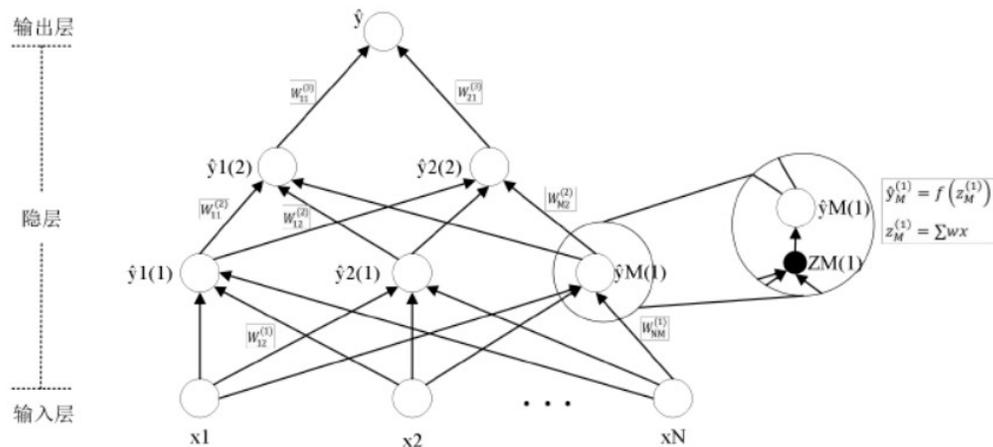
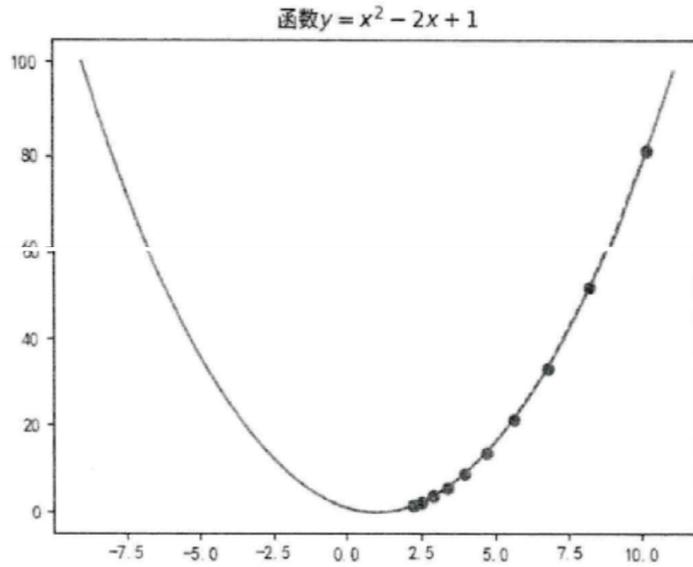


图 4 网络模型图

### 2.2.4 梯度下降算法

梯度下降算法是一种常用的优化算法，其中心思想是沿着目标函数梯度的方向来更新参数值，以实现目标函数最小值(或最大值)。除了深度学习，梯度下降算法还用于其他很多场合。核心思想是找到一个参数  $\theta$ ，使得在参数  $\theta$  的条件下能够使得训练数据集上局部损失函数  $L(\theta)$  获得最小值。损失函数越小，说明模型的分类结果与真实值越接近，神经网络的精度就越高。其工作原理如图 5 所示。

图 5 梯度下降算法原理



从图 5 中，可以看出，当参数  $\theta$  位于曲线底部时，损失函数  $L(\theta)$  将会获得最小值，此时模型的参数为最优参数，使用当前模型获得的分类精度最高。参数  $\theta$  可以通过求偏导数的方式来更新，其数学公式如式 (3)。

$$\theta_{n+1} = \theta_n - \eta \cdot \frac{\partial}{\partial \theta} L(\theta) \quad (3)$$

式中， $\eta$  为学习率，表示参数更新每次移动的步长， $\frac{\partial}{\partial \theta} L(\theta)$  为损失函数对参数  $\theta$  的偏导数。

通过公式可以看出，参数  $\theta$  的更新幅度与学习率  $\eta$  有很大的关系。使用较大的学习率  $\eta$  时，可以使模型损失函数只使用较少的训练次数就可以达到较小的损失值，快速的完成收敛，但学习率较大可能导致损失函数按照既定步长在某一非最小值附近波动，导致永远到不了最小值，因此模型的分类准确率也达不到高精度。相反，使用较小的学习率可以使得模型得分率精度更高，但训练次数需要更多，导致花费更多训练时间。所以可以将学习率  $\eta$  设为一个变量，在较少的训练次数时可以使用较大的学习率，使模型损失函数快速收敛至最小值附近，然后更换为较小的学习率，移动步长会随之减小，使损失函数更接近最优值，便可以使模型快速收敛至较高得分率。按照如式

(4) 对学习率的大小进行设置。

$$\eta_n = \eta_0 - \lambda \frac{\eta_n}{B} \quad (4)$$

式(4)中,数,  $\eta_n$ 表示当前训练次数。  $\eta_n$ 表示当前学习率,  $\eta_0$ 表示初始学习率,  $\lambda$ 表示学习率的衰减率(定值), B表示每轮学习的次

由于深度学习网络逐层深入,层层嵌套的特性,在计算深度网络目标函数的梯度时,有必要用反向传播的方式由深到浅计算以及更新参数。因此,反向传播方法是梯度下降方法在深度网络上的特定实现。无论是卷积神经网络模型还是传统的全连接神经网络模型都可以使用反向传播算法来进行训练。反向传播算法是将损失值进行前向传播的逆过程,从分类层逐层向前,最终传播到输入层,在逆向传播的过程中,使用梯度下降算法逐个更新参数权重。传播过程依赖于链式法则。

### 2.2.5 过拟合

泛化能力指的是我们所选择的神经网络模型对于未知数据的处理能力,好的模型必然具有较好的泛化能力,可以在验证集中有较高的准确率。欠拟合的原因通常是因为模型过于简单以及参数量太少,不能准确提取到每一类的关键特征,解决方法可以是增加神经网络的深度和更改神经网络模型的结构。过拟合则与欠拟合相反,它是指为了得到一致假设而使假设变得过度严格的现象,该现象会将训练数据集本身的某些特殊特征作为一般特征进行记忆的一种现象,过分的追求细节反而使得模型对于未知数据集的识别能力过低。所以为了提高模型的泛化能力,就需要防止模型出现欠拟合与过拟合现象。防止过拟合的方法有:(1)提取终止训练训练时,时刻关注模型的效果,选取合适的停止训练标准,当验证集的损失值不再降低甚至提高时将训练终止。

#### (2) 数据增强

扩充训练集的数据,从而获得更多数据进行交叉验证。可以通过旋转、变焦、平移、改变尺寸等方法生成更多的样本。

#### (3) 正则化

即在进行目标函数或代价函数优化时,在目标函数或代价函数之后加上一个正则项,一般有 L1 正则与 L2 正则等

。

#### (4) Dropout

在训练时随机地让部分隐藏层神经元进行失活处理，即对神经元权重进行置零或者隐藏处理，缓解特征之间的依赖程度，从而避免部分偏差较大的信息传递。

### 2.3 卷积神经网络

#### 2.3.1 卷积神经网络概述

卷积神经网络是深度学习的其中一种重要形式，它是一种前馈神经网络，其中包括卷积计算，并且具有深度结构，可以执行监督学习和非监督学习，适合处理 RGB 图像、三维立体图像等可以呈现在 2D 平面的输入数据。其特点是可以通过多层网络的特征提取最终能够提取到输入数据的高维特征，且每个层次输出的特征图特点不同 (Deng Z, Sun H, Zhou S, 2018)。通常情况下浅层卷积层提取到的特征为纹理、颜色等特征；深层卷积层提取到的特征属于全局特征，比较抽象。

#### 2.3.2 卷积神经网络基本结构

卷积神经网络的基本结构主要包括输入层、隐藏层、输出层。其中隐藏层可以将输入数据加权求和并通过激活函数生成非线性特征输入到下一层，直到最后一层隐藏层的输出传递到输出层。隐藏层包括几个不同功能的层：卷积层、池化层以及全连接层。卷积神经网络输出的结果是每幅图像的特定特征空间。当处理图像分类任务时，我们会把卷积神经网络输出的特征空间作为全连接层或全连接神经网络的输入，可以用全连接层来完成从输入图像到标签集的映射，即分类。图 6 为卷积神经网络的结构模型。

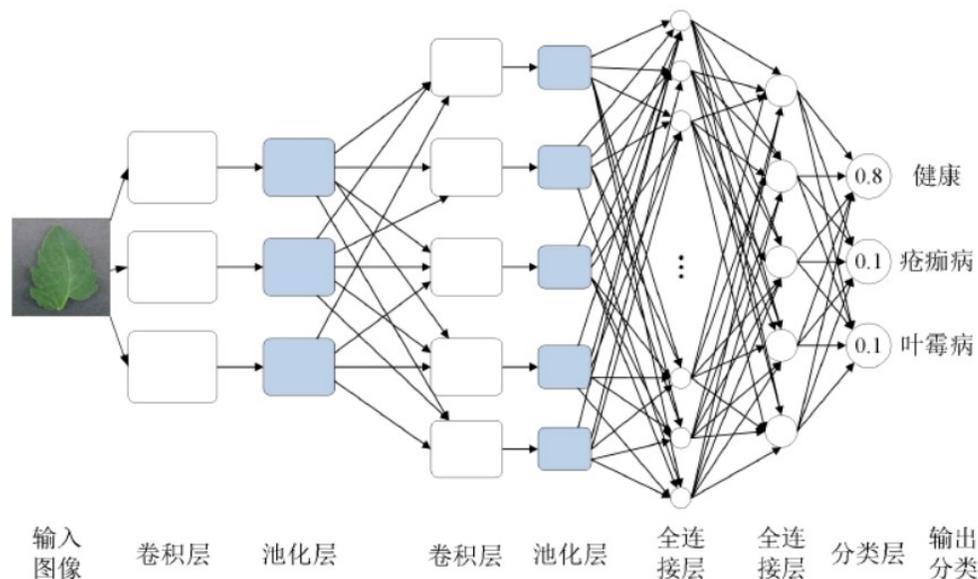


图 6 卷积神经网络的结构模型通过分析图 6，对卷积神经网络的结构模型进行简单分析：

### (1) 输入层

输入层是作为模型对外的接口用于数据的输入，可以处理多维数据，处理的图像可以是灰度图像，也可以是 RGB 图像，输入图像的像素值作为输入。

### (2) 卷积层

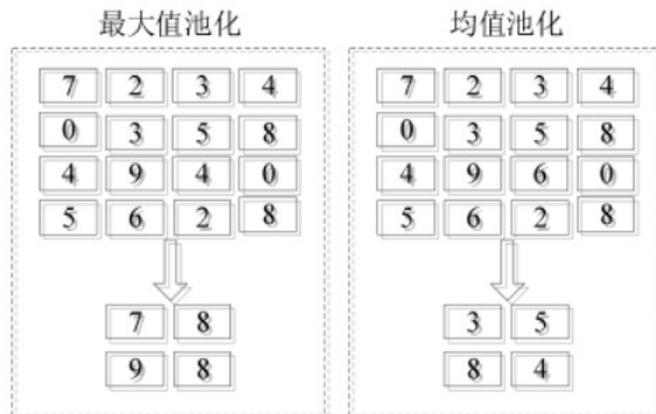
卷积层是卷积神经网络模型中最核心的部分，其主要作用是对输入数据进行特征提取，然后在后续网络中将其抽象为更高维度的特征，便于分类。不同层次的卷积层可以学习到不同特征，浅层卷积操作的输出为局部域特征，深层卷积操作的输出为更加抽象一些的特征，所提取的特征由浅到深，输出特征逐步抽象化。卷积层内部包含多个卷积核，卷积核的每个元素都对应一个权重系数和一个偏差量，类似于前馈神经网络的神经元。常用的卷积核尺寸为  $3 \times 3$  或者  $5 \times 5$ ，卷积核的参数是可以训练的，并且在训练过程中可以自动调整。卷积层的参数包括卷积核大小、填充和步长，这些都决定了卷积层输出特征图的大小。其中卷积核大小可以指定为比输入图像尺寸小的任意值，使用的卷积核越大，可提取的输入特征就越复杂，卷积计算方式如图 7 所示，以卷积核尺寸为  $3 \times 3$ ，步长为 1 滑动实例卷积计算方式，其输出是格式为  $[1, 3]$  的特征矩阵。



图 7 卷积计算方式

### (3) 池化层

在卷积层中提取特征后，输出的特征图将会传递到池化层用来进行信息过滤和特征选择。池化层包含预先设定的池化函数，其作用是减小特征图的尺寸，但不改变特征图的深度，从而减少后续网络结构的尺寸大小，可以缩短训练时间，使参数收敛更快，同时还能起到防止过拟合的作用。池化层的核尺寸越大，输出尺寸与输入尺寸的比值越小。池化方法通常分为均值池化和最大值池化，分别将池化视野中的元素取均值和取最大值。池化窗口以设定好的步幅沿着输入样本的空间维度进行滑动，直接计算池化窗口内的像素的平均值和最大值。如图 8 所示为池化窗口尺寸为 2x2，步长为 2 进行滑动的池化计算方式。



## 图 8 池化计算方式

### (4) 全连接层

卷积神经网络中的全连接层等价于传统前馈神经网络中的隐藏层，该层的全部神经元与上一层的全部输出彼此相连，再计算出每个类别对应的分数为输出层做准备。全连接层的功能是对提取的特征进行非线性组合以获得输出，全连接层本身是不具有特征提取功能，而是尝试使用现有的高阶特征完成学习目标。全连接层一般由一到两层组成，因为其参数量非常庞大，所以全连接层的层数越多，训练难度越大，越容易造成过拟合。如果数据由十个类别，那么全连接层最后一层就有十个神经元结构。

### (5) 输出层

卷积神经网络中输出层也称为分类层，其上方连接的通常是全连接层，所以其结构以及工作原理与前面介绍的传统前馈神经网络中的输出层相同。对于图像分类问题，输出层使用归一化指数函数或者逻辑函数输出分类标签。通常将最后一层计算出的每个类别的数值相加归一化，重新分配每个类别的概率，概率和为 1。

## 2.4 深度学习框架

在深度学习的发展初期，每个深度学习研究人员都需要编写很多重复的代码，为了提高工作效率，这些研究人员就将这些代码写成了一个框架，并将其放到网上供所有研究人员一起使用。接着，网上就出现了不同的框架。深度学习之所以能够快速流行起来，很大程度得益于这些算法完备的深度学习框架。在每一个深度学习项目开始之前，选择一个合适的框架是非常重要的，因为选择一个合适的框架可以提高效率，减少大量的工作量。

目前，全世界最为流行以及较受研究人员青睐的深度学习框架有 PaddlePaddle、Tensorflow、Caffe、Theano、MXNet、Keras 和 PyTorch。其中 Theano 可以说是深度学习的鼻祖，它是为处理大型神经网络算法所需的计算而专门设计，被认为是深度学习领域的行业标准；后来 Caffe 因为具备高效性，很快抢占了一部分市场，但其缺点是不够灵活，同时内存占用高；TensorFlow 是世界上拥有最多用户和最大社区的框架，由于该框架是 Google 公司出品，因此更新比较频繁，同时因为其具有 Python 和 C++ 的接口，使用起来更易于上手。TensorFlow 自带可视化功能，可以实时监测训练情况，具有高度灵活性和强大的可移植性。

2.5 经典卷积神经网络模型目前，研究人员常用的经典卷积神经网络有以下几种：

(1) AlexNet

AlexNet 诞生于 2012 年，是当年 ImageNet 竞赛冠军选手 Hinton 以及他的学生 Alex Krizhevsky 共同设计的。该网络由八层卷积神经网络组成，有大量的神经元和自由参数，是对 LeNet 结构的进一步改进和加深。AlexNet 首次将 Dropout 技术实用化，应用了 ReLU 激活函数，并且使用 CUDA 加速训练，利用 GPU 的强大的计算能力，相对于使用 CPU 大大提高了运算效率。AlexNet 的结构如图 9 所示。

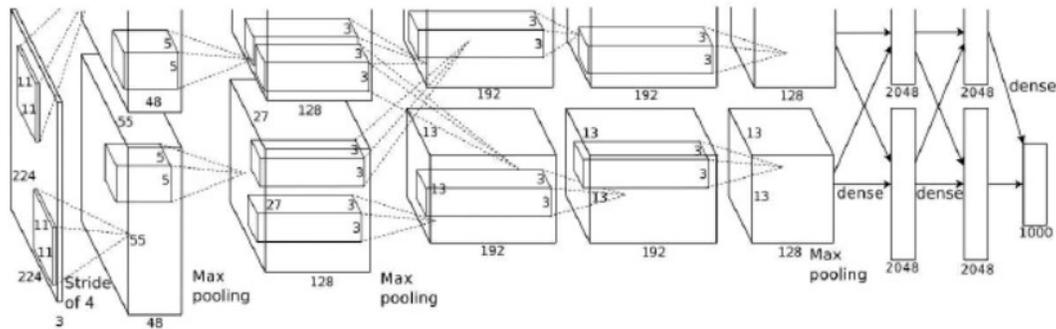
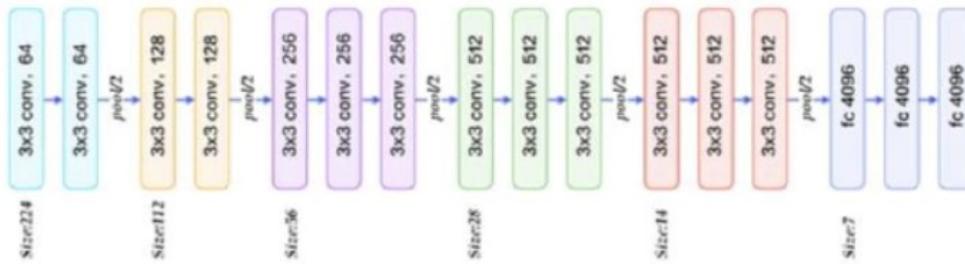


图 9 AlexNet 结构图

(2) VGGNet

VGGNet 是牛津大学视觉几何团队开发的一种卷积神经网络算法。该网络着重于研究卷积神经网络结构中其深度与性能之间的关系，使用 2x2 的最大池化与 3x3 的小卷积核进行反复堆叠，使其网络深度加深并且性能也得到大大提升。VGGNet 因为其简洁的结构，所以在特征提取以及迁移学习任务中应用广泛。VGG 的结构如图 10 所示。



以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/768051063111007003>