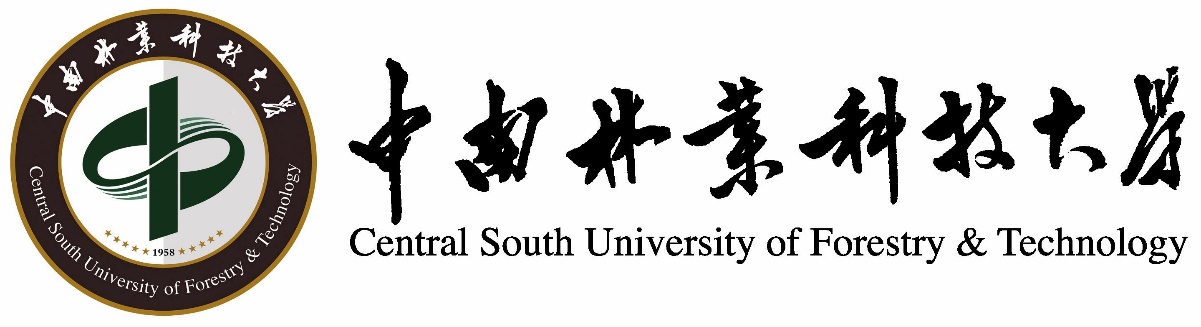
分类号： 密 级：

学校代号： 10538 学 号： 20202524





本科毕业（学士学位）设计

基于多特征融合和深度学习的植物叶片识别

学生姓名：  曾祥文

指导教师：邝祝芳教授

培养学院：计算机与数学学院

年级专业：2020级计算机科学与技术

提交日期：2024年6月1日

中南林业科技大学

毕业设计原创性声明

本人郑重声明：所呈交的设计是本人在导师的指导下独立进行研究所取得的研究成果。除了文中特别加以标注引用的内容外，本设计不包含任何其他个人或集体已经发表或撰写的成果作品。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律后果由本人承担。

学生签名： 日期：2024年 月 日

毕业设计版权使用授权书

本毕业设计作者完全了解学校有关保留、使用设计的规定，同意学校保留并向国家有关部门或机构送交设计的复印件和电子版，允许设计被查阅和借阅。本人授权中南林业科技大学可以将本设计的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本设计。

本设计属于

1、保 密 ，在 年解密后适用本授权书。

2、不保密 。

（请在以上相应方框内打“√”）

学生签名： 日期：20 年 月 日

导师签名： 日期：20 年 月 日

基于多特征融合和深度学习的植物叶片识别

摘 要

植物在农业、医疗、环保等行业是必不可少的。近年来，由于环境污染的不断加剧，每年都有大量植物物种死亡或灭绝。由于保护植物多样性的价值与日俱增，植物鉴定识别尤为重要。 随着计算机视觉技术的不断进步，研究已逐渐转向基于叶片图像的植物识别技术。

本文以深度学习算法为基础，对植物叶片分类识别进行研究。选用Flavia植物叶片数据集，包括32种植物叶片，分别使用GoogLeNet模型和ResNet模型来对植物叶片进行分类识别。实验过程中首先对植物叶片进行预处理操作，其中包含了图像压缩、图像增强和数据归一化等操作；使用GoogLeNet模型对测试集图像进行识别，平均准确率为89.06%，使用ResNet18模型对测试集图像识别的平均准确率为92.37%，相比于GoogLeNet模型提高了3.31%，通过改进ResNet18模型。使用５５卷积替代７７卷积，采用残差块的瓶颈结构代替捷径结构，改进之后识别的准确率达到了96.95%，相比于改进前的ResNet18模型提高了1.18%，最后使用ResNet34模型对测试集图像识别，平均准确率达到了96.95%，相比与ResNet18模型提升了4.58%。

关键词：植物叶片识别；深度学习；卷积神经网络；GoogLeNet；ResNet

Plant Leaf lmage Recognition Based on Multi-feature Integration and Deep Learning

**Abstract**

Plants are essential in agriculture, medical care, environmental protection and other industries. In recent years, due to the increasing environmental pollution, a large number of plant species die or become extinct every year. Because the value of protecting plant diversity is increasing day by day, plant identification is particularly important. With the continuous progress of computer vision technology, the research has gradually turned to plant identification technology based on leaf images.

Based on the deep learning algorithm, this paper studies the classification and recognition of plant leaves. Flavia plant leaf data set, including 32 kinds of plant leaves, is selected, and GoogLeNet model and ResNet model are used to classify and identify plant leaves respectively. During the experiment, the plant leaves were preprocessed, including image compression, image enhancement and data normalization. Using GoogLeNet model to identify the test set images, the average accuracy rate is 89.06%, and using ResNet18 model to identify the test set images is 92.37%, which is 3.31% higher than GoogLeNet model. Using 5× 5 convolution instead of 7× 7 convolution, and using the bottleneck structure of residual block instead of shortcut structure, the improved recognition accuracy reaches 96.95%, which is 1.18% higher than that of the improved ResNet18 model. Finally, the average accuracy of the test set image recognition using ResNet34 model reaches 96.95%, which is 4.58% higher than that of ResNet18 model.

**Keywords:** Plant leaf identification;Deep learning;CNN;GoogLeNet;ResNet

目 录

[目 录 IV](#_Toc167810232)

[1 引言 1](#_Toc167810233)

[1.1 研究背景和意义 1](#_Toc167810234)

[1.2 国内外研究现状 3](#_Toc167810235)

[1.3论文结构 6](#_Toc167810236)

[2 相关基础知识 7](#_Toc167810237)

[2.1 深度学习算法 7](#_Toc167810238)

[2.2 深度学习框架 14](#_Toc167810239)

[2.3 植物叶片识别常用网络模型 16](#_Toc167810240)

[2.4 总结 22](#_Toc167810241)

[3 基于GoogLeNet网络模型的植物叶片识别 23](#_Toc167810242)

[3.1 数据集及预处理 23](#_Toc167810243)

[3.2模型设计 24](#_Toc167810244)

[3.3 实验过程 25](#_Toc167810245)

[3.4 实验结果与分析 28](#_Toc167810246)

[4 基于ResNet网络模型的植物叶片识别 31](#_Toc167810247)

[4.1模型设计 31](#_Toc167810248)

[4.2 实验过程 34](#_Toc167810249)

[4.3 实验结果与分析 36](#_Toc167810250)

[4.4 总结 40](#_Toc167810251)

[结论 41](#_Toc167810252)

[参 考 文 献 42](#_Toc167810253)

[致 谢 45](#_Toc167810254)

**1 引言**

* 1. **研究背景和意义**

1.1.1 研究背景

植物在食品、医学、工业和日常生活中发挥着重要作用，如提供食物，保持大气中的氧平衡。然而，随着人类的生产活动的不断增加和城市化进程的加速推进，过度的开发、全球气候的变暖、环境的恶化以及污染等因素对植物的生态环境造成了严重破坏。许多植物物种每年都在消失，如果再不采取行动，人类社会将会带来非常严重的后果。全球生物多样性正在迅速减少[24]。根据Knapp等人[25]的调查，已知的植物物种大约有25万种多，其中将近13%濒临灭绝。植物物种的大规模消失将给人类社会和生态系统带来严重的负面后果，比如将会导致土地的荒漠化、洪水频发以及极端气候事件增多。同时，由于不同物种之间相互依赖，一种植物的消失可能将导致其他类别物种的灭绝。越来越多的证据表明，生物多样性的减少会对生态系统带来不利的影响。因此，保护植物物种就会显得尤其迫切和重要[26]。

对于未来的植物物种多样性的保护，准确掌握植物的地理分布信息和特性至关重要[27]。识别和了解植物物种的特征和分布是植物物种保护的主要目标，农业信息化和生态保护需要植物分类研究，然后植物物种分类是一项具有挑战性的任务。因为地球上存在的27万多种植物大部分还没有得到充分的研究，使用传统的人工方法识别植物物种困难、费时、容易出错，外行也很难理解。由于需要使用特定的植物学词汇，几乎不可能手动识别大多数已知的植物种类。

植物物种的自动化识别系统对植物分类学的研究者以及广大群众来说都具有非常重要意义[28]。随着计算机技术的飞速发展，例如更好的软硬件、图像处理、便携式设备的普及和机器学习的快速进步等，植物种类的自动识别成为了可能。

最近几年，越来越多的研究人员和团队开始专注于植物物种自动识别算法的研究，其中包括了计算机视觉等相关领域的专家和学者。植物物种的分类和识别可以基于植物各种器官的特征，如花、茎、叶、果实、根以及种子等。植物叶片形状和结构比较稳定，通常平坦且接近平面。因为植物叶片具有寿命长、特征鲜明、易得等优点，专家以及研究人员更加倾向去选择植物叶片图像来进行识别分类的研究。



图1-1植物的六大器官

植物物种种类的识别可以通过叶片的形状、纹理和颜色等特征。尽管叶片一般都呈现出绿色，但受到自然环境的影响，所以颜色特征并不常用于识别。不同种类的植物叶片在形状上常常有着明显的差异，因此形状特征成为识别植物种类的重要依据。叶片的纹理特征主要包括叶脉及其内部的结构信息。通常情况下，同一种植物的叶脉结构相似，而不同种植物的叶脉结构则存在着较大差异，这使得纹理这项特征成为植物识别的非常重要依据之一。

传统的机器学习方法通常利用植物识别算法中的形状和纹理等特征。然而，这些方法通常需要进行大量的预处理，且所设计的算法通常只在特定或某类数据集上表现良好。当有足够多的叶片图像样本时，深度学习可以自动学习到高度识别的叶片特征，从而实现对植物叶片图像的有效识别。因此，基于卷积神经网络（Convolutional Neural Networks, CNN）[29]和叶片图像的植物物种识别算法正在逐渐兴起。系统通过卷积神经网络，可以自动提取和学习叶片的复杂特征，这将会显著提高识别的精确性和效率。相较于传统方法，基于CNN的深度学习算法减少了对人工特征设计的过度依赖，所以能够在更广泛的数据集上表现出色。随着大量叶片图像数据的积累和深度学习技术的不断发展，基于CNN的植物识别算法在研究和应用中正变得越来越重要。

**1.2 国内外研究现状**

科技的不断进步促使不同学科之间建立了密切联系，形成了一种融合的趋势。在农业领域计算机图像识别技术被广泛的应用，并且取得了迅速的发展。图像识别技术主要用于植物叶片分类识别，通过对植物叶片图像进行特征提取和种类的判别来实现的，以解决人眼识别效率低下的问题。近些年来，国内外的专家和学者以及相关领域对植物叶片识别进行了深入的研究，这已经成为了一个研究的热点了。针对植物叶片识别研究在深度学习的算法帮助下也不断涌现出新的思路和方法。

1.2.1 植物叶片识别领域的研究现状

图像识别技术正不断地日益成熟，并且在不断扩大其应用的范围。早在1986年，国外就已经对植物叶片识别进行了一系列的研究，展开了深入的探讨。通过采用多特征融合的方法，Imgrouile提取了橡树叶片的特征，并运用降维算法对其进行分类[13]。Osika和他的团队采用BP前馈神经网络算法，将面积参数和叶片形态特征作为识别的标志进行了分类[31]。Rashad及其团队从植物图像数据库中挑选了10种不同类型的植物叶片，对纹理特征进行提取并通过模型进行训练，因此取得了非常不错的成果 [32]。通过利用傅里叶变换将频域信号作为特征依据，Lee提出了一种新型的植物叶片分类方法，他通过提取叶脉特征的方式来提高了识别的准确率[33]。Mehdipour Ghazi和他的团队通过利用PCANet算法对叶片进行识别分类。值得注意的是，他们发现在单一背景叶片的情况下，识别率远远高于在复杂背景的叶片，达到了90% [34]。DivyaTomar和她的团队建立了一个植物种类叶片识别系统，采用了基于双梯度最小二乘支持向量机（DAG-MLSTSVM）的分类器。他们的研究显示，该系统的识别准确率为84.70%[35]。Krizhevsky和她的团队以ImageNet数据集作为训练对象，通过构建AlexNet这一经典卷积神经网络，测试的失误率为15.4%[36]。SoonJyeKho及其团队利用图像处理和模式识别技术，结合提取叶片形态特征，采用了人工神经网络和SVM来训练分类模型。他们成功开发了三种榕属植物的自动识别系统，其识别准确率达到了83.3%[38]。EMImah提出了一种方法利用全特征属性可以自动分类15种植物的方法，该方法基于竞争学习算法GRLVO，通过实验表明，这种方法的识别准确率可达92.98%[38]。

在国内，植物叶片的识别技术相对来说起步比较晚。2001年，傅弘及其团队为了提高叶片识别的准确率，采用了叶片脉络纹理作为特征，并运用了边缘梯度等统计特征方法进行分类识别。经过实验的训练，识别的准确度得到了显著提高，这为之后的植物叶片识别奠定了坚实的基础[1]。对于植物叶片识别的研究，杜吉祥提出了一种方法，即移动中值重心超球分类器（MMC），对选取的15种叶片特征进行了提取和分类[2]。王晓峰及其团队采用了传统的识别方法，通过形状参数和叶片图像特征进行分类识别，最终取得了非常不错的识别率[3]。朱宁利用了LBP特征来提取植物叶片图像的纹理特征，然后又结合了数字图像处理技术和模式识别技术，最终设计并实现了一套树叶识别系统[4]。贺鹏针对阔叶树识别研究，首次使用概率神经网络(PNN)及其分类器，通过实验表明，准确率达到了98.3%[5]。阚江明和他的团队提出了很多中对于植物叶片的处理方法，通过对植物叶片进行预处理之后再进行训练，平均识别准确率为70.83%[6]。张宁结合了图像分析技术，对植物叶片识别相关理论方法进行了深入研究，并详细地介绍了植物识别技术的现状和发展趋势[7]。王丽君采用了支持向量机（SVM）分类器，提取了植物的26个特征，最终将观叶叶片的识别准确率提升至91.41%[8]。杨天天和其团队对几种柳属（Salix）植物的单一背景叶片进行了分析，综合考虑了多种叶片特征信息，并通过分类提取的方法，显著提高了识别效果[9]。于慧伶等人利用双路卷积神经网络结构对Flavia公共叶片数据集进行分类，提高了植物叶片的识别准确率[10]。朱良宽和其团队提出了一种融合了DCGAN与迁移学习的识别方法，他们运用了CVL研究领域中的瑞典植物叶片集进行了实验，实验取得了良好的效果[11]。李龙龙等人采用LBP算法度量和提取图像的局部纹理信息，这一方法有效地提高了训练速度[12]。

目前，传统的图像识别技术是对植物叶片进行识别研究的主要方法。这些方法通常涉及将提取到的植物叶片特征输入机器学习模型进行训练，已经取得了令人满意的效果。然而，在图像识别中，常常会使用人工设定的特征进行提取，而所采用的数据集也通常是特定的。若在后续实验中更换数据集，可能会导致结果出现较大差异。相比之下，深度学习通过利用其独特的深度网络结构，自动学习数据的特征信息，从而加快了识别速度，并且更好地适用于各个领域。因此，随着机器学习领域的进步，越来越多的专家学者开始广泛地将深度学习运用于图像识别研究中。

1.2.2 深度学习在图像识别领域的研究现状

人工智能的崛起是信息时代的亮点之一，并且机器学习又是人工智能的一个重要组成部分，使用深度学习算法更加的显著。传统机器学习算法中，植物叶片的识别方法相对来说比较固定，一般是通过提取植物叶片的纹理、形状和脉络等特征，然后通过分类器对这些特征进行识别。然而，传统的算法对植物叶片数据集要求比较高，不同的特征提取方法可能会影响识别的效果，导致模型的泛化能力较差，识别的准确率就会较低。随着时间的推移和硬件技术的进步，深度学习得到了广泛发展，并且广泛的应用在图像识别领域。其高效的学习能力和独特的架构模式引起了国内外的广泛关注，受到了世界各地专家的重视。

在国外，2006年，深度学习迎来了一股热潮。Hinton等人构建了深度信念网络(DBN)模型，采用无监督训练方法，通过方向传播算法来优化模型的参数[39]。随后，2012年，杰弗里等人提出了卷积神经网络模型AlexNet，这个模型具有多个隐藏层的。哥伦比亚大学、马里兰大学和史密森尼学会联合发起了Leafsnap项目，采用了美国的东北部将近185种的植物叶片，开发了一款植物叶片识别的APP，APP的使用截图如图1-2所示。Jeon等人针对Flavia数据库中的32总植物叶片通过构建了卷积神经网络模型GoogLeNet模型来提取特征信息进行分类训练，并通过调整超参数来提高其识别率[40]。在PlantCLEF2016比赛中，Ghazi等人发现使用改进后的AlexNet、GoogLeNet和VggNet模型用来识别植物种类的算法明显优于其他的方法[41]。

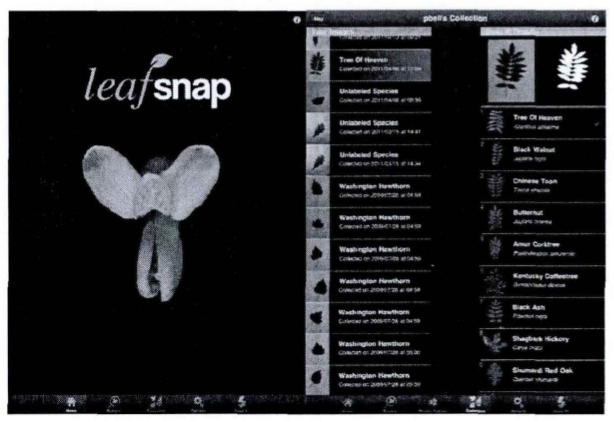


图1-2 ipad版本Leafsnap的截屏

在国内，许多的研究者将深度学习方法应用到植物叶片的识别领域。尽管深度学习在国内起步比较晚，但是其发展迅速，并且广泛应用在农业领域，例如农作物自动化采摘、病虫害识别和植物叶片识别等方面。朱静等人通过对对象进行预处理，然后提取叶片图像特征，成功地将植物叶片识别准确率提高到了92%[14]。龚丁禧等研究人员利用深度学习理论，通过构建剪辑神经网络模型，他们对瑞典叶片数据集中的15种叶片进行了训练，不断调整参数，结果比传统的叶片识别算法更为优异[15]。丰晓霞及其团队引入了深度学习中常用的模型分类方法，他们将提取到的特征用SVM代替了softmax分类器进行分类[16]。张帅将PlantNet叶片数据集以及自己扩展的叶片作为训练样本集，采用分割算法分别对单一和复杂环境的叶片进行处理，并结合CNN与分类器对植物叶片进行训练，实验表明，SVM分类器的识别率要高于softmax分类器，在单一背景取得了相对较好的分类效果，但是在复杂背景下识别率较低[17]。薄琪苇利用深度学习算法开发了植物叶片识别系统，并将其应用于Android端。该系统不仅对15种叶片图像获得了良好的识别率，还在识别模式方面进行了创新[18]。为了提高储粮害虫图像的识别率，程曦等人构建了GoogLeNet和AlexNet卷积神经网络模型进行训练，实验结果显示，他们获得了良好的分类效果[19]。2019年，黄志国以中草药叶片为识别对象，改进LeNet-5模型，经过对比实验分析，改进后的模型减少训练时间的同时，还提高了识别的准确率[20]。王艳等人通过利用深度学习相关的算法对中草药进行研究，取得了良好的成果[21]。孙颖异等研究人员提出了一种新的叶片识别方法，在卷积神经网络AlexNet模型的基础上进行改进，引入了残差连接，他们对PlantVillage数据集中的12种植物叶片数据集进行了识别实验，实验证明带有残差连接的AlexNet卷积神经网络能够高效识别植物叶片[22]。樊湘鹏及其团队使用了复杂背景下的玉米病害图像，他们设计了一个包含5层卷积、4层池化和2个全连接层的卷积神经网络结构，并开发了一个玉米田间病害识别系统, 实验结果表明，该系统的平均识别准确率达到了83.33%[23]。

综上所述，本研究计划采用多特征融合和深度学习的方法，旨在解决传统方法在植物叶片识别中的局限性。通过综合利用形状、纹理和颜色等多个特征，并通过深度学习进行训练和分类，本研究旨在提高植物叶片识别的准确性和鲁棒性，从而推动植物叶片识别技术的发展，提高植物资源的利用效率，保护生态环境。

1.3论文结构

本论文是基于多特征融合和深度学习的植物叶片识别，文章的组织结构如下：

第一章，引言部分。本章首先介绍的是植物叶片识别的研究背景和意义；然后，对国内外在该领域的研究历程进行了梳理，罗列了一些具有代表性的成果。

第二章，植物叶片识别和深度学习的相关基础知识部分，本章节首先介绍的是深度学习的基础知识，详细阐述了CNN 的各个组成部分。然后介绍了深度学习的两种框架Pytorch和TensorFlow；最后，介绍了常用的植物叶片识别的网络模型，包括AlexNet，VGG，GoogLeNet和ResNet模型。

第三章，详细介绍了GoogLeNet的网络模型结构并且使用GoogLeNet模型对植物叶片进行识别实验。

第四章，首先使用ResNet18模型对植物叶片进行识别，为了提升植物叶片识别的准确率对ResNet18模型进行改进；同时也采用了ResNet34网络模型和ResNet18模型做对比实验；最后通过改变ResNet34模型的实验参数Batch\_Size来获得识别率最高的Batch\_Size。

**2 相关基础知识**

**2.1 深度学习算法**

2.1.1 深度学习相关概述

随着人工智能的快速演进，机器学习已经成为了实现人工智能的重要手段之一。机器学习的核心目标是从数据中捕获模式，通过对数据的分析来推断出预测模型，以实现对现实事件的识别和决策。在图像识别领域，机器学习有着广泛的应用，其基本原理是通过获取图像数据并提取特征，然后利用分类器对历史数据进行学习，以实现对真实场景的准确识别。

然而，传统的机器学习方法通常需要人工设计特征，如尺度不变特征（SIFT）、DOG金字塔（Difference of Gaussian）和全局特征（HOG）等，这给特征选择和分析带来了一定的挑战。近年来，为了解决这个问题，出现了一种新的研究方向——深度学习。深度学习利用多层神经网络自动提取特征，从而降低了特征提取的时间和工程投入。

深度学习的历史十分悠久。自从2006年以来，深度学习在机器学习这项领域取得了重大进展，尤其是在图像的识别方面。2012年，Hinton教授及其学生提出了AlexNet网络模型，在ImageNet大规模视觉识别竞赛中获得了冠军，这标志着深度学习在图像识别方面取得了成功。随后，VGG模型和ResNet模型相继问世，进一步推动了深度学习在图像识别领域的发展。

相对于传统的神经网络，深度学习算法通过增加网络的层数来自动学习主要特征，从而减少了特征提取的复杂性。深度学习网络结构包括多个隐藏层，每一层的节点都与相邻层的节点连接，并且具有不同的激活函数，这有助于提高模型的泛化能力。深度学习算法的设计理念受到了人脑分层结构的启发，使得网络结构更贴近人脑神经系统的运作方式。

深度学习的进步将进一步助推人工智能技术的发展，为解决各类现实世界问题提供了全新的思维和方法。

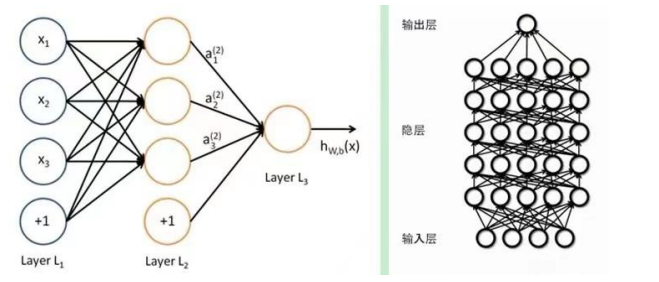


图2-1单隐层和多隐层深度学习模型对比图

深度学习的核心理念在于利用深度网络模型作为主要工具，通过构建特征来自动训练大规模数据，从而不断提升模型在识别和决策方面的准确性。深度学习算法的主要优势集中体现在以下的两个方面：首先是网络结构的“深度化”，即模型包含多个隐层和节点，能够逐层提取数据信息，使用较少的参数表示复杂的函数，因此分类结果更为准确；其次是特征提取的“关键性”，即模型能够自动学习特征规律，丰富数据的内在信息，从而提高了模型的识别能力。

深度学习在人工智能领域已被证实是最先进的技术之一，其发展势头迅猛。无论是模式识别技术、分类任务，还是个性化推荐系统、购物助手，抑或是医疗预防诊断和疾病预测等，深度学习在各领域展现了强大的应用潜力，为我们的生活带来了惊喜和便利。

2.1.2 深度学习训练过程

在对于传统网络的训练方法中，一般都会采用梯度下降法。这一方法通过设定初始参数，计算网络输出与标签之间的差异，然后不断调整参数直至达到收敛。然而，这种训练方法存在一些不足之处。例如，BP神经网络在训练过程中自上而下的顺序可能导致梯度下降变得稀疏，从而引起较大的误差，而且对于没有标签的数据样本也存在一定的局限性。与之不同，深度学习网络能够在训练过程中纠正这些不足之处。

在图像识别领域，深度学习网络的训练过程通常包括以下五个步骤：

数据预处理：对训练数据进行标记和归一化处理，旨在提高模型的识别性能。

正向传播：采用有监督学习方式，逐层训练神经网络，将标定好的数据输入网络中，调整每个神经元的权重，并使用激活函数计算中间变量。

计算误差：降低损失函数的值，以防止模型过度拟合，通过准确值与实际识别率之间的误差来评估识别效果。

1. 反向传播：采用无监督学习方式，对有标签或无标签的数据进行训练，通过反向传播算法迭代地计算误差函数的梯度，从而调整参数值以使误差函数最小化。
2. SGD算法：通过梯度向量调整每个权值，使误差逐步趋于最小值。对于非凸损失函数，调整收敛趋势以获得更好的训练效果。

2.1.3 卷积神经网络

卷积神经网络（CNN）是一种前馈神经网络，是深度学习领域中备受关注的强大工具。CNN能够自动学习相关特征，无需人工监督，这是其主要优势之一。目前，CNN在多个领域都得到了广泛应用，尤其是在计算机视觉领域。

卷积神经网络（CNN）是一种逐层连接的多层网络结构，每一层都包含多个特征图和独立的神经元。通常，CNN由三个主要部分组成：输入层负责接收原始数据信息，中间包含多个卷积层和池化层，主要用于对输入图像进行卷积操作以提取有用信息和学习特征，最后是全连接层，将提取的特征输入分类器进行分类，然后由输出层呈现结果。在卷积层中，滤波器用于提取输入图像的特征，而池化层则对输入图像进行降维操作，以防止模型过度拟合。如图2-2所示，是卷积神经网络的示意图

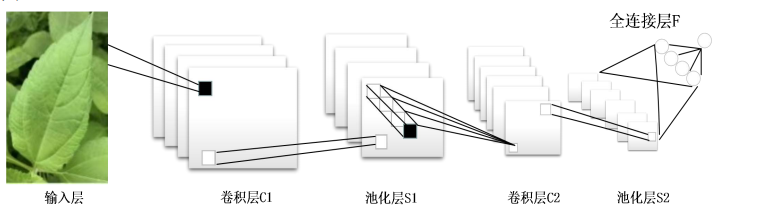


图2-2卷积神经网络结构示意图

(1)卷积层

卷积层是卷积神经网络（CNN）中最关键的部分之一，用于从输入图像中提取特征。卷积层的主要功能是通过在输入图像上滑动固定大小的滤波器或卷积核来执行卷积操作，从而生成一个新的特征图。在这个过程中，滤波器中的权重与输入图像的一部分进行乘积运算，然后将结果相加，得到特征图中对应像素的值。图2-3展示输入大小为5\*5，卷积核大小为3\*3的卷积运算。

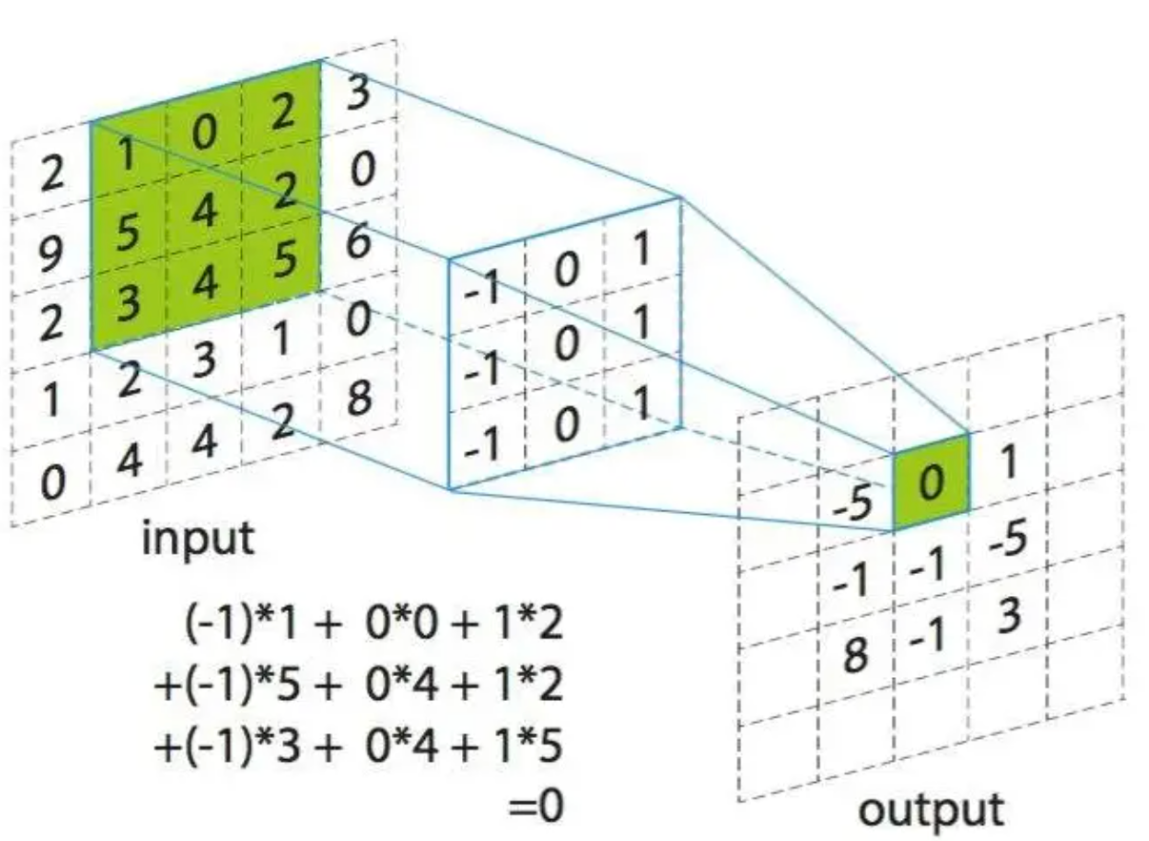


图2-3卷积运算示意图

卷积层通常包含一个或多个卷积核，每个卷积核能够捕捉输入图像中的不同特征。举例来说，某些卷积核可能捕捉边缘、线条或纹理等低级特征，而其他卷积核则可能捕捉更高层次的语义特征，比如形状、物体部分或场景。

卷积层具有一些重要的超参数，包括卷积核的大小、步长和填充。卷积核的大小决定了它在输入图像中捕获特征的空间范围。步长决定了每个滤波器在输入图像上滑动的距离，从而影响输出特征图的空间分辨率。填充是在输入图像的边缘周围添加一些像素，以控制输出特征图的大小和形状。

卷积层的输出通过使用非线性的激活函数来增强网络模型的非线性能力。其中，ReLU函数是常用的激活函数之一。在训练过程中，通过反向传播算法来更新卷积层中的权重，以最小化损失函数并提高模型的准确性。

在CNN中，卷积层具有权重共享的优势，这意味着相邻层的神经元之间没有独立的权重分配。相反，整个卷积层使用一组权重与输入矩阵的所有像素一起进行运算。这种共享权重的机制大大减少了训练过程中所需的参数数量，降低了模型的复杂度，并显著减少了训练时间和计算成本。

(2)池化层

池化层（Pooling layer）是卷积神经网络（CNN）中常见的一种层次。通常在卷积层之后使用，用于减小特征图的尺寸，从而减少参数数量并降低过拟合的风险。在池化过程中，每个池化步骤都会保留主要信息或特征。

常见的池化操作一般主要包括平均池化（Average Pooling）、最大池化（Max Pooling）和全局平均池化（Global Average Pooling，GAP）。最大池化通过在一个池化窗口内选择窗口中的最大值作为输出，以突出显著特征，有助于保留图像中的重要信息。平均池化则是在一个池化窗口内计算窗口中像素值的平均值，并将其作为输出值，这有助于平滑图像并减少对细节的过度关注，从而提高模型的鲁棒性。而全局平均池化（Global Average Pooling，GAP）是一种池化层的形式，它对输入的每个特征图的每个通道的元素进行平均，得到一个标量输出。利用图2-4来说明这三种池化操作。

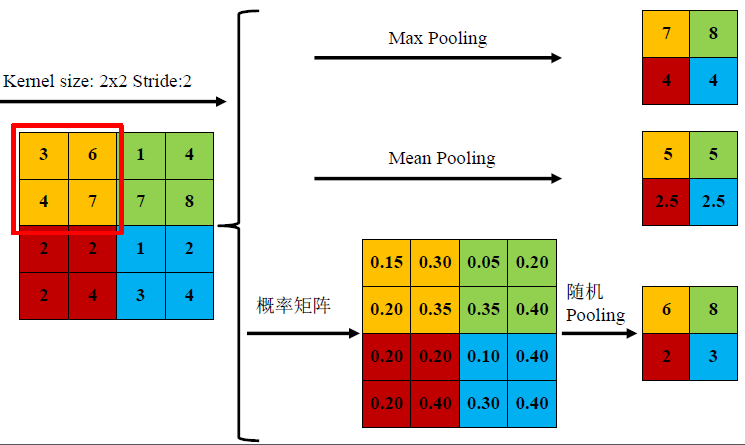


图2-4三种池化操作

在上图2-4中，输入特征图为4×4的矩阵，步长为2，池化窗口大小为2×2。因此，经过最大池化和平均池化后，输出的特征图为2×2的矩阵，而全局平均池化的输出为一个值。

(3)激活函数

激活函数在神经网络模型中扮演着非常关键的角色，它能够将输入数据进行非线性转换并输出。在卷积神经网络（CNN）的架构中，非线性激活函数通常被应用在可学习的层（例如卷积层）之后。这些激活函数的非线性特性使得输入与输出之间的关系变得非线性，从而使得CNN能够学习到更加复杂的特征。此外，激活函数需要具备可微性，这对于使用误差反向传播算法来训练网络是至关重要的。以下是CNN中最常用的几种激活函数：

Sigmoid：如图2-5，输入是实数，而输出范围在0到1之间。sigmoid 函数曲线是 S 形的，表达式如下：

Tanh：如图2-6，它类似于 sigmoid 函数，输入是实数，但输出范围在-1和1之间，表达式如下：

ReLU：如图2-7，是一种简单且使用频率较高的激活函数。当输入值大于零时直接返回输入值，而在输入值小于等于零时返回零。表达式如下：

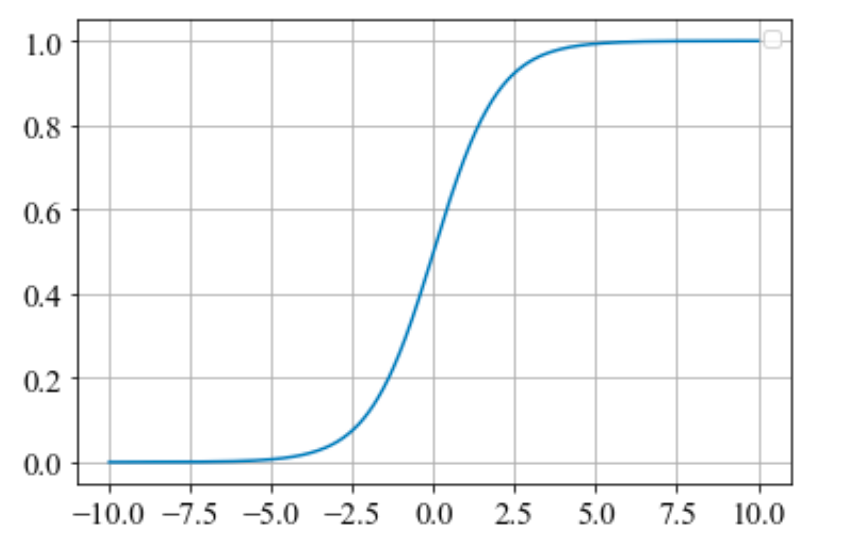


图2-5 Sigmoid 函数图

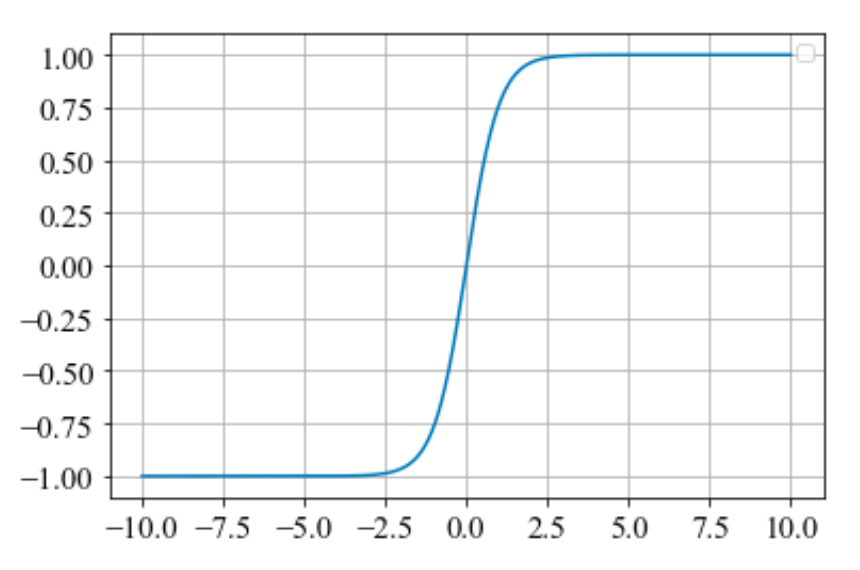


图2-6 Tanh函数图

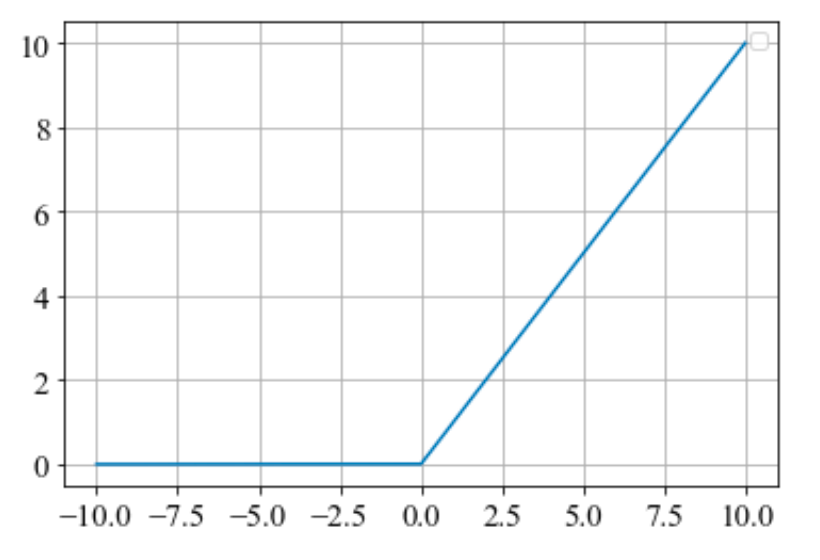


图2-7 ReLU函数图

(4)全连接层

全连接层是卷积神经网络（CNN）中特别常用的一种神经网络层，通常设计在CNN网络的末端。如图图2-8所示，在全连接层中，每个神经元都与前一层的所有神经元相连接，这称为全连接（FC）方法。每个神经元在全连接层中都与前一层的所有神经元相连接，这样的话就能够考虑前一层所有特征的组合。使得全连接层能够对提取到的特征进行更高级别的抽象，从而更好地区分不同的类别。然而，由于全连接层的参数数量非常大，因此会增加模型的计算和存储开销。

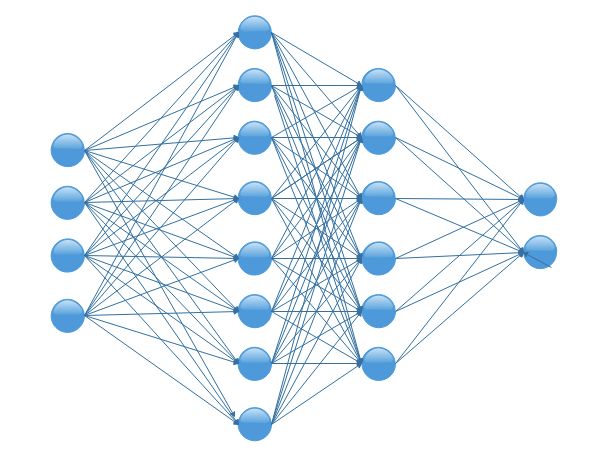


图2-8 全连接层示意图

(5)损失函数

在深度学习中，损失函数（Loss function）用于衡量模型预测值与真实值之间的差异。CNN中的损失函数通常用于指导模型的训练，并在训练过程中进行优化。目前来说被广泛使用了的交叉熵损失函数可以用以下公式表示：

其中，N表示样本数，k表示总类别数，表示类别为m对应的真实标签，表示模型对类别m的预测概率。如果损失函数的损失值越小，那么就表明模型越好

2.2 深度学习框架

在深度学习领域，选择合适的框架至关重要。框架的选择不仅影响着使用者的体验，也直接影响着开发者的效率。因此，一个优秀的框架不仅对使用者友好，同时也对开发者十分友好。为了实现研究目标，研究人员们开发了各种各样的框架来满足不同的需求。目前，深度学习领域有许多框架可供选择，其中包括PyTorch、TensorFlow、Theano等。在本节中，我们将重点介绍两个重要的深度学习框架：PyTorch和TensorFlow。

2.2.1 Pytorch

PyTorch和TensorFlow、Caffe、CNTK等框架一样，用于构建深度神经网络，如多层感知机、卷积神经网络等。这些框架就像建造房子时的地基一样重要，如果地基不牢固，整个房子都会不稳定。因此，在深度学习算法领域，选择一个优秀框架进行实验显得至关的重要。在特定任务中选择适合的框架可以促进任务的顺利完成。PyTorch不仅支持CPU计算，还支持GPU计算，利用GPU可以实现高速运算。

PyTorch有许多优点，其中之一就是其设计追求最少的封装，这体现在其简洁性上。这种设计让使用者能够更轻松地阅读和理解代码。换句话说，相比于TensorFlow，PyTorch在实现相同功能时，代码量只有十分之一左右。PyTorch的动态计算图也非常简洁，当需要构建简单的深度神经网络时，只需几行代码就可以轻松实现复杂的神经网络结构。例如，要搭建一个多层感知机，只需调用torch.nn.Linear即可实现一层神经网络。如果需要构建图神经网络，可能需要使用卷积层和池化层，而相应的代码也非常简洁，可直接调用torch.nn.Conv2d和torch.nn.MaxPool。

PyTorch不像TensorFlow那样需要重新学习复杂的计算规则和语法结构，它更像是Python的扩展，因此人们只需要花费更少的时间就能很快适应。它具有简单的调试机制和一些通俗易懂的API，这对于初学者来说非常有帮助。此外，PyTorch还能对输入的数据进行并行化处理，这种并行化处理使得数据的执行效率比串行化更高。它允许将数据分成多个批次，然后将这些批次同时发送到多个图形处理单元进行并行运算。通过这样的操作，PyTorch逐渐将原本由CPU执行的大量任务转移到GPU，从而大大减轻了CPU的使用压力。

2.2.2 TensorFlow

在深度学习领域，TensorFlow同样也是最常用的框架之一。它备受学生和研究者的欢迎并非没有原因，因为它是专门为方便人们使用而设计的框架。TensorFlow不仅是一种深度学习框架，还是一个端到端的学习平台。在这个平台上，人们可以创建用于机器学习或深度学习的库或软件。它也可以被理解为一个库，本质上是一个包含各种符号和定义的数学库。通过这个库中的各种符号和定义，人们可以构建和计算深度学习领域中各种复杂的任务。

TensorFlow是由谷歌的研究与开发人员构建的。在TensorFlow问世之前，深度学习凭借其出色的求解和计算性能在很长一段时间内遥遥领先于其他传统机器学习算法。当时，谷歌的研发人员也意识到了深度学习的优势，他们希望将这种优势应用于改进谷歌自身的软件功能，例如谷歌搜索引擎、谷歌地图和谷歌邮箱等服务。为此，他们开发了一种名为TensorFlow的框架。这个框架不仅适用于谷歌公司自身的服务和应用程序，也可以被其他开发人员使用。由于其开源性，TensorFlow得到了快速的发展，现在已经成为一个规模越来越庞大的项目。

TensorFlow的优势之一是其强大的社区功能。由于其早期出现以及代码的开源性，TensorFlow的社区不断壮大。其库中包含许多简单易用的接口供开发人员选择，如循环神经网络（RNN）、长短期记忆网络（LSTM）和残差网络等。然而，TensorFlow仍然存在一些缺点，主要体现在代码方面。由于其代码相对底层，用户需要编写大量代码，这可能导致代码的重复性，对于初学者来说并不友好。

2.3 植物叶片识别常用网络模型

2.3.1 AlexNet模型

AlexNet[43]是CNN中的经典网络模型，它包含8个深层网络，其中包括5个卷积层和3个全连接层。在卷积层中，引入了ReLU（Rectified Linear Unit）激活函数，以加快模型的训练速度，并有效地解决了梯度消失的问题。此外，还利用局部响应归一化（LRN）技术，对卷积层输出进行归一化处理，以防止较小的神经元输出对模型的干扰，从而提高了模型的拟合效果。在全连接层之后，引入了Dropout技术，以避免网络过拟合的问题。总体而言，AlexNet在图像分类和识别领域具有显著的性能优势。AlexNet模型的网络示意图如图2-9 AlexNet卷积神经网络架构示意图所示。表2-1为 Alexnet 模型的网络参数

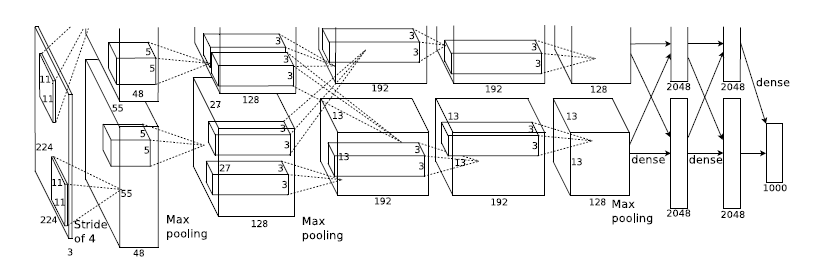


图2-9 AlexNet卷积神经网络架构示意图

表2-1 Alexnet 模型的网络参数



在AlexNet的第一个卷积层中，输入尺寸大小为224×224×3的RGB彩色图片，其中有3个通道。接着，使用96个尺寸为11×11×3的卷积核，设置步长为4，对输入的图像进行卷积操作。接着，将第一个卷积层的输出用来作为第二个卷积层的输入。在第二个卷积层中，采用256个尺寸为5×5×48的卷积核，步长设置为1，通过池化操作来降低参数量。接下来，第三个卷积层包含384个尺寸为3×3×256的卷积核，然后是384个尺寸为3×3×192的卷积核，第五层使用256个尺寸为3×3×192的卷积核，并进行池化降维操作。

在接下来的三个全连接层中，第六层、第七层和第八层分别包含4096个通道，将之前学习到的图像特征相融合。最后一层设置为1000输出，并通过softmax函数进行分类识别。这一结构能够有效地提取图像特征，并在最后的全连接层中实现分类。

2.3.2 VGG模型

VggNet[44]的提出起源于 2014年，这一网络模型主要通过加深网络结构以提升性能。VggNet的架构包括Vgg-11、Vgg-13、Vgg-16和Vgg-19，它们都以多个3x3的小型卷积核和2x2的max pooling为特征。其中，Vgg-16和Vgg-19因其简洁结构和出色分类性能，在深度学习领域广受欢迎。在训练过程中，VggNet要求输入图像的尺寸为224x224，并进行批量均值处理。它采用连续的3x3卷积核和13个卷积层来提取图像特征信息，最后利用3个全连接层和softmax分类器对图像进行分类，这样就可以最大程度地提取图像的特征了。

相较于AlexNet模型，Vgg-16在模型架构方面做出了重要改进，主要体现在卷积核的选取上。Vgg-16采用了多个3x3的轻量级卷积核，而不是像AlexNet那样使用较大的11x11、7x7、5x5的卷积核，这一改进有助于降低分类错误率。采用小卷积核的优势在于它能够增加网络的层次结构，使得网络可以更好地提取抽象的高阶特征。这样，代价函数的值就能够更快地收敛于最小值，有利于复杂网络学习模式，从而在一定程度上提升了模型性能和训练精度。举例来说，通过将三个步长为1的3x3小型卷积核叠加，可以得到与7x7卷积核相同的像素感受野大小。整个网络的参数量为3x(9xC2)，其中C为模型中输入层和输出层的通道数。

Vgg-16是一个深度网络结构，共包含16层，其中13层为卷积层，3层为全连接层。这一网络架构是在VggNet基础上延伸而来的。相对于Vgg-19，在分类任务中，Vgg-16具有更好的识别率。图2-10展示了Vgg-16的网络架构，而具体的模型网络参数可见表2-2。

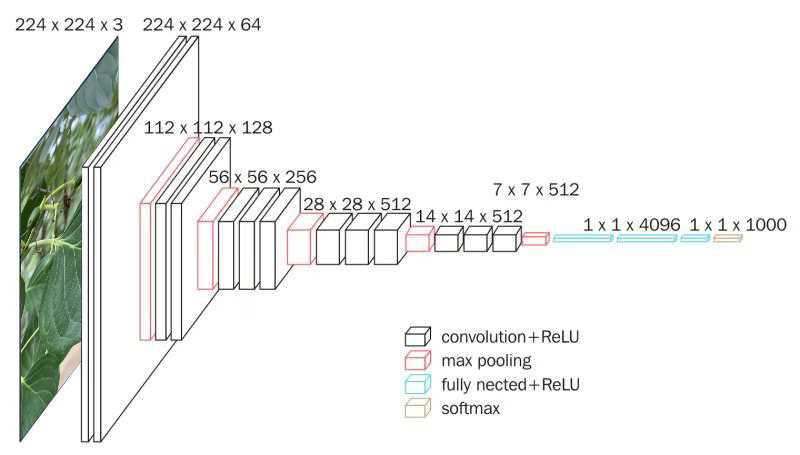
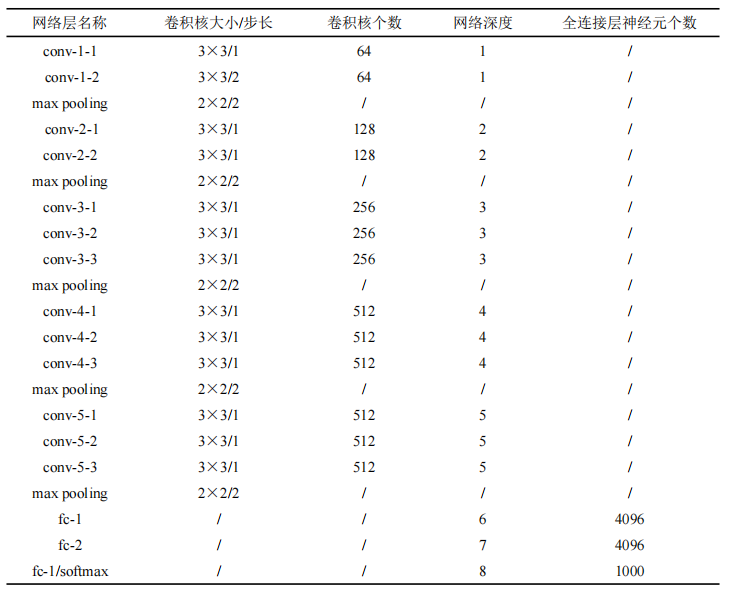


图2-10 VGG-16的网络架构图

表2-2 VGG-16 模型网络参数



2.3.3 GoogLeNet模型

GoogLeNet[45]是一种新型的网络结构，于2014年提出，并在ImageNet比赛中夺得第一名。尽管GoogLeNet有22层的深度，但其参数量仅为6.8百万，远少于传统网络如AlexNet和VGG等。然而，从模型表现来看，GoogLeNet的性能却更胜一筹。这是因为GoogLeNet采用了一种全新的Inception结构，Inception模块结构采用了四个并行路径，其中前三个路径分别通过1×1、3×3和5×5的卷积核获取不同区域的感受野信息，而第四个路径则使用3×3的最大值池化核提取特征点数据信息。在卷积过程中，引入了1×1的卷积核，虽然增加了深度，但有助于防止过拟合。为了降低模型复杂性和数据维度，第二条和第三条路径中增加了1×1的卷积操作。尽管这加深了网络结构，但却能有效提高精确度。Inception模块的结构如图2-11所示。

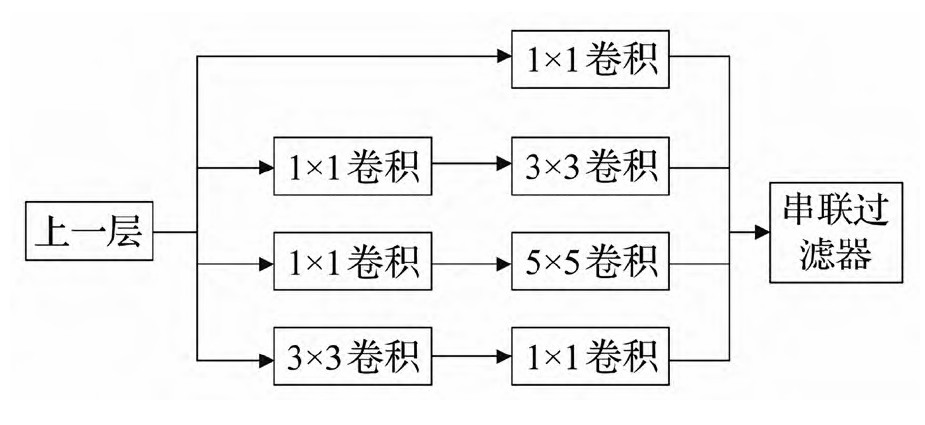


图2-11 Inception模块结构示意图

因此，使用GoogLeNet模型提取植物叶片图像的特征有助于迅速识别植物叶片。在CNN结构中，全连接层在目标图像的分布式特征表示中发挥着关键作用，能够充分呈现图像的全局特征信息。图2-12是GoogleNet的网络架构示意图

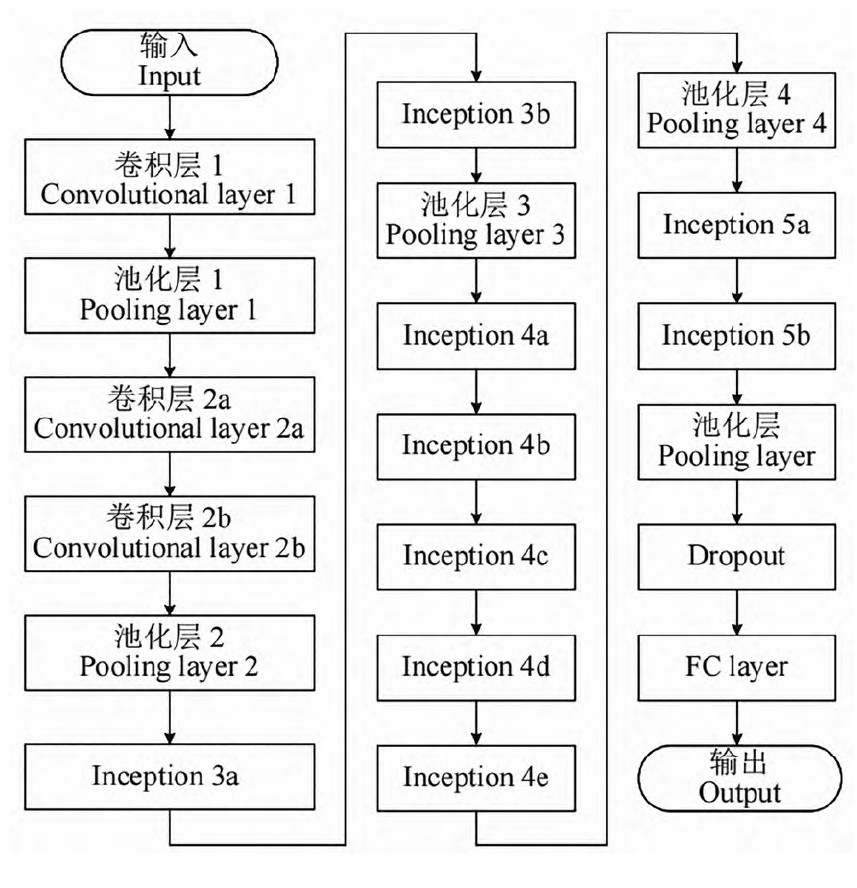


图2-12 经典GoogleNet网络结构

2.3.4 ResNet模型

2015年，何凯明等人提出了ResNet[46]网络，这一创新标志着卷积神经网络领域的又一次突破。该网络模型最多可达152层，具有相当深的层次结构，在当年在多个领域取得了显著的成绩，包括目标检测、语义分割和图像识别等方面，取得了具有历史意义的突破，其成果引人注目。与传统神经网络模型相比，ResNet不仅具有良好的学习效果，而且参数数量较低。ResNet的问世为研究者们提供了更深层次的训练方法。传统的深度学习方法通过反复堆叠卷积层来获取更好的特征，但随着层次深度的增加，模型的精度往往会出现先增加后下降的趋势。为了解决这一问题，ResNet团队引入了残差网络的概念。残差网络的核心思想是“恒等映射连接”，即每一层的输出保持不变，这样做的好处是减少了数据的冗余，更容易进行特征学习，并且残差映射比原始映射更容易优化。残差块的结构如图2-13所示。

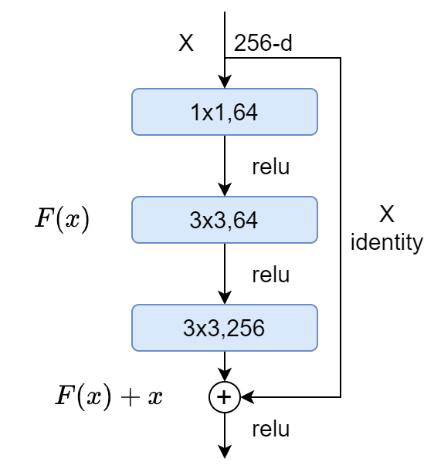


图2-13 残差块结构

在残差块中，假设x为模型的输入。由于存在Identity Mapping，H(x) = F(x)+x，其中F(x)代表模型的真实学习部分，可表达为F(x) = H(x)-x。

残差块在某些情况下能够增强网络模型中的特征提取，而不会带来负面影响。因此，残差网络能够在减少模型参数的同时提升神经网络的性能，这正是近年来残差网络被广泛应用的关键原因之一。

ResNet可以根据其深度分为不同版本，包括ResNet18（18层）、ResNet34（34层）、ResNet50（50层）、ResNet101（101层）和ResNet152（152层）。每个版本的ResNet具有不同的结构参数，如图2-14中所示。以ResNet18为例，其网络结构如图2-15所示。。

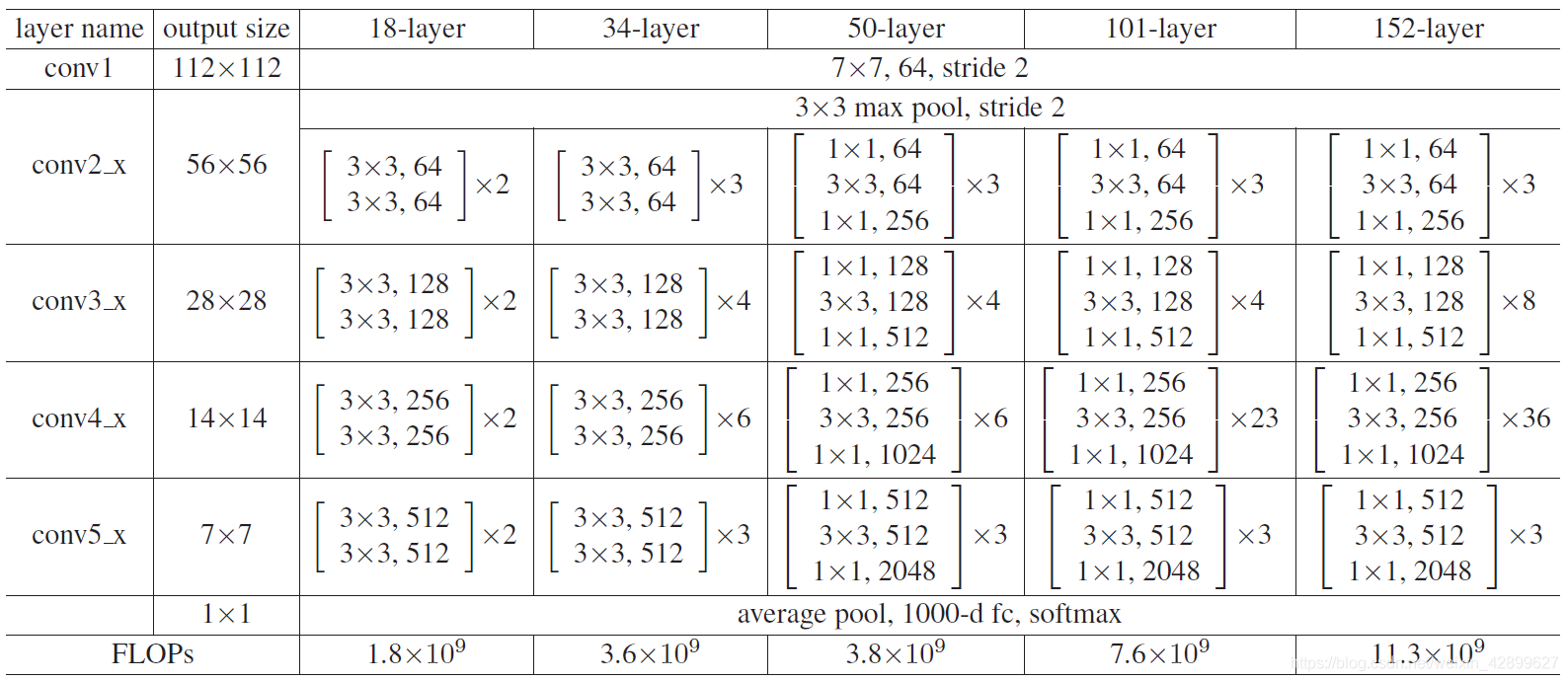


图2-14 不同层数的ResNet网络参数

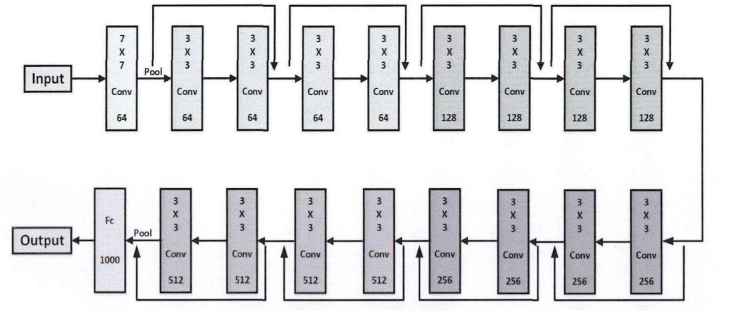


图2-15 ResNet-18结构图

2.4 总结

本章节首先介绍的是深度学习的基础知识，详细阐述了卷积神经网络的各个组成部分。然后介绍了深度学习的两种框架Pytorch和TensorFlow；最后，介绍了常用的植物叶片识别的网络模型，包括AlexNet，VGG，GoogLeNet和ResNet模型。

3 基于GoogLeNet网络模型的植物叶片识别

在深度学习领域，卷积神经网络（CNN）被广泛应用于植物叶片的识别任务。相比传统的基于特征识别方法，CNN模型能够自动学习和提取叶片的特征，从而实现更加精准的识别结果。为了训练CNN模型，需要大量的标注好的训练样本集作为网络的输入数据。因此，在进行训练之前，通常需要对植物叶片图像进行图像预处理，以减少噪声、增强特征等，然后将处理过的图像输入到相应的CNN模型中进行训练。在训练过程中，通过不断地调整模型的参数，可以逐步提高植物叶片的识别率[1]。

本章的主要研究目标是以GoogleNet网络模型为基础，将其应用到植物叶片识别的研究中，探索其在识别植物叶片方面的性能和适用性。

**3.1 数据集及预处理**

3.1.1 数据集描述

Flavia数据库[42]是一个被广泛应用于植物叶片分类和识别领域的公共数据库。该数据库包含了来自32个不同植物物种的1907个样本。每个样本都是彩色图像，其大小均为1600x1200像素，分辨率为300-DPI。在我们的研究中，为了保证样本的多样性和代表性，我们按照80%的比例将这些样本划分为训练集，20%划分为验证集，同时在对实验进行测试时选择20%作为测试集。这样的数据划分方式旨在充分利用数据库中的样本信息，同时保证了训练、验证和测试集之间的独立性和公平性。

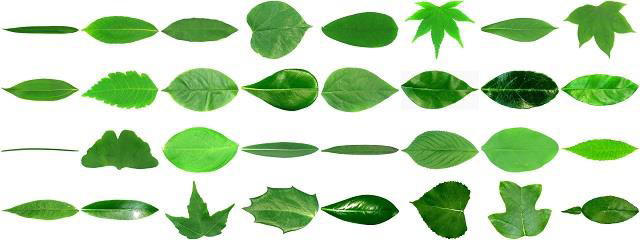


图3-1 Flavia数据库示例

3.1.2 数据预处理和增强

在计算机视觉领域，图像预处理是为了在将图像输入到CNN或其他深度网络模型中进行学习和识别之前对其进行处理。由于图像中的纹理特征受图像大小和方向的影响较大，因此在提取纹理特征之前，需要对图像进行适当的预处理。

在本次植物叶片分类识别实验中，数据预处理过程对于提升模型的训练效果至关重要。本文采用了一系列图像增强技术来丰富训练数据的多样性，并使用标准化方法来保证模型的输入一致性。首先，调整所有输入图像的大小至224×224像素，以适应模型的输入要求。统一的图像尺寸不仅减少了计算复杂度，还能确保模型对每个输入图像的处理一致性。接下来，为了增加训练数据的多样性，本文采用了随机垂直翻转和随机水平翻转两种图像增强技术。这种随机翻转操作可以有效防止模型对图像方向的过度依赖，从而提升模型的泛化能力。此外，对图像进行随机旋转，最大旋转角度为180度。这一处理进一步增加了数据集的多样性，使模型能够更好地适应不同角度下的叶片图像。最后，对植物叶片图像数据进行归一化处理。归一化操作使用预先计算的均值和标准差，使得每个通道的图像数据分布更加均匀。归一化不仅可以加速模型收敛，还能避免因输入数据差异过大导致的训练不稳定问题。

综上所述，这些预处理步骤通过多样化和标准化输入数据，提高了模型的训练效果和鲁棒性，为后续的植物叶片分类识别任务奠定了坚实的基础。

3.2模型设计

GoogleNet \_V1的网络结构框架如图3-2所示。从图3-2 中可以清楚的得知各层数的输入，输出以及卷积核规模。整体网络分为了三个不同的输出结果并在最终输出结果中合并。

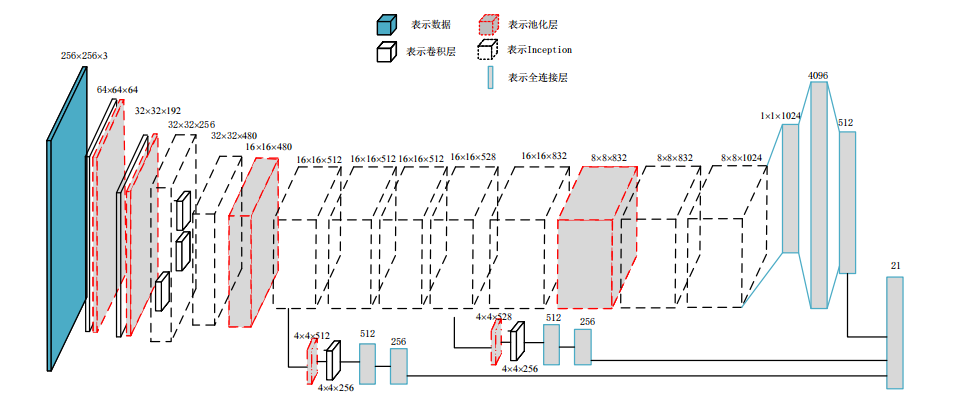


图3-2 GoogLeNet\_V1网络结构示意图

GoogLeNet包含22个层次，初始输入数据的尺寸为224×224×3。

第一个卷积层conv1具有3个填充，生成64个特征，卷积核大小为7×7，步长为2，输出特征为112×112×64。接着进行ReLU激活，然后通过pool1进行3×3的池化，步长为2，生成特征为56×56×64，最后进行归一化。

第二层卷积层conv2具有1个填充，生成192个特征，卷积核大小为3×3，输出特征为56 x 56 x 192。然后进行ReLU激活，并进行归一化。然后通过池化层pool2进行3×3的降采样，步长设为2，得到的特征图大小为28×28×192，最后将输出分成四个分支进行处理。

Inception模块在第三层开始发挥作用，它利用多尺度的卷积核进行特征提取。从第二层到第三层，网络被分为四个支路：

(1) 使用64个1×1的卷积核进行计算，经过ReLU激活后得到28×28×64的特征图。

(2) 使用96个1×1的卷积核进行维度缩减，得到28×28×96的特征图，然后经过ReLU激活后，再进行128个3×3的卷积，填充为1，得到28×28×128的特征图。

(3) 使用16个1 x 1的卷积核进行维度缩减，得到28×28×16的特征图，然后经过ReLU激活后，接下来进行32个5×5的卷积操作，填充设为2，得到的特征图大小为28×28×32。

(4) 进行3×3的池化，填充为1，得到28×28×192的特征图，接着执行32个1×1的卷积操作，生成的特征图大小为28×28×32。

最后将四个结果进行连接，得到的输出特征图大小为28×28×256。

在第二层到第三层的四条支线之后，开始Inception模块的具体操作如下所示：

(1) 使用128个1×1的卷积核进行计算，经过ReLU激活后得到28×28×128的特征图。

(2) 使用128个1×1的卷积核进行维度缩减，得到28×28×128的特征图，然后经过ReLU激活后，再进行192个3×3的卷积，填充为1，得到28×28×192的特征图。

(3) 使用32个1×1的卷积核进行维度缩减，得到28×28×32的特征图，然后经过ReLU激活后，再进行96个5×5的卷积，填充为2，得到28×28×96的特征图。

(4) 进行3×3的池化，填充为1，得到28×28×256的特征图，然后进行64个1×1的卷积，得到28×28×64的特征图。

最后将四个结果进行连接，得到的输出特征图大小为28×28×480。

3.3 实验过程

3.3.1 实验环境

在本研究中，我们选择了Python作为主要开发语言，因其具备跨平台兼容性和丰富的第三方库支持。我们采用了Anaconda3作为编译器，并结合PyCharm作为主要开发工具，以便于项目环境配置和代码调试。为了进行图像预处理，我们引入了OpenCV计算机视觉库，用于实现图像的灰度化、降噪和分割等处理方法。在深度学习方面，我们选择了PyTorch作为主要框架，以便于构建深层神经网络模型。在实验结果的展示方面，我们利用可视化工具matplotlib，根据训练数据的权重文件，绘制了准确率和损失函数曲线，以直观地呈现模型训练的效果；同时为了更好地观察实验结果，将混淆矩阵可视化。这样的工具组合和方法选择，旨在提高开发效率和实验结果的可解释性，为研究成果的展示提供了良好的支持。

表3-1 实验环境

|  |  |
| --- | --- |
| 硬件环境 | 软件环境 |
| CPU：AMD Ryzen 7 5800H with Radeon Graphics 3.20 GHz  GPU：NVIDIA GeForce RTX 3050 Laptop GPu  内存：16G | 操作系统：Windows11 64位  开发工具：Pycharm 2023  编译器：Anaconda3, Python3.10 |

3.3.2 参数设置

在模型训练过程中，网络使用随机梯度下降（SGD）优化算法，并且使用L2范数作为正则化系数，L2范数的权重衰减系数初始值为0.01, 可以帮助减少模型过拟合。学习率的初始值为0.02，使用指数衰减调整学习率 ExponentialLR来衰减学习率。将学习率在每个周期内按照以下调整公式进行衰减。

上述表达式中e是epoch计数器，即当前训练所处的epoch，gamma参数来控制学习率的衰减速度，设定为0.95。批处理量Batch\_size设定为32，共训练100个epoch。训练一个epoch需要花费1～3min。

3.3.3 算法评价指标

在我们的研究中，我们使用了多个评价指标来全面评估我们提出的植物叶片识别算法的性能。在数据集中，如果各个类别的分布相对平衡，通常会选择准确率（Accuracy）作为评估模型性能的指标。在我们的实验中，准确率反映了我们提出的算法在识别植物叶片时的整体分类准确性。假设总样本数为S，而正确分类的样本数为m，则准确率可表示为：

当数据集中各类别样本数量不平衡时，即各类别样本数量存在显著差异，模型在训练过程中会倾向于学习样本数量较多的类别特征。这会导致模型在识别样本数量多的类别时表现较好，而在识别样本数量少的类别时表现较差。在这种情况下，准确率并不是评估模型分类性能的最佳指标。

因此，为了在样本分布不均衡的情况下更准确地评估模型的分类性能，通常采用精确率（Precision）、召回率（Recall）和F1值（F1 Score）等指标。通常通过混淆矩阵计算得到这些性能评估指标，假设一个二分类混淆矩阵如表3-2所示。

表3-2 二分类混淆矩阵

|  |  |  |
| --- | --- | --- |
| 混淆矩阵 | 预测为正 | 预测为负 |
| 实际为正 | TP | FN |
| 实际为负 | FP | TN |

其中TP表示真正例，即实际为正样本且被模型正确预测为正样本的样本数量；FN表示假负例，即实际为正样本但被模型错误预测为负样本的样本数量；FP表示假正例，即实际为负样本但被模型错误预测为正样本的样本数量；TN代表真负例，即模型正确将实际为负样本且被正确预测为负样本的样本数量。

根据混淆矩阵和准确率的定义，准确率还可以表示为下式，它表示预测正确的样本数量在总样本数量中所占的比例。

精确率（Precision）可用以下公式表示，它指的是真实值为正且预测值也为正的样本数量在所有预测值为正的样本中的比例。在植物叶片识别的任务中，精确率衡量了模型在预测为某一类别时的准确性，高精确率意味着模型的预测结果中真正为正类的比例较高。

召回率（Recall）可由公式表示，表示真实值为正且预测值也为正的样本数量与真实值为正的样本数量之比。在植物叶片识别的任务中，召回率衡量了模型对正样本的识别能力。高召回率表示模型能够较好地捕获正样本。

根据公式和，可以得知，精确度和召回率这两个评估标准是相互对立的。当模型的精确度提高时，召回率就会降低；而当精确度下降时，召回率会增加。因此，需要一种合适的指标，来综合考量精确度和召回率，这个评估指标称为F1值。

F1-score值可由方程式表示，它代表着精确率和召回率的调和平均，综合考虑了这两个指标。F1值越高，F1-score的取值范围在0到1之间，越接近1表示模型的性能越好。

混淆矩阵也可叫做误差矩阵，以有n行n列的矩阵形式显示准确度的评价。这种精确度指标体现了不同类型的图像分类精度。在图像的准确性评价中，重点研究了分类结果与实际情况的对比，从而可以得出分类结果的准确性。混淆矩阵通过对图像中的每个特征的位置和顺序进行对比，对分类图像中的对应位置和顺序进行对比。

在评估我们的算法时，我们不仅关注了单一指标的表现，还综合考虑了多个指标，以确保对模型性能的全面评估。通过分析这些评价指标，我们能够更好地理解我们提出的算法在植物叶片识别任务中的优劣势，并且为进一步改进和优化算法提供了指导。

3.4 实验结果与分析

在上述实验过程中，将植物叶片数据集输入到GoogLeNet模型中训练，对模型参数进行设置和调整，经过训练后得到不同的识别率,并保存识别精度高的模型。利用Python语言引入了 OpenCV计算机视觉中的 matplotlib 库绘制模型的准确率及损失函数变化曲线图，同时为了更好地观测测试集识别的效果，将混淆矩阵进行可视化。

3.4.1 基于GoogLeNet网络模型的训练结果

根据模型最初设置的参数，得到训练结果。

如图3-3 GoogLeNet模型的准确率与Loss变化曲线为GoogLeNet模型的准确率与 Loss曲线，从图中可以看出，GoogLeNet的准确率随着迭代次数的增加在第40轮有了明显提升且呈缓慢上升状态，损失值迭代到 40 轮，Loss 曲线有了较为明显的下降趋势。由于验证集图像样本较少，从图中能够看到验证集的准确率波动较大。

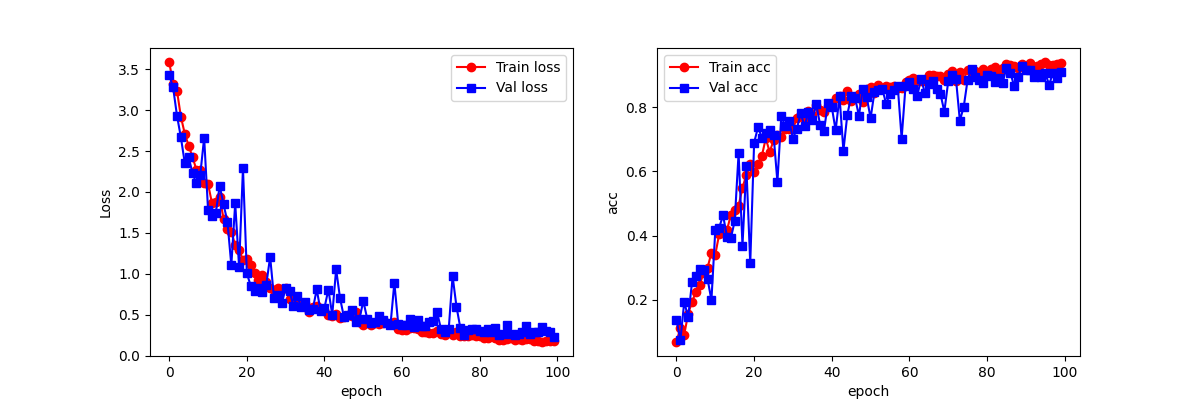


图3-3 GoogLeNet模型的准确率与Loss变化曲线

图3-4为GoogLeNet模型对测试集识别的结果，可以看出实验的测试准确率约为89.06%，精确率Precision约为90.29%，召回率Recall约为88.51%，F1-score约为88.47%；

图3-5为GoogLeNet模型是对于测试集32种植物叶片进行测试的混淆矩阵可视化，在混淆矩阵中，横轴上的标签标示样本的预测label，纵轴上的标签标示样本的实际label，对角线上的数字表示预测label和真是label一致的数量，也就是预测正确的数量。对于其他位置的数字就表示预测错误的。整体上来看对于测试集的预测正确率较高

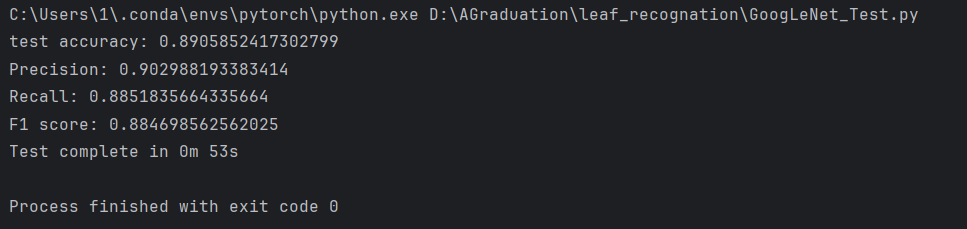


图3-4 GoogLeNet模型的测试结果

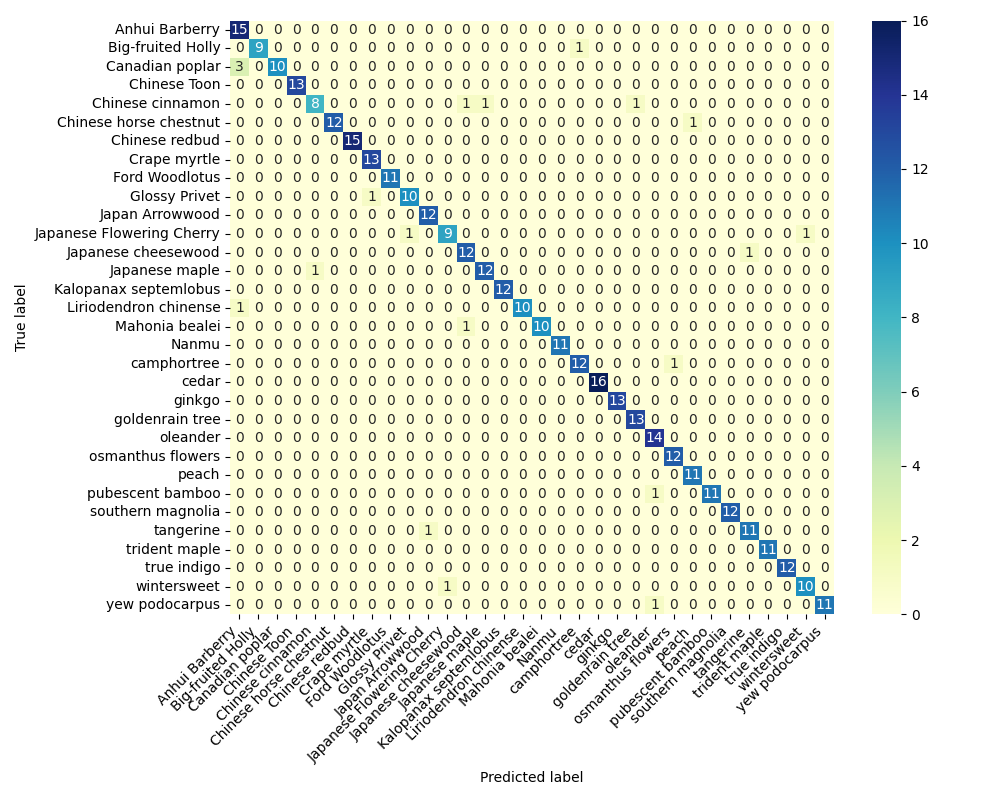


图3-5 GoogLeNet模型测试的混淆矩阵可视化

表3-3是GoogLeNet模型下的 Flavia 数据集中各类植物叶片识别准确率。

表3-3 GoogLeNet模型下各类植物叶片识别准确率

|  |  |  |  |
| --- | --- | --- | --- |
| 植物种类 | 识别率 | 植物种类 | 识别率 |
| 毛竹 | 91.67% | 雪松 | 100.00% |
| 中国七叶树 | 92.31% | 银杏 | 100.00% |
| 安徽小檗 | 100.00% | 紫薇 | 100.00% |
| 中国紫荆花 | 100.00% | 夹竹桃 | 100.00% |
| 蓝靛果 | 100.00% | 紫杉梅松 | 90.91% |
| 日本枫树 | 92.31% | 日本樱花 | 81.82% |
| 楠木 | 100.00% | 女贞 | 90.91% |
| 龙牙 | 100.00% | 中国香椿 | 100.00% |
| 中国肉桂 | 72.73% | 桃树 | 100.00% |
| 金雨树 | 100.00% | 木莲 | 100.00% |
| 大果冬青 | 80% | 三叉戟枫 | 100.00% |
| 日本干酪木 | 92.31% | 阔叶十大功劳 | 90.91% |
| 梅花 | 90.91% | 南方玉兰 | 100.00% |
| 樟树 | 92.31% | 加拿大杨树 | 76.92% |
| 日本箭树 | 100.00% | 鹅掌楸 | 90.91% |
| 桂花 | 100.00% | 柑橘 | 91.67% |

从表3-3中可看出，对于大部分物种GoogLeNet都可以得到很好的识别效果，但同样

也会出现误分类情况，比如日本樱花，加拿大杨树，中国肉桂，大果冬青。

4 基于ResNet网络模型的植物叶片识别

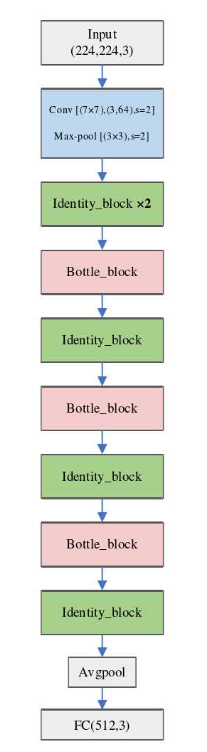
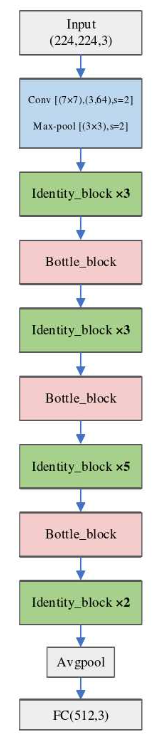
4.1模型设计

4.1.1 ResNet18模型结构

Resnet18的网络结构如图4- 1 ResNet18和ResNet34模型结构图 a)所示，该神经网络共包含18个层，其中包括17个卷积层和1个全连接层。数据首先经过一个卷积层，该卷积层使用的卷积核尺寸为7×7，步长为2。随后再经过一个最大池化层，池化操作会使特征图的尺寸缩小为原来的四分之一，即从224缩小至56，同时特征图的通道数也从3增加至64。

接着，特征图将先经过两个Identityblock残差块进行相应的处理，保持其通道数和特征图尺寸不变。随后，特征图会经由一个Bottleblock残差块来处理，导致特征图尺寸减小至原来的一半，通道数增加为原来的2倍，之后再经过一个Identityblock残差块，通道数和特征图尺寸保持恒定。在交替经过三个Bottleblock和Identityblock后，特征图尺寸减小至原始的1/8，由之前的56变为7，通道数增加至原来的8倍，由之前的64变为512。

最后，特征图将通过进行全局平均池化处理，然后被展开为一维数据，随后输入至一个包含512个神经元的全连接层。全连接层之后，会加入一个用于执行分类任务的分类器，用于完成植物叶片的识别分类。

a) ResNet18模型结构 b)ResNet34模型结构

图4- 1 ResNet18和ResNet34模型结构图

4.1.2 对ResNet18模型进行改进

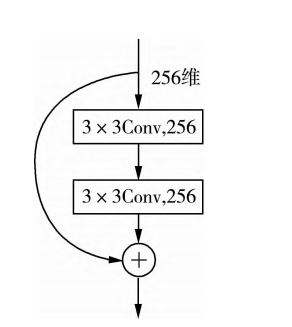
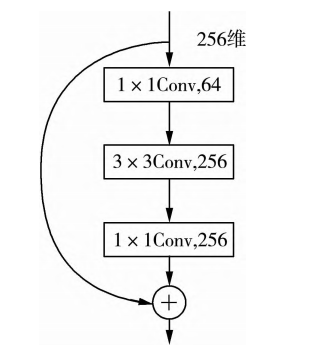
为了精确识别植物叶片图像类别，ResNet18模型表现出了明显的优势。然而，在实验中发现，ResNet18仍然存在一些无法避免的问题：(1)底层卷积层参数过多；(2)模型训练速度慢。为了解决这些问题，我们对ResNet18模型进行了调整，包括调整底层卷积尺寸和改进模型残差块结构。

1. 卷积核尺寸调整

在深度卷积神经网络中，虽然大尺寸卷积核能够捕获图像中更为复杂的特征，但其参数量较大，不适合用于植物叶片分类识别研究。相反，小尺寸卷积核参数量较小，但可能无法充分表达图像信息。为了在尽量减少模型精度损失的前提下降低参数量，我们选择将原有的7×7卷积替换为5×5卷积，以平衡模型参数和特征表达。

1. 残差块结构优化

为了减少ResNet18网络模型的参数量，我们对残差块的结构进行了优化。我们采用了1×1卷积、3×3卷积和1×1卷积串联的瓶颈结构，来替代原有的两个3×3卷积串联的结构。改进前和改进后的残差块结构如图4-2所示。

（a）改进前的结构 （b）改进后的结构

图4-2 残差块结构

以输入256维数据为例，使用改进前的残差块结构进行特征提取需要1179648个参数；而使用改进后的残差结构进行特征提取仅需要69632个参数。通过数据对比可知，本文改进的残差块结构显著降低了参数量，降幅达到了94%。改进后的残差块结构在计算量上有明显减少，有助于提高模型训练速度，并且提升了泛化性能。

经由上述卷积核尺寸调整、残差块优化后，所设计改进后的植物叶片识别模型如图4-3所示。

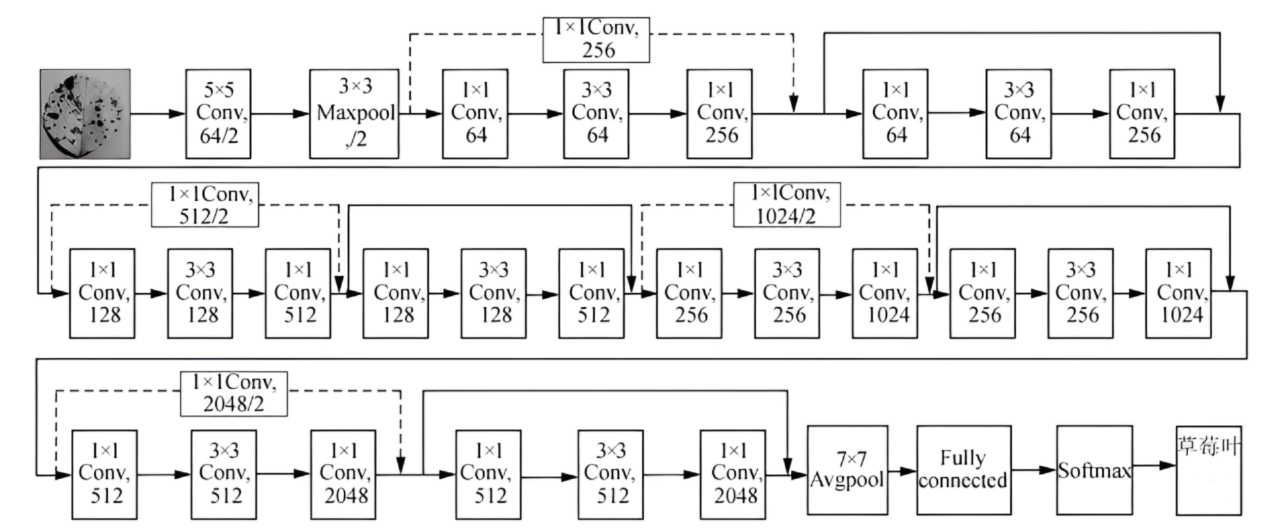


图4-3 改进后的ResNet18模型结构示意图

4.1.3 ResNet34模型结构

Resnet34的网络结构如图4- 1 b)所示，与ResNet18相似。输入数据首先通过一个7×7的卷积核，步长为2的卷积层，然后经过一个3×3大小的最大池化层。这使得特征图的尺寸从224缩小到56，通道数从3增加到64。

不同于Resnet18，Resnet34在相应位置增加了Identityblock的数量，以增加网络深度，而具有下采样功能的Bottleblock数量保持不变。在经过一个Identity block残差块后，通道数和特征图的尺寸保持不变；在经过一个Bottleblock残差块后，特征图的尺寸缩小为原来的1/2，通道数增加为原来的2倍。因此，经过3个Identity block、1个Bottle block、3个Identity block、1个Bottle block、5个Bottle block、1个Bottle block、2个Identity block的顺序后，特征图的尺寸经过缩小，从56变为7，同时通道数增加为原来的8倍，从64增至512。

接下来，特征图经过全局平均池化层处理，将其展平为一维数据，然后传递到一个包含512个神经元的全连接层。最后，全连接层后添加了一个用于执行分类的分类器，以完成植物叶片的识别分类任务。

4.2 实验过程

4.2.1 实验环境

在本研究中，我们选择了Python作为主要开发语言，选择PyTorch作为主要框架。

表4-1 实验环境

|  |  |
| --- | --- |
| 硬件环境 | 软件环境 |
| CPU：AMD Ryzen 7 5800H with Radeon Graphics 3.20 GHz  GPU：NVIDIA GeForce RTX 3050 Laptop GPu  内存：16G | 操作系统：Windows11 64位  开发工具：Pycharm 2023  编译器：Anaconda3, Python3.10 |

4.2.2 参数设置

在模型训练过程中，网络使用随机梯度下降（SGD）优化算法，并且使用L2范数作为正则化系数，L2范数的权重衰减系数初始值为0.01, 可以帮助减少模型过拟合。学习率的初始值为0.02，使用指数衰减调整学习率 ExponentialLR来衰减学习率。将学习率在每个周期内按照以下调整公式进行衰减。

上述表达式中e是epoch计数器，即当前训练所处的epoch，gamma参数来控制学习率的衰减速度，设定为0.95。批处理量Batch\_Size设定为32，共训练100个epoch。训练一个epoch需要花费1～3min。

4.2.3 算法评价指标

在我们的研究中，我们使用准确率（Accuracy）和混淆矩阵来评估我们提出的植物叶片识别算法的性能。

（1）准确率（Accuracy）：准确率是评价分类模型整体性能的指标，它表示模型正确分类的样本数量与总样本数量之比。在我们的实验中，准确率反映了我们提出的算法在识别植物叶片时的整体分类准确性。公式表示为：

这里，其中TP表示真正例，即实际为正样本且被模型正确预测为正样本的样本数量；FN表示假负例，即实际为正样本但被模型错误预测为负样本的样本数量；FP表示假正例，即实际为负样本但被模型错误预测为正样本的样本数量；TN表示真负例，即实际为负样本且被模型正确预测为负样本的样本数量。

（2）混淆矩阵:混淆矩阵也可叫做误差矩阵，以有n行n列的矩阵形式显示准确度的评价。这种精确度指标体现了不同类型的图像分类精度。在图像的准确性评价中，重点研究了分类结果与实际情况的对比，从而可以得出分类结果的准确性。混淆矩阵通过对图像中的每个特征的位置和顺序进行对比，对分类图像中的对应位置和顺序进行对比。

在评估我们的算法时，我们不仅关注了单一指标的表现，还综合考虑了多个指标，以确保对模型性能的全面评估。通过分析这些评价指标，我们能够更好地理解我们提出的算法在植物叶片识别任务中的优劣势，并且为进一步改进和优化算法提供了指导。

4.3 实验结果与分析

在上述实验过程中，将植物叶片数据集输入到ResNet18以及ResNet34模型中训练，对模型参数进行设置和调整，经过训练后得到不同的识别率,并保存识别精度高的模型。利用Python语言引入了 OpenCV计算机视觉中的 matplotlib 库绘制模型的准确率及损失函数变化曲线图，同时为了更好地观测测试集识别的效果，将混淆矩阵进行可视化。

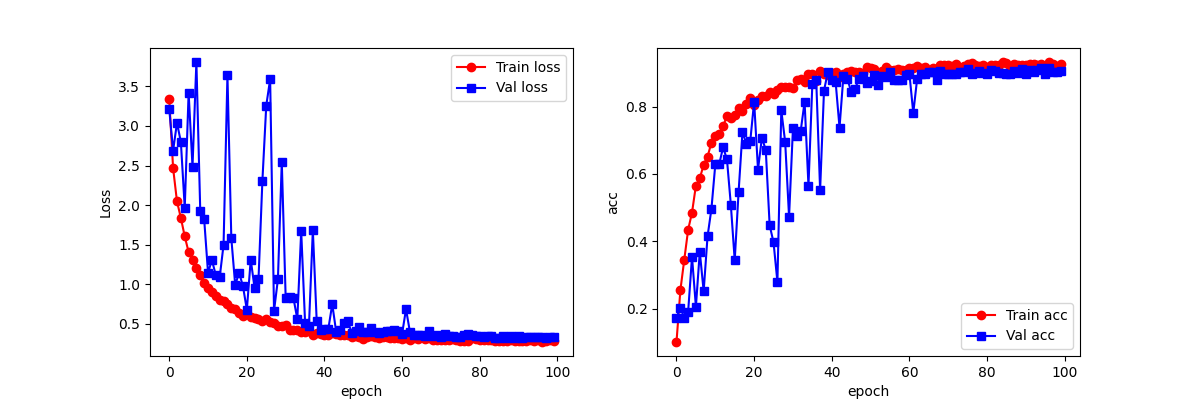


图4-4 ResNet18模型的准确率与Loss变化曲线

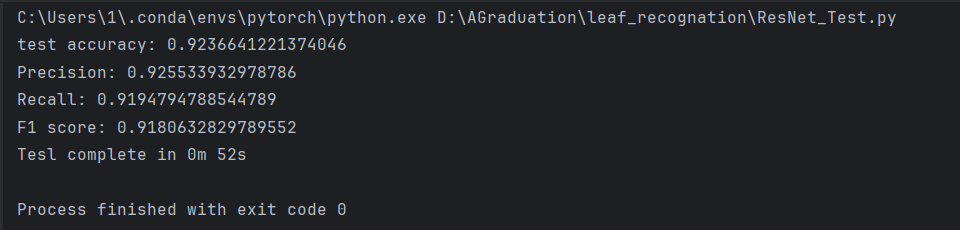


图4-5 ResNet18模型的测试结果

4.3.1 ResNet18改进前后的实验对比

为确定ResNet18模型改进中所采用的卷积核尺寸调整、残差块优化对ResNet18模型植物叶片识别能力的具体提升效果，设计测试试验如表4-2所示，其中方案4为本章改进后的ResNet18模型

表4-2 测试设计方案

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 编号 | 底层卷积核 | | 残差快 | |
| 7×7卷积核 | 5×5卷积核 | 捷径结构 | 瓶颈结构 |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
| 3 |  |  |  |  |
| 4 |  |  |  |  |

表4-3各方案实验结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 编号 | 准确率 | 精确率 | 召回率 | F1分数 |
| 1  2 | 92.37%  92.89% | 92.55%  92.81% | 91.95%  92.62% | 91.81%  92.35% |
| 3  4 | 92.94%  94.15% | 93.22%  94.54% | 93.54%  94.02% | 92.96%  94.02% |

由根据表4-3的测试结果，方案4改进后的ResNet18模型在植物叶片识别方面表现最佳，准确率达到了94.15%。

相比方案1，在采用瓶颈结构替代捷径结构后，模型的准确率提升了0.52%。而在方案3中，通过使用5×5卷积核替代原有的7×7卷积核，模型的准确率提升了0.57%

综合而言，方案4采用了5×5卷积核替代原有的7×7卷积核，并采用了瓶颈结构替代捷径结构，使模型在准确率得到了有效的提升。

4.3.2基于ResNet网络模型的训练结果

在本实验中，还使用了ResNet18和ResNet34两个网络模型进行对比实验，实验中的实验环境、使用到的训练集、验证集、测试集以及参数设置均保持相同，这样就可以确保实验结果具有可比性和可靠性。

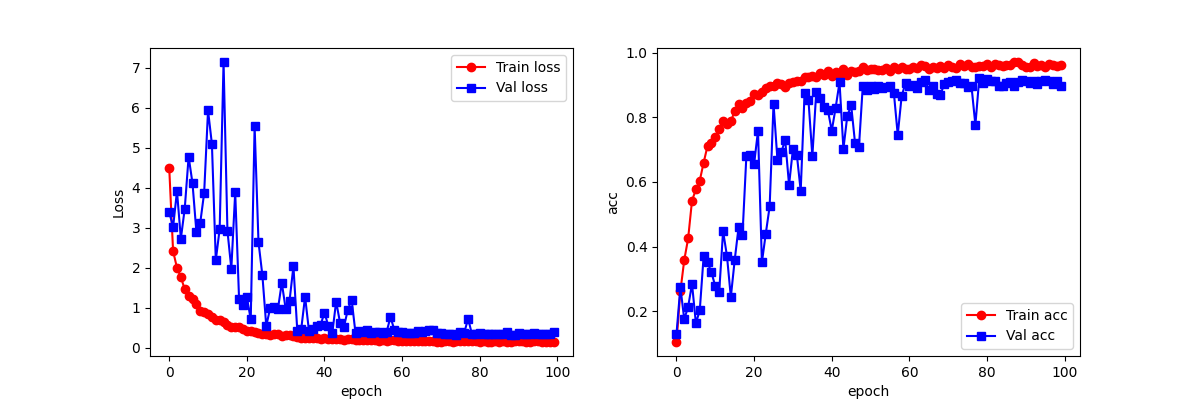


图4-6 ResNet34模型的准确率与Loss变化曲线

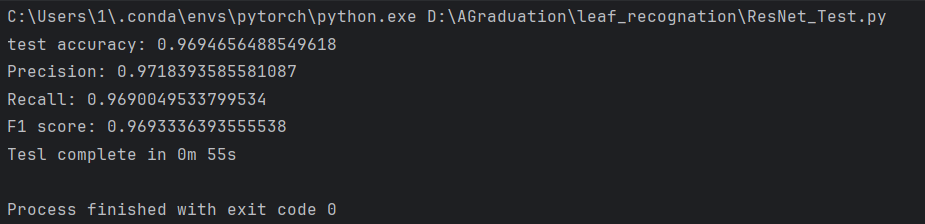


图4-7 ResNet34模型的测试结果

表4- 4 ResNet18和ResNet34模型实验结果对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模型 | 准确率 | 精确率 | 召回率 | F1分数 |
| ResNet18 | 92.60% | 92.34% | 92.47% | 92.42% |
| ResNet34 | 96.95% | 97.18% | 96.90% | 96.93% |

通过实验可以看出，ResNet34模型相比于ResNet18模型（改进前）对植物叶片识别准确率平均提升了约4.35个百分比

4.3.3 不同实验参数下的ResNet34模型的实验分析

在实验中，采用了批量训练的方法，通过设置合理的批次大小来确保模型梯度向量的准确性，从而保证以较快的速度收敛到局部最小值。初始时，每批次训练包含24张图片，根据训练过程中损失值的变化进行动态调整，逐步增加至32和48张图片。初始学习率设置为0.008，在实验过程中改变初始学习率为0.01和0.02进行实验，通过不同的batch\_size和学习率来分析得到最佳的参数。

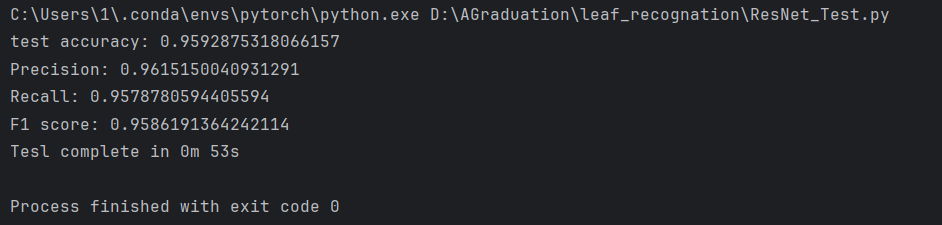


图4-8 ResNet34模型(batch\_size=24)的测试结果

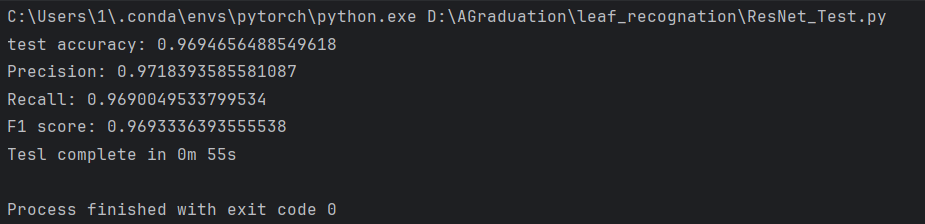


图4-9 ResNet34模型(batch\_size=32)的测试结果

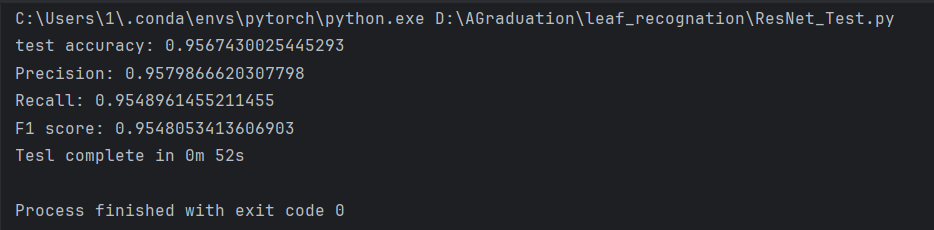


图4-10 ResNet34模型(batch\_size=48)的测试结果

表4- 5 ResNet34模型不同Batch\_Size实验结果对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Batch\_Size | 准确率 | 精确率 | 召回率 | F1分数 |
| 24 | 95.93% | 96.15% | 95.78% | 95.86% |
| 32 | 96.95% | 97.18% | 96.90% | 96.93% |
| 48 | 95.67% | 95.80% | 95.49% | 95.48% |

批量训练数量（batch\_size）对模型性能有着重要影响。batch\_size 对模型训练结果有一定的影响，从实验结果可以看出，batch\_size较小或者较大，准确率和精确率都会有一定程度的下降。合理的batch\_size能得到较高的准确率,通过实验可以看出训练模型在 batch\_size为32时训练准确率最高，为93.64%。

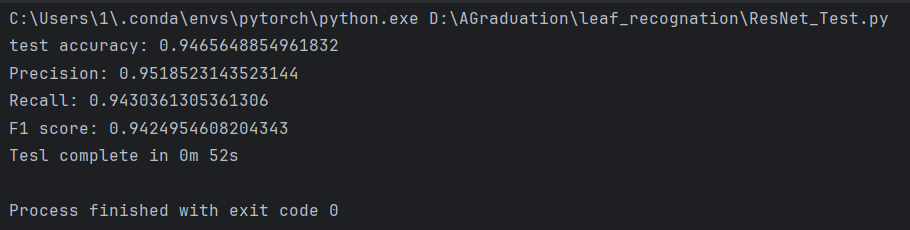


图4-11 ResNet34模型(lr=0.008)的测试结果

表4-6 ResNet34模型不同学习率实验结果对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 学习率lr | 准确率 | 精确率 | 召回率 | F1分数 |
| 0.008 | 94.66% | 95.19% | 94.30% | 94.24% |
| 0.01 | 95.96% | 95.34% | 95.47% | 95.42% |
| 0.02 | 96.95% | 97.18% | 96.90% | 96.93% |

超参数——学习率lr对模型训练有着较大的影响，实验过程中通过改变初始学习率为0.008和0.01,0.02来观测实验结果，通过表4-6对植物叶片测试集进行测试，可以看出学习率为0.02时训练模型的测试准确率最高，为96.95%

4.4 总结

本章通过ResNet18和ResNet34网络模型对植物叶片进行分类识别。实验过程中首先对ResNet18模型进行训练，发现ResNet18仍然存在一些无法避免的问题：(1)底层卷积层参数过多；(2)模型训练速度慢。为了解决这些问题，我们对ResNet18模型进行了调整，包括调整底层卷积尺寸和改进模型残差块结构。实验测试表明这两种改进方法可以有效地提升植物叶片的识别准确率；接着通过ResNet34模型进行训练，和ResNet18模型识别效果进行对比，可以得出层数的增加可以提升植物叶片的识别准确率。

结论

本论文利用Flavia植物叶片数据集32种植物叶片，基于深度学习算法对植物叶片进行分类识别进行了相应的研究，本文主要是通过GoogLeNet模型和ResNet模型来实现植物叶片识别，总结如下：

(1)第一个是通过GoogLeNet网络模型对植物叶片识别，在实验过程中不断调整超参数比如批训练量batch\_size，学习率lr，对植物叶片测试集进行测试实验，通过混淆矩阵进行可视化，实验表明，GoogLeNet对植物叶片的识别准确率平均达到了89.06%，对于大部分植物叶片都有良好的识别率。

(2)第二个是通过ResNet网络模型对植物叶片识别，实验过程中首先对ResNet18模型进行训练，对测试集测试平均识别准确率为92.37%，同时也发现ResNet18仍然存在一些无法避免的问题：(1)底层卷积层参数过多；(2)模型训练速度慢。为了解决这些问题，我们对ResNet18模型进行了调整，包括调整底层卷积尺寸和改进模型残差块结构，将原有的7×7卷积替换为5×5卷积并且采用了1×1卷积、3×3卷积和1×1卷积串联的瓶颈结构，来替代原有的两个3×3卷积串联的结构。改进之后的识别准确率道道了94.15%，相比于改进前提升了1.18个百分比，实验测试表明这两种改进方法可以有效地提升植物叶片的识别准确率；接着通过ResNet34模型进行训练，和ResNet18模型识别效果进行对比，ResNet34对植物叶片的平均识别准确率达到了96.95%，相比于ResNet18模型识别准确率提高了4.58个百分比，可以得出层数的增加可以提升植物叶片的识别准确率。

最后，由于我的水平以及计算机的硬件有限，有些问题还有待研究。

参 考 文 献

1. 傅弘, 池哲儒, 常杰. 基于人工神经网络的叶脉信息提取——植物活体机器识别研究Ⅰ[J].植物学通报, 2004, (04): 429-436.
2. 杜吉样, 植物物种机器识别技术的研究[D]. 安徽: 中国科学技术大学,2005.
3. 王晓峰, 黄德双, 杜吉祥, 张国军. 叶片图像特征提取与识别技术的研究[J].计算机工程与应用, 2006, (03): 190-193.
4. 朱宁. 基于 LBP 的树叶识别系统研究与实现[D]. 北京林业大学, 2008.
5. 贺鹏. 基于叶片综合特征的阔叶树机器识别研究[D]. 西北农林科技大学, 2008.
6. 阚江明, 王怡萱, 杨晓微, 冷萃. 基于叶片图像的植物识别方法[J]. 科技导报, 2010, 28(23):81-85.
7. 张宁. 基于图像分析的植物叶片识别算法研究[D]. 北京林业大学, 2013.
8. 王丽君. 基于叶片图像多特征提取的观叶植物种类识别[D]. 北京林业大学, 2014.
9. 杨天天, 潘晓星, 穆立蔷. 基于叶片图像特征数字化信息识别 7 种柳属植物[J]. 东北林业大学学报, 2014, 42(12): 75-79.
10. 于慧伶, 麻峻玮, 张怡卓.基于双路卷积神经网络的植物叶片识别模型[J]. 北京林业大学学报, 2018, 40(12): 132-137.
11. 朱良宽, 晏铭, 黄建平. 一种新型卷积神经网络植物叶片识别方法[J]. 东北林业大学学报, 2020, 48(04): 50-53.
12. 李龙龙, 何东健, 王美丽. 基于改进型 LBP 算法的植物叶片图像识别研究[J/OL]. 计算机工程与应用: 1-10[2021-01-24].
13. 张露. 基于深度学习的植物叶片图像识别方法研究[D]. 北京林业大学, 2019
14. 朱静, 田兴军, 陈彬, 吕劲紫. 植物叶形的计算机识别系统[J]. 植物学通报, 2005(05): 89-94.
15. 龚丁禧, 曹长荣. 基于卷积神经网络的植物叶片分类[J]. 计算机与现代化, 2014(04): 12-15.
16. 丰晓霞. 基于深度学习的图像识别算法研究[D]. 太原理工大学, 2015.
17. 张帅. 基于深度学习的植物叶片识别算法研究[D]. 北京林业大学, 2016.
18. 薄琪苇. 基于卷积神经网络的植物叶片识别研究[D]. 浙江农林大学, 2018.
19. 程曦, 吴云志, 张友华. 基于深度卷积神经网络的储粮害虫图像识别[J]. 中国农学通报, 2018, 34(01): 154-158.
20. 黄志国. 基于深度学习的中草药叶片识别算法的研究与应用[D]. 湖北工业大学, 2019.
21. 王艳, 孙薇, 周小平. 基于深度学习的中草药植物图像识别方法研究[J]. 医药信息, 2020, 37(06):21-25.
22. 孙颖异, 李健, 时天等. 基于改进的 AlexNet 卷积神经网络的植物叶片识别[J]. 种子, 2020, 39(02): 77-81.
23. 樊湘鹏, 周建平, 许燕, 彭炫. 基于改进卷积神经网络的复杂背景下玉米病害识别[J]. 农业机械学报:1-12[2021-01-11].
24. Pimm S. L., Jenkins C. N., Abell R., et al. The biodiversity of species and their rates of extinction, distribution, and protection[J]. Science, 2014, 344(6187): 1246752.
25. Knapp A. K., Fay P. A., Blair J. M. , et al. Rainfall Variability, Carbon Cycling, and Plant Species Diversity in a Mesic Grassland[J]. Science, 2002, 298(5601): 2202-2205.
26. Zhang S., Huang W., Huang Y., et al. Plant species recognition methods using leaf image:Overview[J]. Neurocomputing, 2020, 408: 246-272.
27. Wldchen J., Mder P.. Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review[J]. Archives of Computational Methods in Engineering, 2018, 25(2):507–543.
28. Thyagharajan K. K., Raji I. K.. A review of visual descriptors and classification techniques used in leaf species identification[J]. Archives of Computational Methods in Engineering, 2019, 26(4):933-960.
29. Tan C., Sun F., Kong T., et al. A survey on deep transfer learning[A]. International conference on artificial neural networks[C]. Cham: Springer, 2018: 270-279.
30. Ingrouillem J, Lairds M. A quantitative approach to oak variability in some north London woodlands[J]. The London Naturalist, 1986, 65: 35-46.
31. Osikar J O. Computer vision classification of leaves from Swedish trees[J]. Linkoping: Linkoping University, 2001, 181-186.
32. Rashad M Z, Eldesouky B S, Khawasik M S. Plants Images Classification Based on Textural Features using Combined Classifier[J]. International Journal of Computer Science & Information Technology, 2011, 3(4): 93-100.
33. Lee K B, Hong K S. An Implementation of Leaf Recognition System using Leaf Vein and Shape[J]. International Journal of Bio-Science & Bio-Technology, 2013, 214(2): 109-116.
34. Mehdipour Ghazi, Mostafa, Berrin, et al. Sabanci-Okan system in LifeCLEF 2015 plant identification competition[C]. CLEF 2015: 23-30.
35. Divya Tomar, Sonali Agarwal. Leaf Recognition for Plant Classification Using Direct Acyclic Graph Based Multi-Class Least Squares Twin Support Vector Machine[J]. International Journal of Image and Graphics, 2016, 16(3).
36. Krizhevsky A , Sutskever I , Hinton G . ImageNet Classification with Deep Convolutional Neural Networks[C]// NIPS. Curran Associates Inc. 2012.
37. Kho S J , Manickam S , Malek S , et al. Automated plant identification using artificial neural network and support vector machine[J]. Frontiers in Life Science, 2017, 10(1):98-107.
38. Imah E M, Rahayu Y S, Wintarti A. Plant Leaf Recognition Using Competitive Based Learning Algorithm[J]. Iop Conference, 2018, 288(1).
39. Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. Neuralco mputation, 2006, 18(7): 1527-1554.
40. Jeon W, Rhee S. Plant Leaf Recognition Using a Convolution Neural Network[J]. International Journal of Fuzzy Logic and Intelligent Systems, 2017, 17(1): 26-34.
41. Ghazi, MM, Yanikoglu, et al. Plant identification using deep neural networks via optimization of transfer learning parameters[J]. Neurocomputing, 2017,685(2): 132-231.
42. Wu S G,Bao F S,Xu E Y,et al.A leaf recognition algorithm for plant classificationusing probabilistic neural network[C]//2007 IEEE international symposium on signalprocessing and information technology.IEEE,2007:11-16.
43. Boya Zhao, Mingjiang Wang,Ming Liu. An energy-efiicient coarse grained spatial architecture for convolutional neural networks AlexNet[J. IEICE Electronics Express,2017,14(15):595
44. Qu X, Lu H, Tang W, et al. A VGG attention vision transformer network for benign and malignant classification of breast ultrasound images[J]. Medical Physics, 2022, 49(09): 5787-5798.
45. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]Proceedings of the lEEE conference on computer vision and pattern recognition. 2015: 1-9.
46. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

致 谢

行文至此，落笔之处皆为序幕。随着毕业论文的结束，这一程大学四年生涯就此谢幕，下一程真正的人生才刚刚掀开帷幕。在这段特殊的旅程中，有太多生命中重要的人为我撑伞，他们是父母、良师亦是挚友，带给我拼搏的勇气、无畏的精神，任我昂首阔步，追逐远方。

感谢父母庇护，任我自由翱翔。你们尊重我所有的决定并给予我最大的信心，你们的支持和鼓励是我成长道路上最坚实的后盾，是我追求梦想的强大动力。

感谢我的毕业设计导师——邝祝芳老师，从选题到实验，从撰写论文到定稿，邝老师给我提出了很多宝贵的意见，邝老师渊博的专业知识、严谨的治学态度对我影响深远。

感谢大学四年教过我的老师们，谢谢你们在专业学习上给我的帮助；感谢辅导员谢丽娟老师，在我担任班长期间，给予了我巨大的帮助和支持。

感谢实习遇到的所有导师们，在实习过程中帮助我提升了技术能力和实践经验，使我在专业领域得以快速成长。

感谢大学期间陪伴我的同学和朋友们。你们在学习和生活中的陪伴和帮助，让我的大学生活充满了欢乐和温暖。感谢你们在我需要时给予的帮助和关怀，让我在大学期间感受到浓浓的友情和真挚的友谊。

最后，感谢参加答辩的各位老师，感谢所有在我成长道路上给予过我帮助和支持的人，你们的关心和支持让我充满信心，勇敢前行。谢谢你们！祝你们生活幸福、万事如意！