

分类号 TP391.1
U D C

密级
编号 10741

兰州财经大学

LANZHOU UNIVERSITY OF FINANCE AND ECONOMICS

硕士学位论文

论文题目 基于深度学习的高原夏菜分级模型优化

研究生姓名: 郭泓

指导教师姓名、职称: 尚庆生 教授

学科、专业名称: 管理科学与工程

研究方向: 数据分析与信息处理

提交日期: 2024年5月31日

独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名：郭泓 签字日期：2024.5.31

导师签名：尚庆生 签字日期：2024.5.31

关于论文使用授权的说明

本人完全了解学校关于保留、使用学位论文的各项规定，同意（选择“同意”/“不同意”）以下事项：

1. 学校有权保留本论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文；

2. 学校有权将本人的学位论文提交至清华大学“中国学术期刊（光盘版）电子杂志社”用于出版和编入CNKI《中国知识资源总库》或其他同类数据库，传播本学位论文的全部或部分内容。

学位论文作者签名：郭泓 签字日期：2024.5.31

导师签名：尚庆生 签字日期：2024.5.31

Optimization of Plateau Summer Vegetable Grading Model Based on Deep Learning

Candidate : Guo Hong

Supervisor: Shang Qingsheng

摘要

随着生活水平的提高,年轻一代对多样化需求的偏好显著增长,尤其在水果和蔬菜的质量及品种上。尽管甘肃省的高原夏菜产业被认为是省级特色产业,但是受到人工分拣成本限制,分级标准并没有在销售环节推广。本文利用深度学习技术的进步,提出了一种基于图像识别的高原夏菜质量分级方法,考虑生产环境硬件资源的限制,提出的模型尽量降低时间和空间复杂度,为方法落地实施奠定基础。具体研究内容包括:

(1) 构建高原夏菜数据集。为应对高原夏菜图像数据集不足的挑战,设计并实施了一套高原夏菜图像收集装置。利用此装置,成功采集到了包括结球甘蓝、娃娃菜、菜花、以及西兰花在内的四种蔬菜图像,共计构建了含 2400 张图片的初始数据集。

(2) 扩充原始数据集。为提高模型训练的准确性和效率,本研究在扩展原始高原夏菜数据集方面进行了细致工作。首先,通过粗略分割技术去除了图像中的背景干扰,确保数据的质量。随后,借助拍摄扩展和数据增强手段扩充了初始数据集,从而为卷积神经网络模型的训练提供了更为丰富和多样化的图像资源,进一步优化了训练过程的效果。

(3) 提出了针对高原夏菜质量分级的多尺度融合 ECA-DS-ResNet50 模型。这种融合模型结合了多种技术,在 ResNet50 和 MobileNetV2 的基础上进行了改进和融合。首先,ResNet50 作为主干网络,通过引入 ECA (Efficient Channel Attention) 注意力机制和多尺度特征提取的 ASPP 模块,进一步增强了对图像特征的抽取能力。其次,针对 ResNet50 中 7x7 卷积核和池化层的计算成本较高的特点,将其 3x3 标准卷积替换成深度可分离卷积,整合 MobileNetV2 的 DW (Depthwise Convolution) 和 PW (Pointwise Convolution) 卷积机制取代 ResNet50 的常规卷积,有效降低了计算量。同时,在每层卷积后引入池化层,进一步减少参数量,优化网络结构,大幅减少了时间和空间复杂度,有效提升模型的运算速度。

本研究通过对预处理后的高原夏菜数据集在改进后的 ResNet 模型以及其他经典深度神经网络模型上进行训练,并通过消融对比实验进行评估。实验结果显示,本文提出的方法在准确率和模型识别速度方面均显著优于其他网络以及原始

未改进的网络,这充分证明了我们所改进的网络模型在有效性和可行性方面的显著优势。

关键词: 高原夏菜 质量分级 ResNet 网络 ECA 注意力模块

Abstract

With the improvement of living standards, the younger generation's preference for diversified needs has significantly increased, especially in the quality and variety of fruits and vegetables. Although the plateau summer vegetable industry in Gansu Province is considered a provincial-level characteristic industry, the classification standards have not been promoted in the sales process due to limitations in manual sorting costs. This article utilizes the advancement of deep learning technology to propose a high-altitude summer vegetable quality grading method based on image recognition. Considering the limitations of hardware resources in the production environment, the proposed model minimizes time and space complexity as much as possible, laying the foundation for the implementation of the method. The specific research content includes:

(1) Build a dataset for plateau summer cuisine. To address the challenge of insufficient data sets for high-altitude summer vegetable images, a set of high-altitude summer vegetable image collection devices has been designed and implemented. Using this device, four vegetable images including cabbage, baby bok choy, cauliflower, and broccoli were successfully collected, and an initial dataset containing 2400 images was constructed.

(2) Expand the original dataset. In order to improve the accuracy and efficiency of model training, this study conducted meticulous work in

expanding the original high-altitude summer vegetable dataset. Firstly, the background interference in the image was removed through rough segmentation techniques to ensure the quality of the data. Subsequently, the initial dataset was expanded using shooting expansion and data augmentation techniques, providing richer and more diverse image resources for the training of convolutional neural network models, further optimizing the effectiveness of the training process.

(3) A multi-scale fusion ECA DS ResNet50 model was proposed for the quality grading of summer vegetables on the plateau. This fusion model combines multiple technologies and cleverly improves and integrates ResNet50 and MobileNetV2. Firstly, ResNet50 serves as the backbone network, further enhancing the ability to extract image features by introducing ECA (Efficient Channel Attention) attention mechanism and ASPP module for multi-scale feature extraction. Secondly, in response to the high computational cost of the 7×7 convolution kernel and pooling layer in ResNet50, the 3×3 standard convolution was replaced with depthwise separable convolution, and MobileNetV2's DW (Depth Convolution) and PW (Pointwise Convolution) convolution mechanisms were integrated to replace ResNet50's conventional convolution, effectively reducing computational complexity. At the same time, introducing pooling layers after each convolution layer further reduces the number of parameters, optimizes the network structure, significantly

reduces time and space complexity, and effectively improves the computational speed of the model.

This study trained the preprocessed high-altitude summer vegetable dataset on an improved ResNet model and other classic deep neural network models, and evaluated it through ablation comparison experiments. The experimental results show that the method proposed in this article is significantly better than other networks and the original unimproved network in terms of accuracy and model recognition speed, which fully demonstrates the significant advantages of our improved network model in terms of effectiveness and feasibility.

Keywords: Plateau summer vegetables; Quality grading; ResNet network; ECA Attention Module

目 录

1 引言	1
1.1 研究背景和意义	1
1.2 研究现状	2
1.2.1 基于机器学习的蔬菜图像分类	3
1.2.2 基于深度学习的蔬菜图像分类	5
1.3 研究内容及创新点	6
1.3.1 研究内容	6
1.3.2 研究框架	8
1.3.3 创新点	8
1.4 本章小结	9
2 相关理论基础介绍	10
2.1 传统神经网络模型	10
2.2 卷积神经网络模型	12
2.2.1 输入层	13
2.2.2 卷积层	13
2.2.3 池化层	15
2.2.4 全连接层	16
2.2.5 输出层	17
2.3 ResNet 的理论基础	17
2.3.1 残差学习	17
2.3.2 ResNet 模型结构	19
2.4 MobileNetV2 的理论基础	20
2.4.1 深度可分离卷积	20
2.4.2 倒残差结构	22
2.5 注意力机制模块	23
2.5.1 CBAM 模块	23

2.5.2 SENet 模块.....	24
2.6 本章小结.....	26
3 高原夏菜数据集构建及其预处理.....	27
3.1 高原夏菜数据集的构建.....	27
3.1.1 图像采集.....	27
3.1.2 数据集存在问题.....	28
3.2 图像分割和背景移除.....	29
3.3 数据增强.....	30
3.3.1 图像的空间几何变换.....	31
3.3.2 图像添加噪声.....	31
3.4 高原夏菜分级标准.....	32
3.5 本章小结.....	34
4 基于深度学习的高原夏菜分级模型构建及优化.....	36
4.1 ResNet50 网络模型结构.....	36
4.2 注意力机制模块.....	38
4.3 卷积替换.....	40
4.4 ASPP 模块.....	41
4.5 优化网络参数.....	42
4.5.1 损失函数.....	42
4.5.2 AdamW 优化器.....	43
4.6 ECA-DS-ResNet50 融合网络模型构建.....	44
4.7 本章小结.....	48
5 实验结果与分析.....	49
5.1 高原夏菜数据集划分.....	49
5.2 实验.....	49
5.2.1 实验内容.....	49
5.2.2 实验环境及参数设置.....	50
5.2.3 评价指标.....	51

5.3 蔬菜类别分类.....	52
5.4 不同质量蔬菜间分级.....	53
5.4.1 与原网络模型性能对比.....	53
5.4.2 与经典深度卷积神经网络模型的对比.....	57
5.4.3 消融实验对比.....	59
5.5 12 类别蔬菜分级与分类.....	61
5.6 本章小结.....	62
6 总结与展望	63
6.1 总结.....	63
6.2 展望.....	64
参考文献	65
攻读硕士学位期间从事的科研工作及取得的成果	70
致谢	72

1 引言

1.1 研究背景和意义

随着食品质量标准的提高和消费者对健康饮食关注度的提高,高原夏菜作为一种独特的农产品变得越来越突出。高原夏菜是利用西北高原夏季凉爽、日照充足、昼夜温差大等气候特点,在高海拔地区生产的优质蔬菜。从每年5月下旬到10月初持续供应,很好地弥补了东南沿海夏季蔬菜供应的缺口。它们的种植不仅解决了对反季节蔬菜的需求,还缓解了通常与玉米种植相关的市场危机。因此,这些蔬菜的品种和贸易量都在增加。种植高原夏菜已成为高海拔地区农业产业多样化和提高农业效率的关键战略^[1],能大大增加农民收入。在甘肃省,高原夏菜产业被视为重要的支柱产业,旨在提供高品质农产品的同时促进农民收入增长和乡村振兴的战略目标。相比于其它季节,由于在夏季生长,良好的阳光照射和清新的空气使其具有较高的营养价值。同时,夏菜的口感受生长环境的限制和高山地区的特殊性,相比于其它地方的夏菜更具有清新的口感、特殊的风味和更高的抗氧化物质含量或其他特殊成分。

除此之外,高原夏菜的优质稳定供应对解决东南部城市蔬菜供应短缺问题起着至关重要的作用。高原夏菜的种植在甘肃省范围内得到了大规模发展,向市场提供了各种优质产品,这满足了消费者对新鲜、健康和环保蔬菜日益增长的需求,不仅提高了农业效率,还促进了高原地区的可持续农业实践和经济繁荣。为进一步壮大高原夏菜产业,各级政府部门还加大了政策支持力度,包括财政援助、土地资源优先保障、科技创新扶持等,这些政策和措施为高原夏菜产业的快速发展提供了坚实的支持与保障。

近年来,某些地区调整并加强了高原夏菜品种的选择,以适应市场需求,这包括根据消费者偏好和不断变化的需求引入和推广新品种,重点培育优质品种。这些举措旨在满足消费者对各种蔬菜产品的需求,同时确保高原夏菜品种具有市场销售潜力和竞争力^[2]。虽然在规范这些蔬菜的生产过程方面取得了一些进展,但其销售过程的标准化还存在滞后问题。蔬菜行业的许多公司通常在没有标准化程序的情况下使用冷藏和运输蔬菜。按照植物学的分类系统,蔬菜可以被归类为

不同的植物种类。而蔬菜的分级是将同一类别分成不同级别，通常是根据外观、大小、质量等因素进行的。在商业上，蔬菜通常会被分为不同的等级，例如特级、一级、二级等，以便于消费者选择和销售。目前，在蔬菜产品分级方面明显缺乏意识，传统的手工分拣方法占主导地位，这种传统的分类在很大程度上基于经验，容易产生主观性，导致分类不一致和不可靠。这一差距使得蔬菜的品质与其价格不相符，进而负面影响了农民的收益。为了应对这些挑战，基于计算机视觉的自动分级方法可以显著提高高原夏菜分级过程的客观性、准确性和一致性，从而优化其市场价值和农民的盈利能力。

随着深度学习和数据驱动方法的兴起，计算机视觉技术的质量和性能也在不断提高，被普遍地应用在蔬菜水果的分类方面。相较于人工分类，计算机视觉可以自动识别和分类大量的图像数据，避免了人为主观误差的影响，提供更可靠的分类结果^[3]。

在水果和蔬菜分类领域，利用机器学习和深度学习技术，计算机视觉能够实现精确的图像分类，自动识别和分类大量图像，从而避免与手动分类相关的不准确和偏见，提供更可靠的分类结果。本文提出了一种高原夏菜质量分级方法，旨在减少人力、物力和成本消耗，并且提高蔬菜分级的性能和速度。特别是 ECA-DS-ResNet50 模型的多尺度融合，有助于快速、高效地进行蔬菜质量分级，丰富了农产品分级领域的理论和应用，并为设计制造蔬菜自动识别和分级流水线提供重要的研究成果，对简化和提高农业生产和加工效率具有重大前景。

1.2 研究现状

自 20 世纪 60 年代以来，数字图像处理已成为一个重要的研究领域，计算能力和算法发展的进步拓宽了其应用范围，包括图像增强、压缩、分割和识别。20 世纪 80 年代，研究人员，特别是国外的研究人员，开始集中关注水果和蔬菜的质量检测和分类，这些研究主要利用计算机视觉和图像处理技术来评估这些产品的外观和质量指标。将特征提取和分类算法应用于水果和蔬菜的图像，可以精确检测和评估其成熟度、疾病存在和物理缺陷。这些技术的发展不仅简化了水果和蔬菜的检测和分类过程，而且在整个种植和销售阶段的质量控制中发挥了关键作用。在中国，蔬菜分类主要依靠人工和基本的机械分类方法，这些方法往往耗时、

劳动密集且效率低下。然而，随着技术的不断进步和学者在蔬菜和水果识别领域进行的广泛研究，这一景象正在迅速变化。这一进展对中国在水果和蔬菜识别领域具有重大潜力，将推动水果和蔬菜行业的智能化发展。

蔬菜图像的分类技术流程涵盖了输入图像、图像预处理、特征提取，以及最终的图像分类与识别等环节。在这一领域，蔬菜识别技术经历了两个不同的研究阶段。最初，研究集中在基于传统图像处理方法的蔬菜图像识别上^[4-8]，这些方法采用了经典的图像处理技术，通过颜色空间转换、边缘检测和形态分析等手段进行蔬菜图像的预处理和特征提取，接着采用支持向量机（SVM）和随机森林等机器学习算法进行分类识别，进一步的研究阶段专注于深度学习^[9-11]基础上的图像识别探索。深度学习通过使用卷积神经网络和循环神经网络等高级神经网络模型，在处理广泛数据集时能自动学习到更复杂的特征，从而提高了蔬菜图像分类的精确度和鲁棒性。深度学习在蔬菜图像分类方面带来的进步是显著的，标志着该领域的实质性飞跃。

1.2.1 基于机器学习的蔬菜图像分类

用于分类和识别蔬菜图像的传统图像处理方法涉及传统图像处理技术和机器学习算法的组合。这种方法从蔬菜图像的预处理开始，随后通过特征提取来辨别这些图像的代表性特征，然后利用机器学习算法对这些提取的特征进行分类和识别。通常，蔬菜是根据颜色、质地和几何特征等一系列属性进行分类和分级的。

Bolle^[12]等人在 1996 年开发出一套创新系统，它能自动从果蔬图像中提取颜色、形状、纹理和密度等特征，这些特征准确地捕捉了图像的颜色分布、轮廓、纹理模式和像素密度利用这些特征，训练了一个最小邻居分类器模型，显著提高了水果和蔬菜识别的准确性和稳健性。同时，刘禾^[13]设计了一个基于神经网络专家系统的水果形状自判别系统，识别苹果形状的准确率超 80%。2010 年，Arivazhagan^[14]等人利用纹理和颜色特征的自相关矩阵，采用最小近邻分类器进行有效的水果分类和识别。Rocha^[15]等人解决了特征融合的挑战，提出了一种简化训练过程并减少训练所需数据量的方法，通过集成各种特征和分类器，为特征融合和数据需求挑战提供了可行的解决方案。同年，Huo^[16]及其团队展示了一种基于图像纹理分析的蔬菜识别和分类技术，他们的实验包括使用灰度共生矩阵

(GLCM)算法从黄瓜、皱叶莴苣和普通白菜中提取纹理特征,平均识别率为 85%。张亚静^[17]等人开发专门用于番茄图像采集的机器视觉系统,应用机器视觉技术分析番茄质量的定性和定量方面,结合 BP 神经网络评估番茄表面的颜色和纹理。他们的研究表明,机器视觉方法在预测番茄生长阶段方面是可行的,为质量预测提供了一种新的方法,并为农业生产监测和管理提供了有价值的见解。

2012 年,东芝集团^[18]推出了一款配备摄像头的创新超市扫描仪,用于过滤农产品典型的复杂环境,利用模式识别软件识别这些产品的各种类型。扫描仪展示了快速准确识别不同农产品的能力,显著提高了超市收银台操作和库存管理的效率,标志着农产品识别和分类效率的大幅提高,从而提高了超市的运营效率。Danti 等人^[19]通过裁剪和调整大小对图像进行预处理,从而启动了特征向量的形成。然后,他们从 HSV 图像中提取色调和饱和度通道的平均值和范围用作特征,利用 BPNN 分类器,他们成功地对 10 种叶菜进行了分类,取得了 96.40% 的分类成功率。2015 年,姜伟^[20]结合了 K 聚类算法、HSV 颜色直方图和 Unser 的纹理描述符,开发了蔬菜识别系统,该系统能够准确识别包括菜花、包菜、土豆和香菇在内的蔬菜,识别率达到 100%。这种组合方法显著增强了系统准确区分各种蔬菜的能力,从而产生了高度准确的识别结果。同年,Zhang Y D^[21]融合了前馈神经网络和混沌人工蜂群算法,有效提取了果蔬的颜色、纹理和形状特征,通过混合分类器实现了 18 种水果图像的高效识别。

2018 年,Wan^[22]等人提出了一种判定番茄成熟度的新策略,该策略将颜色特征值与反向传播神经网络相结合。该方法首先将番茄表面最大的内切圆识别为颜色特征提取区域,然后将该区域分割为五个子区域,每个子区域由不同的同心圆半径定义。利用三层 BP 神经网络对两个不同番茄品种的成熟度进行了分类,实现了高达 99.31% 的准确率。2020 年,芦范^[23]开发了径向基函数神经网络的一种优化策略,通过集成量子遗传算法,优化了神经网络模型,显著提升了识别蔬菜图像形状、纹理和颜色特征的能力,表现超越了其他算法。

总之,依赖于传统图像处理的蔬菜识别方法通常涉及人工定义各种属性,如颜色、纹理和形状,以对水果、蔬菜和其他农产品进行分类和检测。然而,这些方法,特别是那些描述纹理特征(如灰度共生矩阵)的方法,往往需要大量的存储空间,并需要较长的计算时间。因此,这对在大规模图像数据集中实时识别蔬

菜提出了挑战。此外，传统的分类器包括最小近邻、支持向量机、决策树和神经网络等需要提取和训练大量特征以实现最佳性能。这一过程显著增加了特征提取和训练中涉及的计算复杂性，对系统设计造成了一定的限制。因此，尽管传统的图像处理技术在蔬菜识别领域发挥了重要作用，但它们并非没有缺点，特别是在存储空间、计算时间和整体系统性能方面。这些局限性凸显了在这一领域继续创新和探索更有效方法的必要性。

1.2.2 基于深度学习的蔬菜图像分类

人工智能技术的发展带来了深度学习卷积神经网络的广泛应用^[24-28]。2013年至2015年期间尤为引人注目，细胞神经网络在图像识别和分类任务方面取得了重大成功，从而引起了广泛关注。这一时期见证了如VGGNet^[29]、GoogLeNet^[30]和ResNet^[31]等卷积神经网络模型的适应与增强，2012年，Alex Krizhevsky及其团队推出的AlexNet模型在ImageNet挑战中取得了重大突破，大大降低了图像分类的错误率，并引发了学术界和工业界对CNN的兴趣激增。2017年，彭俊^[32]提出了一种基于卷积神经网络的蔬菜及水果品种识别方法。通过在LeNet-5模型中添加采样和池化层，他将这种自适应模型应用于蔬菜和水果的图像分类。这种基于CNN的分类模型在九种不同类型的蔬菜和水果中实现了平均识别率约95%。2018年，马越^[33]设计了双边区域推荐网络结构，该结构在水果和蔬菜检测中被证明是高效的。这种结构提高了识别候选区域位置的精度，从而提高了水果和蔬菜的整体检测精度。此外，马越使用MobileNet网络对神经网络模型进行了显著压缩，使其大小仅为原始模型的十分之一。这种优化大大提高了模型的适用性，使其更适用于各种设备和场景。韩朋朋^[34]设计了一种旨在提高蔬菜识别准确性和减小模型大小的新型深度学习网络结构。这种结构利用较小的3x3卷积核来代替传统的较大卷积核，并最大限度地减少了池化层的使用，背后的原理是通过小卷积核的分层堆叠来提取更丰富的特征表示，并减少池化层的特征损失，从而提高蔬菜识别的准确性，这一方法还有助于减少模型参数数量。

2019年，赖佩霞^[35]进行了一项研究，重点是区分细粒蔬菜和粗粒蔬菜。她还介绍了一种新的神经网络newNet和受跨层转换概念启发的网络LSNet，以准

确识别各种蔬菜。此外，还应用了迁移学习技术，并开发了一种识别粗粒蔬菜集的方法，对小样本量特别有效。这些模型在实验中表现出出色的泛化能力和稳健性，有效地对不同类型的蔬菜进行了分类。2020年，Raikar^[36]及其同事对秋葵手指的分类和分级进行了比较研究。他们评估了三个模型——AlexNet、GoogLeNet和ResNet50——并为网络训练实施了迁移学习。研究结果显示，尽管ResNet50需要最长的训练时间，但其准确性显著超过了其他两个模型，这表明ResNet50在秋葵手指的分类和分级方面更为熟练。2021年，许学斌^[37]等人提出了一种称为LBP-PSO-KELM的分类方法，旨在应对蔬菜和水果实时、高精度分类的挑战。在Fruits-360数据集上进行测试，该方法不仅减少了模型的计算时间，而且提高了模型的准确性，达到了98.8658%的显著分类准确率。Jayakumari^[38]和团队开发了一种基于深度卷积神经网络的方法，用于从3D角度对作物进行语义分割。这种方法实现了点云中作物和植物对象的高效特征学习和基于对象的分割。实验结果表明，在识别茄子和卷心菜的植物级作物方面，准确率约为90%，在作物识别方面具有很高的效率。王豪男^[39]提出了对SSD网络的改进，主要涉及用级联双模网络取代传统的特征提取网络，并引入深度密集卷积神经网络。这一增强实现了本地化和分类任务的双重输出，改进后的模型提高了识别水果和蔬菜的速度和准确性。

此外，Li^[40]等人利用ResNet18、ResNet34和VGG16等模型，添加残差块，专注于果蔬病害的识别和分类研究。实验结果表明，与分层较多的ResNet34和VGG16模型相比，具有较少网络层的ResNet18模型表现出优越的识别能力。

尽管基于深度学习的研究取得了鼓舞人心的成果，但实验的执行需要依赖高质量的环境条件。上述算法模型是为高配置服务器量身定制的，并且需要大量的硬件资源。然而，为较小、功能较弱的计算机设计一个优化的蔬菜识别分级系统将大大拓宽这些模型的范围。这将标志着深度学习的图像分类朝着实现智能农业愿景迈出实质性的一步，使先进的分析能力更容易获得，更适用于更广泛的农业应用。

1.3 研究内容及创新点

1.3.1 研究内容

本论文将采用深度学习对高原夏菜质量分级问题进行研究,设计出基于深度学习的高原夏菜质量分级模型,对高原夏菜图像分类的研究工作主要体现在以下几个方面:

(1) 针对目前缺乏的高原夏菜数据集这一情况,设计一套实验室环境下的图像采集装置,通过采集结球甘蓝、娃娃菜、菜花和西兰花四种高原夏菜的分类、分级图像数据,并经过预处理和标注,构建蔬菜分类、分级标准数据集。每一种高原夏菜分为特级、一级、二级,每种蔬菜每一个级别接近 200 张数据集,一共 2400 张高原夏菜图片。

(2) 对原始数据集进行扩充,使其更有利于后续卷积神经网络模型的训练。首先,数据集扩增,对每一类别的数据集进行重新拍摄扩增数据集,扩增拍摄后每种蔬菜每一个级别接近 360 张数据集,一共 4320 张高原夏菜图片。其次,对图像进行粗分割,排除背景干扰,实现样本增强,再使用数据增强的方法对原始图像进行处理,包括翻转、旋转、噪声等多种操作,最终扩充后的数据集达到了 5000 张。

(3) 提出了针对高原夏菜质量分级的多尺度融合 ECA-DS-ResNet50 模型。这种融合模型结合了多种技术,在 ResNet50 和 MobileNetV2 的基础上进行了改进和融合。该网络模型由三个主要组成部分构成: ECA-DS-ResNet-Extractor、MobileNet-Extractor 和 ASPP Module。

第一部分是 ECA-DS-ResNet-Extractor,以 ResNet50 作为主干网络,通过引入 ECA (Efficient Channel Attention) 注意力机制进一步增强了对图像特征的抽取能力。针对 ResNet50 中 7x7 卷积核和池化层的计算成本较高的特点,将其 3x3 标准卷积替换成深度可分离卷积。

第二部分是 MobileNet-Extractor,整合 MobileNetV2 的 DW (Depthwise Convolution) 和 PW (Pointwise Convolution) 卷积机制取代 ResNet50 的常规卷积,有效降低了计算量,同时在每层卷积后引入池化层,进一步减少参数量,优化网络结构,大幅减少了时间和空间复杂度,有效提升模型的运算速度。其中 MobileNet-Extractor 的最后一个 Bottleneck 后有一个 1x1 的卷积,用于维度扩展。

第三部分是多尺度特征提取的 ASPP Module,图像分别经过 ECA-DS-ResNet-Extractor 和 MobileNet-Extractor 模块进行特征提取,两个模块提取的特征图进行

逐元素相加，合并后的特征图送入 ASPP Module，通过不同尺度的感受野来获取图像的多尺度信息，从而提升图像分类的准确性和鲁棒性。

最后，经过一个全连接层的分类器，得到最终的分类结果。这种融合模型综合利用了 ResNet50 和 MobileNetV2 各自的优势，通过特征提取、信息融合和多尺度感知，显著提升了图像分类任务的性能和效率。

1.3.2 研究框架

本文共有六个章节，各个章节的安排如下：

第一章引言：第一部分详细阐述了甘肃省高原夏菜分级的研究背景及对其进行质量分级的意义；第二部分主要介绍了国内外关于机器学习与深度学习的蔬菜图像分类研究现状，引出课题内容；第三部分阐述了本文的研究内容及论文创新点。

第二章相关理论基础介绍：首先主要介绍了蔬菜图像分类分级时用到的理论基础，包括卷积神经网络的基本概念、ResNet 和 MobileNetV2 的理论知识、注意力机制。

第三章高原夏菜数据集构建及其预处理：实现图像采集及其预处理工作，对采集的图片进行裁剪，扩充等操作，达到了数据增强的效果，同时还介绍蔬菜图像的分级原则。

第四章基于深度学习的高原夏菜分级模型构建及优化：章节中依次介绍了高原夏菜图像分级的改进后神经网络的各组成部分及其实现。

第五章实验结果与分析：将改进的模块在高原夏菜数据集上进行训练，通过准确率、参数量、inference time (median)中值推理时间和每秒帧数 (FPS) 以及 F_1 值来验证改进模块的有效性和时间空间复杂度，随后通过消融实验，验证改进模型相较于原网络有所提升。

第六章总结本文工作并提出展望：总结提出的高原夏菜分级方法并提出下一步展望。

1.3.3 创新点

(1) 对于 ResNet50 中 7x7 卷积核和池化层的计算成本较高的特点, 将其 3x3 标准卷积替换成深度可分离卷积, 整合 MobileNetV2 的 DW (Depthwise Convolution) 和 PW (Pointwise Convolution) 卷积机制取代 ResNet50 的常规卷积, 有效降低了计算量, 同时在每层卷积后引入池化层, 进一步减少参数量, 优化网络结构, 大幅减少了时间和空间复杂度, 有效提升模型的运算速度。

(2) 通过两个模块提取的特征图进行逐元素相加, 合并后的特征图送入 ASPP Module, 在本文中仅使用 ASPP 模块网络结构中的中间部分, 通过不同尺度的感受野来获取图像的多尺度信息, 从而提升图像分类的准确性和鲁棒性。这种融合模型综合利用了 ResNet50 和 MobileNetV2 各自的优势, 通过特征提取、信息融合和多尺度感知, 显著提升了图像分类任务的性能和效率。

1.4 本章小结

本章全面概述了研究背景、研究意义、研究现状以及研究的主要内容。首先阐述了蔬菜图像分类研究的背景和重要性, 并指出了使这项研究成为一项重要工作的原因。在此之后, 对先前在蔬菜图像分类领域的研究进行了综合概述, 特别是侧重于涉及机器学习和深度学习的方法。本综述的目的是为后续阶段奠定基础框架和起点, 包括蔬菜数据集的构建和模型的优化。这一综述有助于引导研究人员走向新的研究思路和方法。接着介绍了论文的主要研究内容、结构和创新之处。这包括对本文的研究方向和创新要素的总结。总的来说, 本章为读者提供了一个信息指南和文章内容的清晰概述, 有助于理解研究的核心概念、思想和组织结构, 从而引导读者了解研究的目标和方法。

2 相关理论基础介绍

深度学习是蔬菜图像分类和分级领域最流行的方法之一。本章深入探讨了蔬菜图像分类和分级的理论基础，包括对深度学习和卷积神经网络基本概念、注意力机制模块、ResNet 和 MobileNetV2、多尺度特征提取技术和 PyTorch 深度学习框架。这一理论阐述旨在深入了解所提出的蔬菜质量分级模型的关键技术和框架。

深度学习^{[41][42]}是机器学习的一个子集，其核心是通过构建和训练多层神经网络模型来识别数据的固有表示和特征。它的创立植根于神经网络研究，并受到人脑神经系统运作原理的启发。深度学习通过多个神经元的分层堆叠和互连来执行复杂的计算。这种方法在计算机视觉、自然语言处理和语音识别等领域取得了重大进展，在实现图像分类、对象检测和机器翻译等任务的里程碑方面发挥了关键作用。硬件能力的增强和不断的算法改进推动了深度学习的进展，深度学习有望将其应用扩展到更广泛的领域，从而推动人工智能技术的进步。

2.1 传统神经网络模型

传统人工神经网络，灵感来源于生物的神经系统，旨在复制人脑处理复杂信息的能力。这种网络模仿的生物系统由亿万神经元构成，这些神经元能够响应外界刺激并产生电信号。然后，这些信号被传递到中枢神经系统整合形成判断和命令，接着身体根据这些指令对这些刺激做出反应。基于此理念，引进了人工神经元，即感知机，为人工神经网络铺垫了基础^[43]，结构如图 2.1 所示。

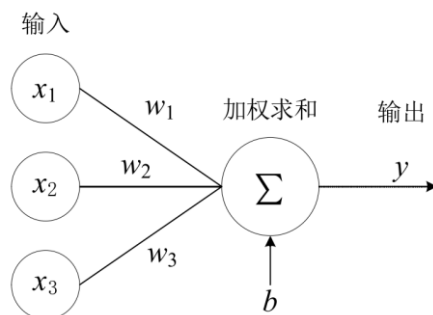


图 2.1 单层感知机模型

其中, x_1, x_2, x_3 为输入信号, 矩阵形式表示为 X , w_1, w_2, w_3 为输入信号所对应的权值, 矩阵形式表示为 W , 与偏移量 b 经过加权求和后, 得到的结果 y 即为该感知机的输出。这个过程表示为:

$$y = w_1x_1 + w_2x_2 + w_3x_3 + b = WX + b \quad (2.1)$$

感知器通过模拟神经元的工作原理, 处理接收到的信号并产生输出。当汇聚多个感知机时, 形成了多层感知机 (Multi-Layer Perception, MLP), 它代表了人工神经网络的基础形态, 是单层网络发展的延伸, 包含多个隐藏层。简单的多层感知机结构如图 2.2 所示:

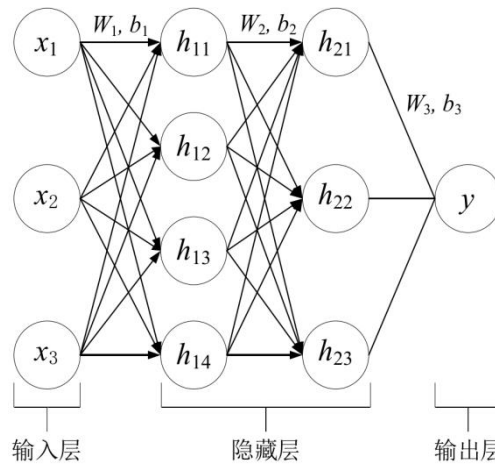


图 2.2 多层感知机模型

该多层感知机包含一个输入层 X , 两个隐藏层 H_1 与 H_2 , 以及一个输出层 Y 。其中, W_1 和 b_1 为隐藏层 H_1 的权值矩阵和偏移量, W_2 和 b_2 为隐藏层 H_2 的权值矩阵和偏移量, W_3 和 b_3 为输出层 Y 的权值矩阵和偏移量, 于是, 该多层感知机的输出 y 可以表示为:

$$y = W_3H_2 + b_3 = W_3(W_2H_1 + b_1) + b_3 = W_3(W_2(W_1X + b_1) + b_2) + b_3 \quad (2.2)$$

尽管增加隐藏层, 模型输出与输入仍保持线性关系, 揭示了仅增加层数并不足以实现有效的多层网络。因此, 引入了非线性变换, 即在每个神经元输出上添加非线性函数 (如 sigmoid、softmax 等), 改进了模型结构, 方程如下 (2.3):

$$\begin{aligned}H_1 &= h(W_1X + b_1) \\H_2 &= g(W_2H_1 + b_2) \\y &= f(W_3H_2 + b_3)\end{aligned}\tag{2.3}$$

其中， h, g, f 为激活函数。这样，输入在通过隐藏层时会进行非线性变换，输出成为对输入进行的一系列非线性变换的结果。

这样的多层神经网络结构更善于拟合和求解非线性问题，增强了其在复杂计算任务中的有效性。

2.2 卷积神经网络模型

卷积神经网络（Convolutional Neural Network，简称 CNN），是计算机视觉和图像处理中广泛使用的一类算法。它通过模仿生物视觉系统中的神经元连接方式，具有良好的特征提取和图像分类能力。作为深度网络的一种，CNN 通过其卷积结构减少了深层网络的内存和参数数量，帮助缓解过拟合问题，包括输入层、特征提取层和分类器等组件。网络的输入层负责将图像传输到系统中，通常特征提取器被集成到多层卷积神经网络中，在这样的网络中，较低的卷积层通常专注于提取局部特征，而较高的层能够学习更复杂的特征表示。激活层配备了非线性函数，显著提高了网络的表达能力，增强了模型对不同输入数据的适应性。在池化层中，诸如最大池化和平均池化之类的流行操作起着关键作用，有效地减小了特征图的大小，同时保留了最关键的特征信息。这一过程不仅提高了模型的计算效率，而且增强了其泛化能力。分类器是卷积神经网络的基本组件，用于将提取的特征映射到特定的类别标签，促进图像分类预测的任务。这一要素是确保基于图像固有特征对图像进行准确分类的关键^[44]。

图 2.3 所示的卷积神经网络，举例说明了一种广泛使用的深度学习模型，该模型能够处理各种类型的图像数据，包括垃圾图像、面部图像和自然场景图像。该网络采用一系列特征提取操作——卷积、激活和池化——将原始输入图像转换为高维特征图，使得能够从图像中提取各种抽象特征，在这些过程之后，所得到的特征图被输入到分类器中。分类器的作用是评估各种特征，并选择最可能的输出标签作为预测。本文的后续部分将深入探讨卷积神经网络的技术架构和内部工作，全面了解其功能和应用。

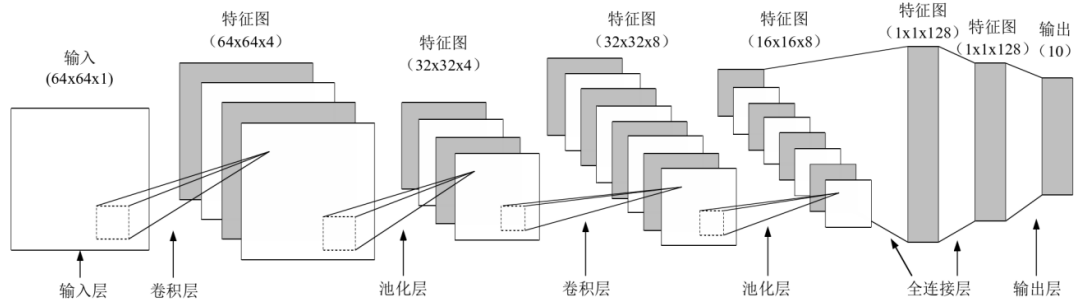


图 2.3 卷积神经网络结构示意图

2.2.1 输入层

输入层作为网络的入口，负责接收并预处理训练和测试数据，这一预处理步骤对提高网络训练效率至关重要。数据预处理中常用的方法包括数据集中和数据标准化等，这些技术确保输入数据具有最佳格式，用于神经网络的后续层的有效处理和分析。

2.2.2 卷积层

卷积层是 CNN 的核心，对图像数据进行处理，通过与卷积核的运算提取图像特征。这些局部运算的结果被保存在它们的相对位置，卷积核以预定序列有条不紊地扫描整个图像，随后产生特征图。在整个过程中，卷积层的输入、输出和任何中间结果都保留了特定的空间结构和相对定位，该属性对于保持图像内空间信息的完整性至关重要。卷积核的设计反映了经典图像处理方法中使用的特征算子。每个卷积层访问并利用这些内核的共享集合，将不同的卷积核应用于图像可以执行各种任务，包括图像去噪、增强和边缘检测。此外，这些内核有助于提取不同的图像特征。卷积层的这种多方面功能突出了它在卷积神经网络中的图像分析过程中的重要性^[45]。图 2.4 展示的是卷积操作中输出特征图像的计算过程。

在图 2.4 所示的二维卷积运算中，若输入特征图大小 $X \in R^{H^a \times W^a}$ ，卷积核 $P \in R^{H^b \times W^b}$ ，则可以得到输出特征图 Y 的计算如下：

$$Y_{i^{a+1}, j^{a+1}} = \sum_{i=0}^{H^b} \sum_{j=0}^{W^b} P_{i,j} x X_{i^{a+1}+i, j^{a+1}+j} \quad (2.4)$$

其中, (i^{a+1}, j^{a+1}) 表示输出特征图上的坐标, 这些坐标按顺序计算, 图 2.4 中输出特征图位置 $(0, 0)$ 的输出结果为 $1 \times 1 + 2 \times 2 + 0 \times 1 + 2 \times 2 = 9$ 。当卷积核设置步长为 1 时, 并且没有对边缘应用填充时, 可以通过顺序计算获得图 2.4 所示的输出图像特征。这个过程有效地展示了输出特征图的每个元素是如何从卷积的应用中导出的。

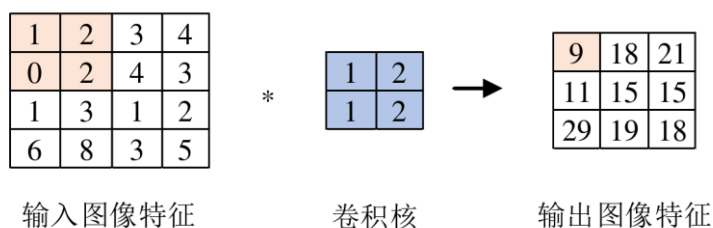


图 2.4 卷积计算

在加权运算之后加入激活函数是多层感知器设计的一个基本方面, 并且对于构建多层神经网络至关重要。这种必要性的产生是因为现实世界生产和生活中遇到的问题往往过于复杂, 无法用简单的线性模型有效地解决。引入非线性激活函数到卷积层不仅提升了模型的表达能力, 也有助于减少过拟合的可能^[46]。这些函数在压缩输出方面是有效的, 但它们有导致梯度消失问题的趋势, 这可能会阻碍网络的学习过程。然而, 随着该领域理论和实验知识的扩展, 研究人员发现, 非饱和函数包括 ReLU 函数, 以及通过它变换得到的 Leaky ReLU 函数等, 提供了优越的激活性能。这些函数随后成为卷积神经网络中激活的首选, 因为它们能够促进更有效的训练和更好的泛化。

激活函数, 如 sigmoid 和 tanh, 分别用于二分类和输出值在 -1 到 1 之间的场景, 数学定义分别在公式 2.5 和 2.6 中详细说明。由于这些特性, tanh 经常被用于网络的隐藏层。虽然 tanh 函数通常被认为是对 sigmoid 函数的改进, 但这两个函数都涉及指数计算, 导致计算复杂度更高。此外, 当使用这两种函数中的任何一种时, 神经元的饱和特性都可能导致梯度消失现象, 这一挑战可能会阻碍神经网络的学习过程。在为深度学习模型选择激活函数时, 这种限制是需要考虑的一个关键因素^[47]。

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2.5)$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.6)$$

在深度学习领域，ReLU 函数是一种常用的不饱和函数，主要是因为它能有效地解决训练过程中遇到的梯度消失现象。ReLU 函数特别擅长保持梯度而不衰减，这对于神经网络的有效学习至关重要，除了减少梯度问题外，ReLU 函数还具有很高的计算效率，与上述饱和函数（sigmoid 和 tanh）相比，在实际应用中收敛得更快。这种增强的效率和更快的收敛速度使 ReLU 成为各种深度学习场景中的首选。ReLU 函数的定义如下：

$$f(x) = \max(0, x) \quad (2.7)$$

虽然 ReLU 函数在缓解过拟合问题方面是有效的，但它有一个局限性：当梯度不存在时，它可能导致无法更新网络的权重。这就是 Leaky ReLU^[48]功能发挥作用的地方，解决了 ReLU 功能的一个关键缺陷——神经元的潜在不活动。Leaky ReLU 函数的数学表达式详见方程 2.8，通常函数在非活动阶段的 α 默认值为 0.01。ReLU 函数的这种轻微修改确保了网络中的所有神经元都有机会进行调整和学习，使其成为深度学习架构中更可靠、更高效的激活函数。

$$\sigma(x) = \max(\alpha x, x) \quad (2.8)$$

2.2.3 池化层

池化层在 CNN 中通过聚合特征图的特征来简化图像表示，采用最大池化和平均池化^[49]等方法降低图像分辨率，减少参数数量和计算复杂度。如图 2.5 所示，一种是最大值池化（max-pooling），即计算出池化窗口内各元素的最大值作为输出结果，池化过程见图 2.5 a)；另一种是平均值池化（mean-pooling），即计算出池化窗口内各元素的平均值作为输出结果，池化过程见图 2.5 b)。

一般来说，池化过程会将卷积后特征图的宽度与高度缩减一半，即特征图的尺寸减少了 3/4，这些技术对于减轻从图像数据中提取特征可能产生的误差至关重要。潜在误差通常可归因于两个主要来源，卷积层中的参数误差会导致估计平

均值的偏移和池化操作中邻域大小的限制会增加估计值的方差。这两个因素都会影响细胞神经网络中特征提取的准确性和效率，因此在这些网络中精心设计池化层很重要。

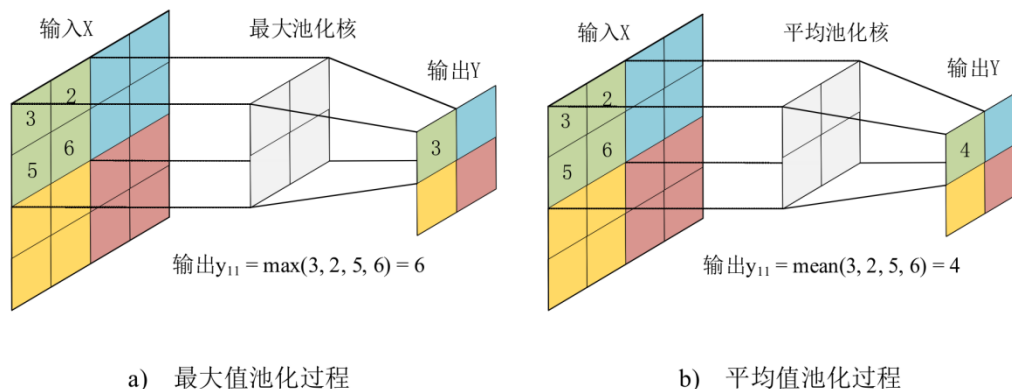


图 2.5 两种池化方式

卷积神经网络中的池化层主要用于降低特征图的维数、消除冗余信息和减少参数的数量，该功能对于减少网络内的过拟合至关重要。此外，池化层减小了特征图的大小，聚合了不同区域的特征，这种特征的收敛增强了模型从不同输入中提取和概括的能力。池操作的另一个显著优点是它对平移不变性的贡献，该属性使模型能够更多地关注特定特征的存在，而不是它们在输入图像中的确切位置，即使图像经历了微小的位移或变换，模型也能够一致地识别特征。池化操作的这一特性增强了神经网络的鲁棒性，在提高卷积神经网络的整体性能和可靠性方面发挥着至关重要的作用^[50]。

2.2.4 全连接层

全连接层（Full Connected Layers, FC）通常被视为一种多层感知器。该层的特点是其全面的连通性，其中每个神经元与相邻层中的每个神经元形成连接，同时排除任何非相邻层的连接。在卷积神经网络领域中，FC 层通常位于输出附近，从卷积层接收其输入。它的主要作用是输出一个用作分类器的向量，包括通过卷积和池化操作将图像映射到多维特征空间中。然后，FC 层综合并细化这些高级特征，有效地定位与特定类别最相关的特征，从而促进特征分类。FC 层的复杂

性之一是其神经元之间成对连接的密集网络,这导致了高参数计数,构成了 CNN 的大部分参数^[51]。

2.2.5 输出层

输出层在神经网络中的主要作用是将从完全连接层接收的数据转换为分类概率输出。这种转换通常通过逻辑回归来实现,逻辑回归是一种用于对二元结果的概率进行建模的统计方法。本质上,输出层充当函数转换器,获取完全连接层的输出并将其转化为最终的分类结果。在处理数值数据时,Softmax 函数通常是这种转换的首选,它是一个数学公式,用于将一组值映射到不同类别上的概率分布,产生的概率与输入数字的指数成比例,从而确保输出值的范围在 0 到 1 之间,总和为 1,使其可解释为概率。Softmax 函数的公式如下:

$$S_i = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}} \quad (2.9)$$

其中 x_i 表示输入信号,即输出层第 i 个神经元的输入值; S_i 是第 i 个神经元的输出值, n 为神经元的数量。

2.3 ResNet 的理论基础

2.3.1 残差学习

在研究图像分类的问题时,通过搭建更深的卷积神经网络结构使分类的准确率得到提高,增加深度模型的层数需要更多的训练样本,计算技术的进步,如多 GPU 并行计算,大大缩短了计算时间,提升了网络性能,但是如果不断地增加 CNN 模型的深度不但不能提高 CNN 模型的性能,反而使 CNN 模型出现梯度消失现象^[52]。

针对深度模型训练的挑战,提出了基于残差网络(Residual Networks, ResNet)结构的 ResNet^[53]模型,通过“恒等映射连接”简化了训练过程。这样做的好处是能够减少数据冗余,更容易进行特征学习,而且残差映射比原始映射更容易优

化。残差网络示意图见图 2.6 所示。在残差模块中，在输入层和输出层之间建立直接链路，使网络中新添加的层能够直接从原始输入层学习新特征。

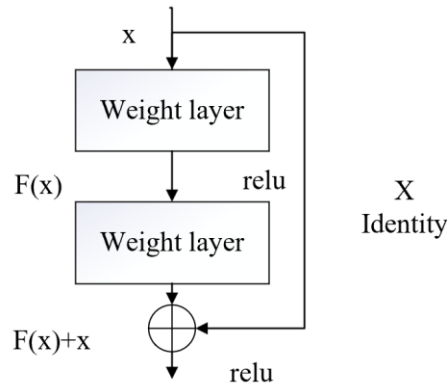


图 2.6 残差网络示意图

在残差块中， x 是模型的输入，残差块可以表示为 $H(x) = F(x) + x$ ， $H(x)$ 为网络输出 X 的期望输出， $F(x)$ 是模型的真实学习部分，当 $F(x) = 0$ 时， $H(x) = x$ 。

在某些场景中，残差块显著增强了网络模型的特征提取能力，而不会引入不利的副作用，这种增强是近年来广泛采用残差网络的一个关键因素。通过集成这些块，残差网络不仅提高了神经网络的整体性能，而且通过减少模型参数的数量来实现这种改进。这种提高效率和降低复杂性的双重优势是残差网络在各个领域流行和广泛应用的主要原因。

$H(x) = F(x) + x$ 函数通过在前馈网络中增加一个的“shortcut Connections”来实现^[54]，残差块的结构图如图 2.7 (a) 所示，捷径以不同的步长跳过一个或多个层与主径汇合，这样结构的输出为：

$$y = F(x, \{W_i\}) + x \quad (2.10)$$

如果输入和输出维度不同，则需要增加一个线性投影，其计算公式如 (2.11) 所示：

$$y = F(x, \{W_i\}) + W_s x \quad (2.11)$$

在训练过程中集成类似于 ResNet 模型的快捷方式可以显著提高性能，使低级错误能够快速传输到后续级别，有效地缓解了梯度消失的问题，这是像 ResNet 这样的深度网络中的一个常见问题，尤其是随着层数的增加。

特别是对于深层网络，引入瓶颈架构以优化训练效率，通过在 3x3 卷积层前后使用 1x1 卷积层来减少输入输出大小。尽管这种瓶颈结构会增加时间复杂性和模型大小，但它显著提高了模型的训练效率。该瓶颈体系结构的细节如图 2.7(b) 所示。

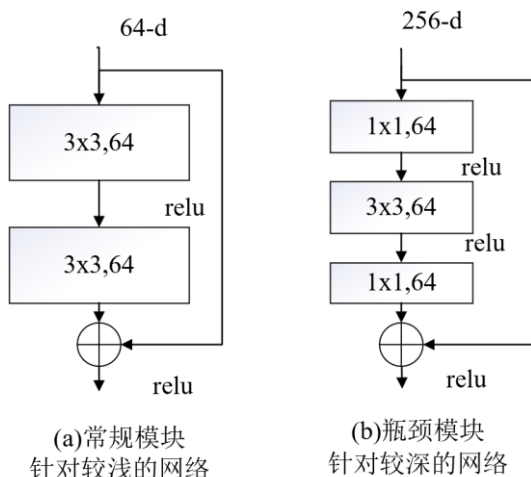


图 2.7 基础残差模块

2.3.2 ResNet 模型结构

ResNet 可以分为 ResNet18 (18 层)、ResNet34 (34 层)、ResNet50 (50 层)、ResNet101 (101 层)、ResNet152 (152 层)。ResNet 各模型的结构参数如表 2.1 所示。以 ResNet18 为例，图 2.8 为 ResNet18 网络结构图。

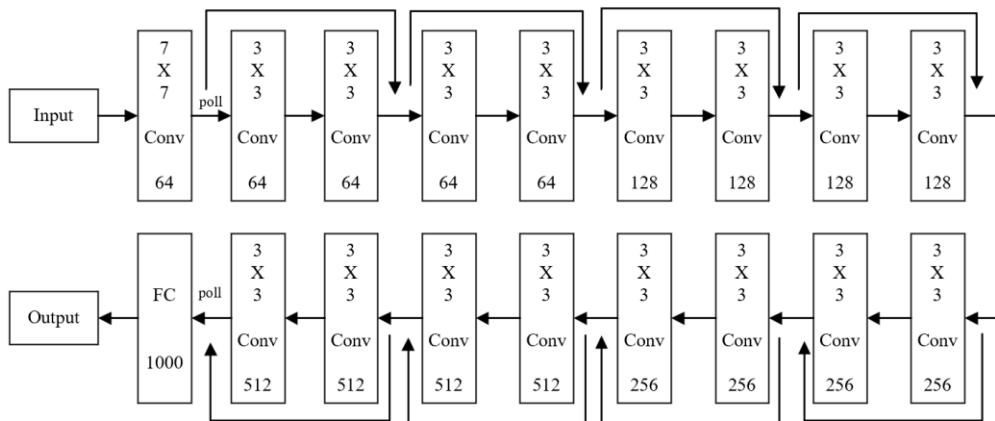


图 2.8 ResNet18 网络结构图

表 2.1 不同层数的 ResNet 网络参数

卷积层名称	输出层大小	ResNet18	ResNet50	ResNet101
Conv1	112x112		7x7,64,S=2	
			3x3 max pool,s=2	
Conv2_x	56x56	$\begin{bmatrix} 3x3, 64 \\ 3x3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 1x1, 64 \\ 3x3, 64 \\ 1x1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1x1, 64 \\ 3x3, 64 \\ 1x1, 256 \end{bmatrix} \times 3$
Conv3_x	28x28	$\begin{bmatrix} 3x3, 128 \\ 3x3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 1x1, 128 \\ 3x3, 128 \\ 1x1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1x1, 128 \\ 3x3, 128 \\ 1x1, 512 \end{bmatrix} \times 4$
Conv4_x	14x14	$\begin{bmatrix} 3x3, 256 \\ 3x3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 1x1, 256 \\ 3x3, 256 \\ 1x1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1x1, 256 \\ 3x3, 256 \\ 1x1, 1024 \end{bmatrix} \times 23$
Conv5_x	7x7	$\begin{bmatrix} 3x3, 512 \\ 3x3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 1x1, 512 \\ 3x3, 512 \\ 1x1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1x1, 512 \\ 3x3, 512 \\ 1x1, 2048 \end{bmatrix} \times 3$
		Average pool,1000-d fc,SoftMax		
FLOPs		1.8x10 ⁹	3.8x10 ⁹	7.6x10 ⁹

2.4 MobileNetV2 的理论基础

2.4.1 深度可分离卷积

MobileNet 由 Andrew G 和谷歌团队^[55]于 2017 年开发，代表了流线型、轻量级深度神经网络构建的范式转变。这种创新架构采用深度可分离卷积作为传统卷积方法的替代方案。MobileNet 提供了根据特定研究需求定制网络的灵活性。它通过两个预设的全局超参数实现了这一点，允许定制设计，以满足移动和嵌入式视觉应用程序的独特要求，同时有效地平衡速度和准确性^[56]。

MobileNet 的架构巧妙地将标准卷积分解为两种类型：深度卷积和点卷积。

这种方法不仅确保了信息的有效流动，而且大大减少了参数的数量和计算负载，在不影响网络性能的情况下提高了计算速度。MobileNet 中使用的滤波器结构如图 2.9 所示，展示了该模型的卷积神经网络设计创新方法。

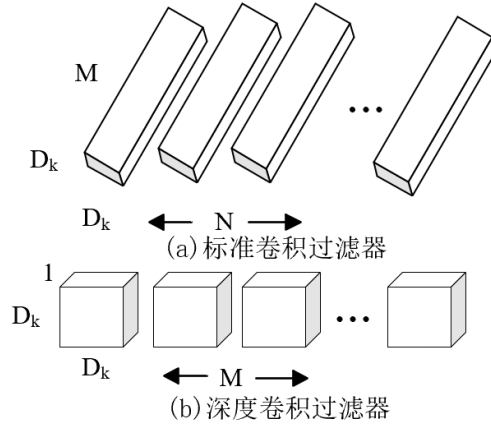


图 2.9 MobileNet 过滤器

假设标准卷积层的输入特征图 F 是具有维数 $D_F \times D_F \times M$ ，使用的标准卷积 K 是 $D_K \times D_K \times M \times N$ ，其中 D_F 代表正方形输入特征图的尺寸， D_K 代表标准卷积核的大小， M 代表输入通道的数量， N 代表输出通道的数量，该卷积核的输出特征图为 $D_G \times D_G \times N$ ，如图 2.9 (a) 所示。

计算标准卷积的输出特征图的公式如下：

$$G_{k,l,m} = \sum_{i,j,m} K_{i,j,m} x F_{k+i-1,l+j-1,m} \tag{2.12}$$

这里标准卷积的计算量为

$$D_K \times D_K \times M \times N \times D_F \times D_F \tag{2.13}$$

标准卷积可拆分为深度卷积和逐点卷积，负责滤波的深度卷积的大小为 $D_K \times D_K \times M$ ，产生的输出特征为 $D_G \times D_G \times M$ ，如图 2.9 (b) 所示。深度可分离卷积的计算公式如下：

$$G_{k,l,m} = \sum_{i,j,m} K_{i,j,m} x F_{k+i-1,l+j-1,m} \tag{2.14}$$

式中 K 是卷积核, K 的第 m_{th} 个卷积核作用在输入特征 F 的第 m_{th} 个通道上, 产生 G 上第 m_{th} 个通道输出, 因此深度可分离卷积的计算量为:

$$D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F \quad (2.15)$$

根据公式 (2.13) 和公式 (2.15), 与标准卷积相比, 深度可分离卷积的计算复杂度降低了, 减少为:

$$\frac{D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F}{D_K \times D_K \times M \times N \times D_F \times D_F} = \frac{1}{D_K^2} + \frac{1}{N} \quad (2.16)$$

MobileNet 采用深度可分离卷积, 彻底改变了特征图的卷积过程^[57]。在这种方法中, 每个卷积核专用于特征图的一部分, 而不是传统的多个核对整个图进行操作的方法。这种巧妙的设计大大减少了模型的计算需求和存储需求, 通常会减少 8 到 9 倍。这样一个高效的框架使 MobileNet 特别适合于计算资源有限的环境, 如移动和嵌入式设备, 同时在处理和分析视觉数据方面保持高性能。

2.4.2 倒残差结构

MobileNet V2 网络代表了原始 MobileNet 架构的高级迭代^[58], 其主要特点是采用了反向残差结构, 这是一个受 ResNet 中残差块启发的概念, 这些剩余区块的架构如图 2.10 所示。随着网络深度的增加会导致更大的训练误差, MobileNet V2 中的残差块是专门为解决这一退化问题而设计的, 其残差函数为 $F(x) = H(x) - x$ 。这种设计有效地避免了梯度消失的问题, 梯度消失通常是由激活函数和链分化过程引起的。此外, 该架构允许根据需要扩展具有附加层的网络, 从而增强其对更复杂任务的适应性和效率。

MobileNet V2 与 ResNet 网络的区别主要在于其特征提取方法。虽然 ResNet 采用标准卷积, 即最初将维度降低到 0.25 倍, 然后进行卷积, 然后再增加维度, ResNet 中的这种方法在视觉上可以被比作沙漏形状。MobileNet V2 采用了不同的策略, 始终使用深度卷积进行特征提取。它首先将维度扩展 6 倍, 然后进行卷积, 然后降低维度。如图 2.11 所示, 当可视化时, 该过程类似于主轴的形状。这种独特的结构差异突出了 MobileNet V2 的创新设计, 优化了它的效率和性能, 特别是在计算资源有限的环境中。

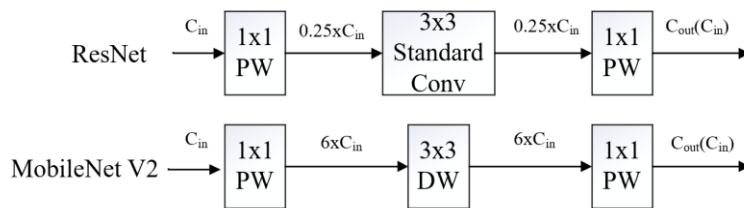


图 2.11 ResNet 和 MobileNet V2 微结构

2.5 注意力机制模块

2.5.1 CBAM 模块

CBAM^[59]注意力机制融合了通道和空间注意力，优化了特征图的处理，如图 2.12 所示，通道注意力在左侧，空间注意力在右侧。与卷积神经网络中主要分析通道域并经常忽略特征图通道之间的相互作用的传统注意力机制不同，CBAM 提供了一种全面的方法。通过分析空间和通道维度的注意力，CBAM 改进了特征图的聚焦效果。空间注意力使神经网络能够更多地集中于对分类至关重要的图像像素区域，这种双向注意力分配通过提高注意力机制的有效性显著提高了模型的性能。

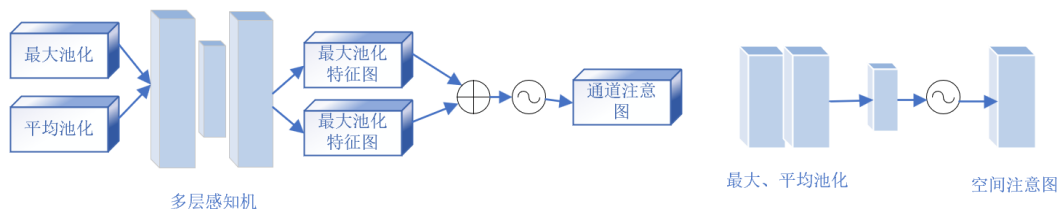


图 2.12 CBAM 注意力机制

在 CBAM 中，通道注意力机制通过全局最大池和平均池操作压缩特征图，然后利用 MLP 网络处理，优化了输入特征的处理。该网络最初使用一个完全连接的层来减少信道数，然后通过另一个完全相连的层来恢复信道数。结合这两种特征图在通道维度上通过 sigmoid 函数对每个通道的权重进行归一化处理，范围设定在 0 到 1 之间。随后，这些调整后的权重与原始特征图进行乘法操作，以此来微调特征图的贡献度。公式为：

$$M_c(F) = \sigma(MLP(AvgPool(F)) = MLP(MaxPool(F))) \quad (2.17)$$

CBAM 模块利用空间注意力机制对经过通道注意力处理后的特征图进行进一步分析, 该机制通过对特征图应用最大池化和平均池化操作开始, 在通道维度对结果进行合并。随后, 使用大小为 7x7、3x3 或 1x1 的卷积核来融合信道信息, 将特征图的维度从 [b, 2, h, w] 改变为 [b, 1, h, w]。卷积输出通过 sigmoid 函数进行归一化, 以导出特征图的空间权重。然后使用该权重来乘以输入特征图, 从而有效地调整其空间特征。这个过程可以通过一个特定的公式来表达:

$$M_s(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \quad (2.18)$$

其中 F 为输入特征, $F \in R^{X \times H \times W}$, σ 表示激活函数, $f^{7 \times 7}$ 代表 7x7 卷积, $AvgPool(F)$ 、 $MaxPool(F)$ 分别为平均池化特征和最大池化特征。

CBAM 流程概述如下: 首先, 特征图通过通道注意力机制进行处理, 与通道权重相乘后, 再经过空间注意力机制处理, 在此过程中, 对空间权重进行归一化并与特征图相乘, 产生最终的加权特征图。

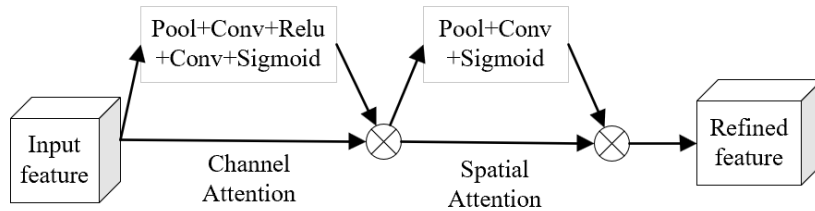


图 2.13 CBAM 模块

2.5.2 SENet 模块

SENet (Squeeze-and-Excitation Networks)^[60]注意力机制是在 2017 年被引入的, 由 Hu 等人提出, 它在 ImageNet 2017 比赛的分类任务中获得了第一名。这种创新方法强调不同维度之间的关系, 特别是将注意力机制纳入渠道维度。通过这样做, SENet 提供了对信道相互关系更细致的理解和处理, 增强了网络的整体分类性能。

SENet 能够熟练地学习每个信道的重要性，从而增强网络辨别和自适应调制各种信道权重的能力。它包含两个关键操作：挤压和激励。挤压运算有效地减少了通道之间的多余信息，而激励运算放大了相关信息，从而增强了模型的性能和泛化能力。这种方法使神经网络能够专注于特定的关键通道，而忽略那些在分类任务中意义较小的通道。这种选择性聚焦类似于人类的视觉感知，即注意力指向特定目标，忽略无关元素。如图 2.14 所示，SENet 的体系结构体现了这两种操作，通过“挤压”操作提取特征图中全局特征，以及“激发”操作帮助分析不同通道间的关系，并确定它们的相对重要性。

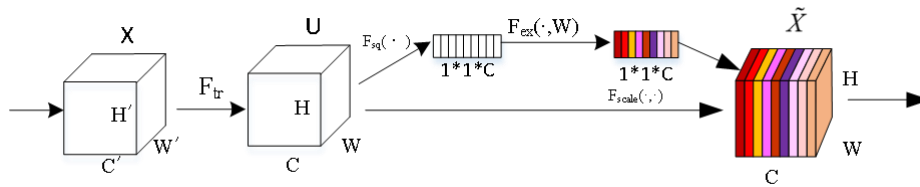


图 2.14 SENet 模块

挤压过程通常把一个通道的所有空间特征压缩成一个全局特征，采用全局平均池化实现这一转换：

$$Z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j), z \in R^c \quad (2.19)$$

其中 u_c 为某张特征图， H, W 为特征图的高和宽， Z_c 为标量值反映特征图的全局特征。

激励操作抓取通道之间的依赖关系，其通过两个全连接层实现：

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \text{ReLU}(W_1 z)) \quad (2.20)$$

其中 W_1, W_2 分别表示第一层和第二层全连接的权重，其中使用不同的激活函数，分别为 ReLU 函数与 Sigmoid 函数。

SENet（挤压和激励网络）的核心概念在于它能够自主确定网络中每个信道的权重。这一过程强调了通道的重要性，放大了具有功能特征的通道，同时抑制了效果较差的通道。尽管这种方法可能会导致某些参数和计算负载的增加，但它给网络带来的性能和效率的明显提高证明了这种权衡是合理的。这种战略加权增

强了网络的整体有效性，尽管有额外的计算需求，但这是一种值得改进的方法。

2.6 本章小结

本章首先深入探讨构成传统和当代卷积神经网络模型基础的五个核心结构。接下来是对 ResNet 和 MobileNet 网络模型和注意力机制的全面探索，阐明它们在深度学习领域的作用和影响，为其在后续讨论中的实际应用奠定了基础。

3 高原夏菜数据集构建及其预处理

3.1 高原夏菜数据集的构建

3.1.1 图像采集

数据构成了深度学习研究的基础,而图像采集及其预处理是图像识别流程中的关键初步步骤。图像样本的质量在开发强大的深度学习模型中起着关键作用。本实验中使用的数据集来自一个高原夏菜蔬菜交易场所,这保证了用于实验目的的数据的真实性和可靠性。为了确保获取高质量数据集,本研究主动在受控的实验室环境中构建了一个专门的计算机图像采集系统,此设置旨在优化图像采集过程,从而提高数据集的整体质量和有效性。

手机配备了内置闪光灯,提供了稳定的光源,并以清晰和高分辨率捕捉图像,非常适合满足模型训练的数据集要求。因此,在本实验中,选择手机作为主要图像采集设备,采集设备还包括计算机和拍照支架等,如图 3.1 所示。图像采集过程主要分为两个步骤:首先,用手机拍摄高原夏菜,闪光灯开关控制照明条件;紧接着,所采集的图像通过数据线传输至计算机,进而进行高原夏菜数据集的整理与深度学习模型的训练。

所采集的高原夏菜数据集为以下四种:结球甘蓝、娃娃菜、菜花、西兰花,如图 3.2 所示,这些蔬菜被进一步分为三个质量等级:特级、一级、二级,共 12 个类别,对于每一类,每种类别蔬菜先拍摄 200 张,共采集的高原夏菜数据集为 2400 张图像。

为了增强数据集的稳健性,随后对其进行了扩增,为每个类别进行了进一步的扩增拍摄,重新拍摄后每一个类别由原来的 200 添加到了 360 张图像,共计采集到的高原夏菜数据集扩增为 4320 张,为分析和模型训练提供了更广泛和多样化的集合。

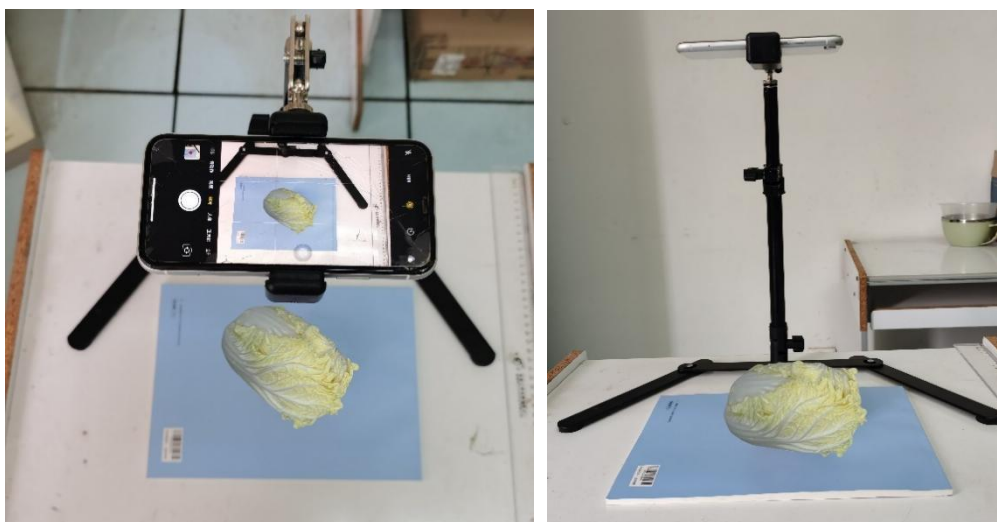


图 3.1 高原夏菜图像采集装置

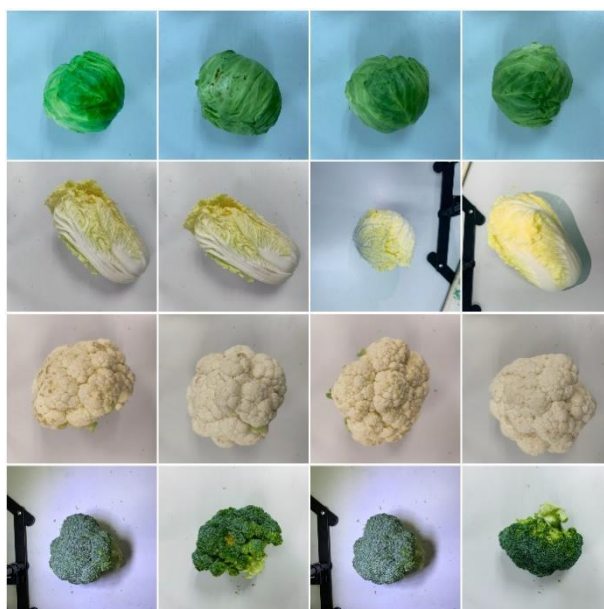


图 3.2 部分初始数据集

3.1.2 数据集存在问题

通过上述方法获取的高原夏菜数据集会出现以下几个问题：

(1) 背景干扰：如图 3.2 所示，根据结球甘蓝和娃娃菜的图像中所观察到的，不同蔬菜品种的背景缺乏一致性，这种不一致性可能会在特征提取过程中引入不需要的噪声，从而可能影响分类的准确性。其次，在进行数据拍摄过程中不可避免地会出现拍摄支架干扰物。

(2) 亮度不一致：由于照明条件或拍摄照片的时间不同，拍摄图像的亮度水平可能会有所不同。光照条件的变化可能会对数据集的视觉统一性产生影响。

(3) 数据集大小有限：考虑到高原夏菜的公开可用数据集的稀缺性，以及该数据集是手动策划的，因此可用的图像数据量受到限制。该数据集的有限大小可能会在模型中造成过度拟合的风险。

为了缓解这些问题并在模型训练中获得最佳结果，对收集的数据集进行预处理至关重要。这涉及到实施策略来标准化背景、标准化亮度水平，并在可能的情况下增强数据集，从而减少特征提取中的干扰可能性，并将过拟合的风险降至最低。

3.2 图像分割和背景移除

为了增强关键特征信息的保留，最大限度地减少干扰，提高识别率，图像的预处理通常需要去噪和分割增强。高原夏菜在拍摄过程中会包含拍摄支架等设备，图像捕获过程中相机支架和其他背景设备的存在会对后续的特征提取过程产生不利影响，从而影响模型的准确性和分类结果。为了解决这个问题，本文利用 Qin 等人提出的 U²-Net^[61]模型从目标特征中去除背景。这种方法有效地减轻了背景造成的干扰，确保它不会对随后的分类结果产生不利影响。

如图 3.3 所示的，U²-Net 架构采用了与 U-Net 相似的设计，充当编码器-解码器模型的角色。该结构的每个阶段包括 U 形剩余单元 (RSU)。该网络巧妙地嵌套和堆叠了两个 U-Net 结构，总共有 11 个 U 形配置。在每个阶段中，模块都被设计用于捕捉多尺度特征，其主要组件包括阶段编码器和显著性图融合模块组。网络的损失函数公式如下：

$$L = \sum_{m=1}^M w_{side}^m l_{side}^m + w_{fuse} l_{fuse} \quad (3.1)$$

其中， M 为网络中层数， l_{side}^m ($M = 6$) 为每个阶段对应输出的显著图， l_{side}^m 为融合模块输出的显著图， w_{side}^m 为对应各个阶段损失函数相关联的权重， w_{fuse} 为最终输出的预测图对应的权重。这些组件在定义网络性能方面发挥着至关重要的作用，每个参数都有助于模型在处理和数据分析时的整体有效性和准确性。

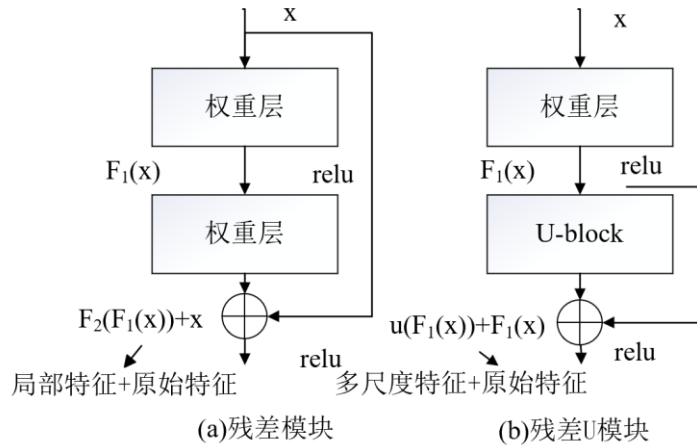


图 3.3 残差模块的网络结构对比

在训练的各个过程中，每阶段的损失计算采用标准的二元交叉熵：

$$l = - \sum_{r,c}^{H,W} [P_{G(r,c)} \log P_{s(r,c)} + (1 - P_{G(r,c)}) \log(1 - P_{s(r,c)})] \quad (3.2)$$

其中， (r,c) 是图像中像素的坐标， (H,W) 是图像的高度和宽度， $P_{G(r,c)}$ 是高原夏菜初始的像素值， $P_{s(r,c)}$ 是已经去除分割的背景之后的图像的像素值。

利用这些检测结果，隔离原始图像中对象的前景，然后去除背景。这一过程确保了整个数据集背景的一致性。该处理的结果如图 3.4 所示，展示了该方法在增强数据集一致性和清晰度以进行更准确的后续分析方面的有效性。



图 3.4 高原夏菜移除背景对比图

3.3 数据增强

在训练网络模型的过程中，尤其是在深度学习领域，受限的训练样本容易引发模型的过拟合问题。为了缓解这个问题，本实验选择增加训练样本，从而增强模型的稳健性。数据扩充是一种通过改变现有数据而不改变相应标签来生成额外训练数据的技术。这种方法在计算机视觉中特别普遍，在计算机视觉中将使用各种图像转换技术来扩展数据集。在本文的图像预处理阶段，使用了以下数据增强方法：

(1) 空间几何变换：涵盖了对原始图像执行随机 90 度旋转、水平翻转和垂直翻转等操作。

(2) 为部分图像样本添加椒盐噪声和高斯噪声：这一措施显著扩充了数据集的规模，提升了训练网络模型的鲁棒性和泛化水平。

接下来详细说明本研究所采用的数据增强策略。

3.3.1 图像的空间几何变换

(1) 旋转操作：通过随机旋转处理，使图像绕原点顺时针旋转 90 度。

(2) 翻转操作：为增强数据多样性，在高原夏菜数据集中随机进行水平和垂直翻转。

每种蔬菜的图像只进行其中一种随机的旋转或翻转几何变换，图 3.5 为原始高原夏菜图像经过几何操作后所生成的新图像。



图 3.5 图像几何变换示例图

3.3.2 图像添加噪声

图像噪声的添加旨在一方面扩大样本量，另一方面提高模型的噪声抵抗力。在理论研究中，高斯噪声较为普遍，视为较难处理的噪声类型，而椒盐噪声则在

视觉上较为突出。模型通过学习这些噪声中的新特征，能够增强其对噪声的抵抗能力，因此本研究主要向原始数据样本中添加这两类噪声。

噪声添加过程包括随机为部分高原夏菜图像引入 0.005 参数设置的椒盐噪声或高斯噪声。需要注意的是，每个图像只接收一种类型的噪声处理，并且并非数据集中的所有图像都要进行噪声添加。这种方法不仅模拟了潜在的现实场景，还用于评估卷积神经网络在噪声条件下的性能，完成后的图像由原来的 4320 张扩增到 5000 张。具体实现如下图 3.6。

其中第一行是处理前的原始图像，第二行是添加高斯噪声后的效果图，第三行是添加椒盐噪声后的效果图。

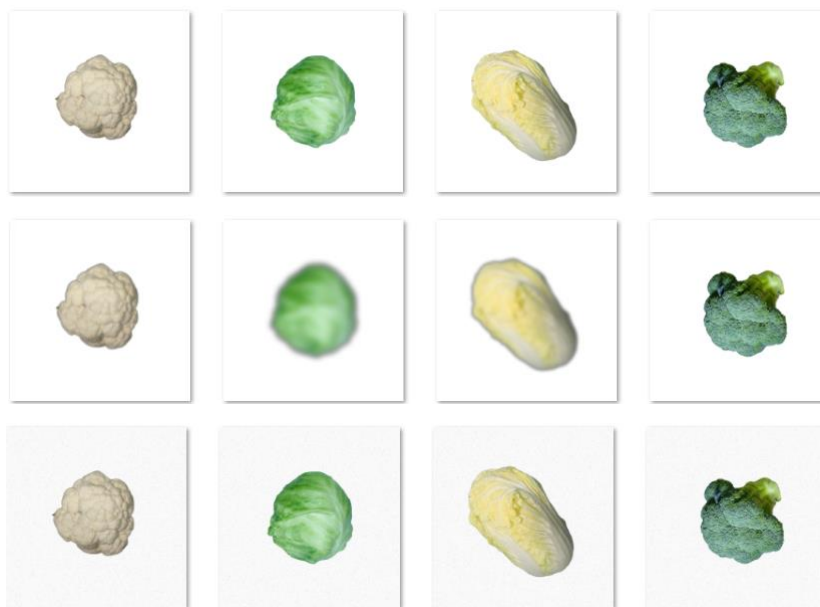


图 3.6 噪声处理后图片

3.4 高原夏菜分级标准

质量分级是根据产品的外观、新鲜度、口感、营养价值等方面的特征，将产品分为不同等级的过程。对蔬菜进行质量分级是为了确保产品的品质和安全，促进市场的有序发展。质量分级通常包括以下几个步骤：

(1) 明确标准：根据产品的特点和市场需求，制定相应的质量分级标准。这些标准可以包括外观、新鲜度、尺寸、形状、颜色、纹理等方面的要求。

(2) 检查样品：从生产的蔬菜中抽取样品，进行外观、口感、新鲜度等方面的检查。检查可以由专业人员或经过培训的工作人员进行。

(3) 评定等级：根据检查结果，将产品分为不同的等级。通常分为一级、二级、三级等。一般来说，一级产品质量最高，三级产品质量最低。

(4) 标注等级：对分级后的产品进行标注，清晰地显示产品的等级。标注可以采用符号、数字或文字等形式进行。

(5) 包装和销售：根据产品等级进行包装，确保产品在运输和销售过程中保持品质。在销售时，清晰地显示产品的等级，方便消费者选择。

2020年，甘肃省质量协会发布了《高原夏菜等级划分标准》(T/GSQA 012-2020)^[62]，对高原夏菜的质量进行了分类，从而建立了评估其等级的权威指南。在我们对高原夏菜质量分级的整个研究过程中，我们严格遵守了这一标准，确保了分级结果的准确性和可靠性。表 3.1 概述了这些分级标准的具体内容。

每种高原夏菜主要考虑蔬菜的大小和外观等因素：它们的紧实性、外表的完整性、表面缺陷、病虫害的损害、机械伤害以及蔬菜的硬度和颜色等。本文重点研究四种高原夏菜的外观、硬度、新鲜度和病虫害的存在。在整理原始和预处理的图像数据后，这些蔬菜被统一分为三个级别：特级产品、一级产品和二级产品。在使用该数据集训练模型之前，必须根据数据各自的等级对数据进行标记，数字表示蔬菜的特定等级：0 表示特级产品，1 表示一级产品，2 表示二级产品，每种高原夏菜的命名惯例都结合了其中文名称、等级和序列号，例如，西兰花的命名格式表示为“西兰花 0 (1) .jpg”。

表 3.1 高原夏菜质量等级 T/GSQA 012-2020

产品	特级产品	一级产品	二级产品
结球甘蓝	叶球规格整齐，外观一致，结球紧实，边缘光滑；没有老帮、焦边、侧发萌芽和机械损伤等，没有病虫害损害	叶球规格基本整齐，外观基本一致，结球较紧实，边缘较为光滑；没有老帮、焦边、侧芽萌发及机械损伤，允许少量虫害损伤等	叶球大小基本整齐，外观相似，结球不够紧实，边缘不够光滑；允许少量焦边、侧芽萌发及机械损伤，允许少量病虫害损伤等
娃娃菜	无虫眼；无焦尾；刀口平外观一致，结球紧实，修整良好；无老帮、焦边、胀裂、侧芽萌发及机械损伤等	无虫眼，允许外叶有少量斑点；无焦尾；刀口平外观基本一致，结球较紧实，修整较好；无老帮、焦边、胀裂、侧芽萌发及机械损伤等	允许外叶有少量虫眼或病斑；允许有焦尾；外观相似，结球不够紧实，修整一般；可有轻微机械损伤等
菜花	花球洁白紧实、形体完整、无褐斑无黄斑，菜球紧凑、菜叶浅绿、新鲜，外叶及茎部去除。	花球洁白紧实、形体完整、无褐斑无黄斑，菜球略有松动、菜叶浅绿。	外观基本一致；花球完好；花球略松散；形体基本完整。
西兰花	外观一致；花球圆整，完好；花球紧实，不松散；色泽浓绿、一致；花蕾细小、紧实，未开放；花茎鲜嫩，分支花茎短；无机械损伤。	外观基本一致；花球较圆整，完好；花球尚紧实，四周略有松散；色泽浓绿、基本一致；花蕾较紧实，但尚未开放；花茎鲜嫩，分枝花茎短；允许有机械损伤，但不明显。	外观基本一致；花球完好；花球略松散；色泽略显黄绿或有少量异色花蕾；花蕾有少量开放；花茎较嫩，分支花茎较长。允许有机械损伤，但不严重

对蔬菜进行质量分级有利于保证产品品质、提高市场透明度、促进产业发展、减少食品浪费等方面，对于整个蔬菜产业的健康发展具有重要意义。

(1) 品质保证：质量分级可以确保消费者购买到的蔬菜符合一定的标准，品质有保证。这有助于消费者放心购买，并提升产品的信誉度。

(2) 提高市场透明度：质量分级可以使市场更加透明，消费者可以清晰了解不同质量等级的蔬菜的特点和价格，从而更好地进行选择。

(3) 促进产品差异化营销：质量分级使得不同质量等级的蔬菜得到了区分，有助于生产者根据产品的质量特点进行差异化营销，提高产品的附加值。

(4) 促进产业发展：质量分级有助于提升整个产业链的效率和竞争力。生产者可以更好地了解市场需求，进行合理的生产计划和管理，提高生产效率和质量。

(5) 标准化生产：通过质量分级，可以对蔬菜的生产、收购、加工和销售等环节进行标准化管理，有利于提高整个产业的规范化水平。

(6) 减少产品浪费：质量分级有助于减少因为质量不佳而导致的食品浪费。消费者可以根据需求和质量等级购买适量的产品，减少因质量不佳而造成的浪费情况。

3.5 本章小结

本章主要关注构建高原夏菜图像数据集的方法，并深入研究图像处理和分级的基本原理。在图像采集阶段，为了保持训练数据的一致性，采用了关于捕获图像的高度和使用的光源标准化方法。这种标准化对于最大限度地减少外部变量对模型训练结果的影响至关重要。同时，为了防止训练过程中的过度拟合，本文结合了数据扩充技术来扩大数据集。此外，U²Net 网络用于背景去除，这有助于提高模型的准确性。高原夏菜的分类遵循甘肃省质量协会制定的团体标准，蔬菜根据大小、质量、颜色和外观等标准进行分类，这种结构化的分类方法确保评分过程符合公认的标准，从而为研究结果提供可信度和准确性。

4 基于深度学习的高原夏菜分级模型构建及优化

如前一章所述，在收集和完善的夏菜数据集后，下一个关键步骤包括将这些图像输入到分级模型中进行分类。本章介绍了专门为评估高原夏菜质量而开发的 ECA-DS-ResNet50 模型。最初，本章深入研究了原始 ResNet50 网络的基本架构，提供了基本的理解；随后，详细阐述改进模型的具体改进方法；最后对本章内容进行总结。

4.1 ResNet50 网络模型结构

ResNet 是深度学习领域广受欢迎的模型，以其在包括图像分类在内的各个领域取得的值得称赞的成就而闻名。它有效地缓解了梯度消失或爆炸的挑战，这在深度网络的训练中很常见，从而简化了复杂架构的学习过程。在这种情况下，我们实验中使用的卷积神经网络模型植根于 ResNet50 的设计理念，具有鲁棒性。该网络不仅具有高识别精度，而且具有卓越的实时性能，特别适合在实际工作条件下执行市场交易中的蔬菜识别任务。

下表描述了 ResNet50 的网络结构。它总共显示了 50 个卷积层，从最初的层延伸到最后的层。这个 50 层的网络被系统地划分为六个主要部分：输入模块、四个顺序块（每个块包括三个残差块）和最终输出模块。该网络模型的基石是残差块结构，每一层都包含 ReLU 激活功能，并辅以批量归一化单元，显著增强了模型的准确拟合能力和高效性能。

(1) 预处理层：通常对输入图像进行一些预处理操作，例如归一化、裁剪或缩放等，以便将图像数据转换为模型可接受的输入格式。

(2) 卷积层：ResNet50 的初始卷积层通常包括一个 7x7 的卷积核，用于对输入图像进行初步的特征提取。通常会跟随一个批量归一化层和 ReLU 激活函数。

(3) 池化层：接下来可能会包含一个最大池化层，用于降低特征图的尺寸和复杂度，以增强模型的平移不变性。

(4) 残差块堆叠：ResNet50 包含了多个残差块的堆叠。每个残差块由若干卷积层和恒等映射（identity mapping）组成，其中恒等映射用于保持特征图的维

度一致。残差块的堆叠使得网络可以学习更复杂的特征表示，同时缓解了梯度消失问题。

(5) 全局平均池化层：在最后一个残差块之后，通常会添加一个全局平均池化层。该层将特征图的每个通道进行平均，将其尺寸降至 1x1，并保留每个通道的特征。

(6) 全连接层：最后，一个全连接层用于将全局平均池化层的输出映射到最终的类别预测。在 ResNet50 中，这个全连接层通常包含一个具有 1000 个神经元的输出层，对应于 ImageNet 数据集的 1000 个类别

(7) 激活函数：在输出层之后通常会使用 Softmax 激活函数，将网络的原始输出转换为概率分布，用于多类别分类任务。

表 4.1 Resnet50 各组块结构和输出尺寸

Layers	ResNet50	Output Size
Conv1	7x7Conv,s=2	64x64
	3x3Max Pool,s=2	
Conv2_x	$\begin{bmatrix} 1x1, 64 \\ 3x3, 64 \\ 1x1, 256 \end{bmatrix} \times 3$	32x32
Conv3_x	$\begin{bmatrix} 1x1, 128 \\ 3x3, 128 \\ 1x1, 512 \end{bmatrix} \times 4$	16x16
Conv4_x	$\begin{bmatrix} 1x1, 256 \\ 3x3, 256 \\ 1x1, 1024 \end{bmatrix} \times 6$	16x16
Conv5_x	$\begin{bmatrix} 1x1, 512 \\ 3x3, 512 \\ 1x1, 2048 \end{bmatrix} \times 3$	16x16
Output	5-Fully Connect,Softmax	1x1

该模型的输入段包括 7×7 卷积层和 3×3 最大池化层。采用 7×7 内核大小背后的原理是扩展感受野，从而使输入层能够捕获并保留尽可能多的全局和相邻特征信息之间的语义相关性。这种设置有助于建立复杂的空间位置关系和提高模型性能。输入之后是模型的核心部分，分为四个不同的块。每个区块围绕三种瓶颈配置进行结构化（如图 2.7 (b) 所示）。在这种设置中，瓶颈结构通过 1×1 卷积修改输入特征向量的维度（增加或减少维度），使用 3×3 卷积进行特征学习，然后将身份映射合并到残差分量中。与传统卷积神经网络的典型线性堆叠相比，这种方法允许更全面地利用深度特征（包括宽度（ w ）、高度（ h ）和深度（ d ）的第三维度），它有效地在每一层内的多个特征图之间建立了稳健的相关性。通过在同一层中线性堆叠不同的特征图，该模型可以快速识别和激活判别区域。

当 100×100 大小的图像被输入到网络中时，它首先用 7×7 、步长为 2 的卷积核进行卷积运算，然后用 3×3 窗口和步长为 2 的池化层进行下采样，得到 64×64 大小的特征图。通过跨越四个块的 12 个瓶颈残差结构（每个包括三个瓶颈残差构造）的后续处理导致输出特征图大小从 32×32 逐渐减小到 16×16 。然后将其引导通过完全连接的层，并最终通过 softmax 层进行分类。

ResNet50 具有 50 层的深度，其中包括堆叠的残差块，以及一些卷积层和全连接层。ResNet50 在 ImageNet 数据集上预训练，并且在图像分类、目标检测和图像分割等任务中都取得了很好的性能。由于其出色的性能和较小的模型规模成为了深度学习领域中应用广泛的模型之一。

4.2 注意力机制模块

ResNet 模型以其准确性而闻名，它在增加模型复杂性和计算需求之间进行了权衡。对此，高效通道注意力网络（ECANet）已经证明，降低通道特征的维度可以影响注意力机制的学习行为。为了解决这一问题，ECANet 采用一维卷积进行多通道维度交互，有效地绕过了多通道降维的需要，这种策略有助于降低模型的复杂性。在蔬菜质量分级的背景下，精确定位缺陷至关重要，本文在模型的每个块中加入了一个有效通道注意力（ECA）模块，如图 4.1 所示，遵循 1×1 卷积。ECA 模块经过战略性设计，可确保在不影响性能的情况下降低复杂性。有效渠

道注意力（ECA）模块的示意图在以下各节中详细介绍，说明了如何实现效率和有效性之间的平衡，增强了该模型准确评定蔬菜质量的能力^[63]。

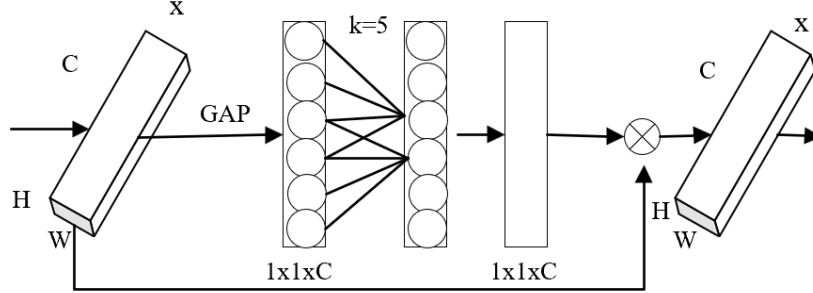


图 4.1 ECA 模块

假设输入特征 $F \in R^{H \times W \times C}$ ，输入特征会经过池化层被压缩得到 $F_{avg} \in R^{1 \times 1 \times C}$ ；随后，采用一维卷积方法来探索每个通道与 k 个相邻通道。这里的 k 不仅表示跨信道覆盖的程度，而且表示一维卷积核的大小的值， k 的数值可以自定义，如果 $k=5$ ，则可以获得 $F_{conv} \in R^{1 \times 1 \times C}$ ；其次使用 Sigmoid 函数对 F_{conv} 进行激活操作，得到权重系数 $M_c(F) \in R^{1 \times 1 \times C}$ ，公式为：

$$M_c(F) = \delta(f^k(\text{AvgPool}(F))) = \delta(f^k(F_{avg}^c)) \quad (4.1)$$

在这里， f^k 表示卷积核尺寸为 k 的卷积操作， F_{avg}^c 表示平均池化操作， δ 表示 Sigmoid 激活函数。公式（4.2）描述了一维卷积神经网络的特征提取过程。将权重系数 $M_c(F)$ 与输入特征 F 逐个元素进行乘积，从而获得新的输出特征 $F1$ 。

$$F1 = M_c(F) \times F \quad (4.2)$$

ECANet，即高效通道注意力网络，是对通道注意力机制的高级形式，通常被看作是 SENet 的增强版。ECA 模块在全局平均池化后直接独特集成了 1×1 卷积层，特别地，它去除了完全连接层。这样的设计选择有效避免了降维，且能够熟练捕捉跨通道的交互。

ECANet 的一个关键优势是它能够在最小的参数参与下获得值得称赞的结果，采用一维卷积层，即 Conv1D，以促进跨信道信息交换。卷积核的大小根据特定

函数进行动态调整,允许具有更多通道的层具有更广泛的跨通道交互。自适应功能定义为: $k = \left\lfloor \frac{\log_2(c)}{\gamma} + \frac{b}{\gamma} \right\rfloor$, 其中 $\gamma = 2$, $b = 1$ 。

ECA 注意力机制的实现过程如下:

- (1) 全局平均池化后,将输入特征图从 $[h, w, c]$ 转换为向量 $[1, 1, c]$ 向量。
- (2) 根据特征图的通道数计算自适应一维卷积核大小 `kernel_size`。
- (3) 使用该卷积核大小在一维卷积中计算特征图每个通道的权重。

这种方法有效地结合了效率和性能,使 ECANet 成为增强各种深度学习架构中特征表示的强大工具。

4.3 卷积替换

ResNet 通过其巧妙的残差结构解决了深度神经网络训练中常见的梯度消失和梯度爆炸问题。在 ResNet 中,一个标准的残差块包含三个卷积层,首先使用 1×1 卷积进行降维以减少参数量和计算量,然后通过 3×3 卷积提取特征,最后使用 1×1 卷积恢复到原始维度。

这种策略通常被描述为“瓶颈”方法,目的是在减少计算复杂度的同时,保持网络强大的表达能力。与此相反,MobileNetV2 引入了一种新颖的设计,称为“反转残差结构”,这种架构颠倒了 ResNet 中的顺序,采用了“小-大-小”的方式。具体来说,它先通过 1×1 卷积进行维度扩展(增强特征空间),然后使用 3×3 深度可分离卷积(简称 $3 \times 3 \text{DW}$ 卷积)进行有效的特征提取,接着,再次使用 1×1 卷积进行降维。深度可分离卷积通过将标准卷积拆分为深度卷积(针对每个输入通道进行单独卷积)和逐点卷积(使用 1×1 卷积合并特征),显著减少了模型大小和计算需求。这种特定的残差结构在下图 4.2 中有详细说明。

传统的 3×3 卷积操作中,如果输入和输出层都有 C 个通道,卷积核尺寸为 $3 \times 3 \times C \times C$ 。相比之下,深度可分离卷积采用了一种更高效的策略:它先通过深度卷积(即对每个输入通道分别使用一个 3×3 的卷积核)进行处理,随后利用 1×1 的卷积核实现通道的融合。这种方法大幅降低了参数数量,因为深度卷积仅需 $3 \times 3 \times C$ 个参数,而 1×1 卷积则需要 $C \times C$ 个参数。这种方式显著减少了模型的参数量和计算需求。MobileNetV2 采用的深度可分离卷积,在减小模型体积和计算

复杂度的同时，仍保持了良好的网络性能。这使得该网络更适合于资源有限的环境，同时还能保持较高的准确率。

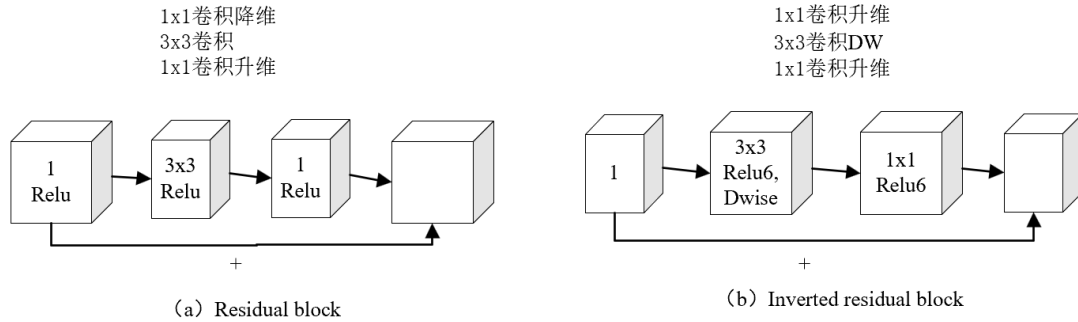


图 4.2 ResNet50 与 MobileNetV2 残差结构对比

考虑到深度可分离卷积的这些优势，我们在本研究中进行了创新性的改进：将 MobileNetV2 中的深度可分离卷积引入替代 ResNet50 中的传统卷积层。这种改进不仅大幅度降低了模型的参数量和计算复杂度，还保留了网络出色的特征提取能力。此举极大地提升了模型在资源有限环境中的实用性和效率，使其更加适用于边缘计算和移动设备，同时确保了高水平的性能表现。通过这一改进，我们的模型在轻量化与高效性能之间实现了完美的平衡。

4.4 ASPP 模块

在深度学习领域，ASPP (Atrous Spatial Pyramid Pooling, 空洞空间金字塔池化) 模块是一种高效的特征提取机制。其核心思想是利用不同膨胀率的空洞卷积 (Atrous Convolution) 来捕获多尺度信息，有效提升网络对不同尺度物体在图像中的识别与分割能力。本研究中，ASPP 模块被置于网络的后端，进行多尺度信息的融合，这一策略显著增强了对图像特征的抽取能力，有效提升了分辨率、精度以及对复杂场景的处理能力。ASPP 模块的结构如图 4.3 所示，它主要由空洞卷积构成。

空洞卷积设计初衷是解决在提取图像特征时如何获得较大感受野同时保持特征图的高分辨率的矛盾。通常，获取较大感受野需要使用较大的卷积核或者在池化操作中采用较大的步长 (stride)，但这两种方法分别面临计算量过大和分辨

率损失的问题。空洞卷积恰好平衡了这两方面的需求：它能够扩大感受野，同时尽可能保持分辨率，以避免丢失图像边缘的细节信息。因此，在本研究中引入 ASPP 模块并将其置于网络后端，是对网络结构的一次重要优化，有效提升了模型对图像的处理能力。

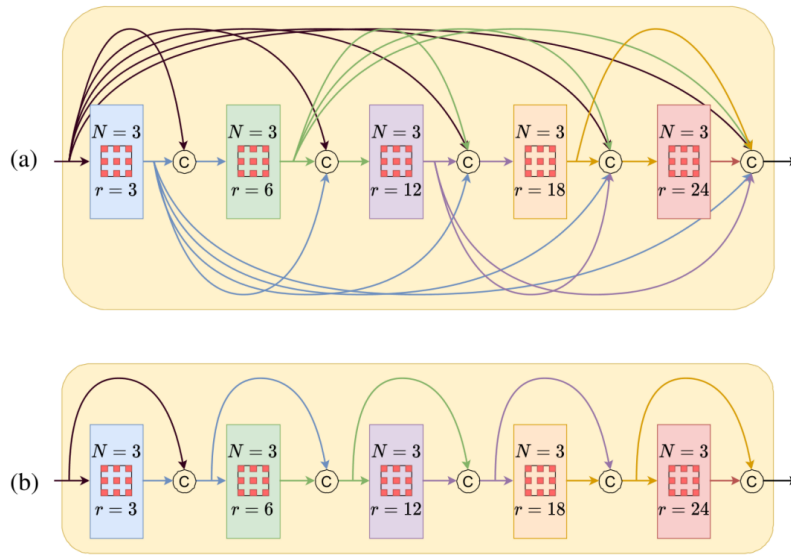


图 4.3 ASPP 模块结构图

4.5 优化网络参数

4.5.1 损失函数

在模型训练时，若仅聚焦于正确标签的损失而忽略其余标签，往往会使模型偏向于仅提升正确标签的识别概率，容易引发过拟合问题。为应对此问题，本研究选用交叉熵作为损失函数，旨在优化目标网络。交叉熵不仅促进正确分类概率的提升，同时有效降低误分类发生，增强了模型的整体学习效果。交叉熵通过计算训练过程中预测的概率分布与真实分布之间的差异来评估模型分类的准确性，并致力于减少这一差异，以提升分类的准确度。数学上，交叉熵衡量了在给定的样本集上，两个概率分布（表示为 p 和 q ）的不同程度，交叉熵的离散函数形式定义为：

$$H(p, q) = -\sum_x p(x) \log q(x) \tag{4.3}$$

在上式中， $p(x)$ 表示真实分布， $q(x)$ 表示使用该数据计算出的估计概率分布。交叉熵 $H(p, q)$ 是用作这两个概率分布之间的相似性的度量，值越大，其差异程度也更大。相反，值越小，较小的值表明估计的概率分布与真实分布紧密一致。假设测试集包含了 N 个样本，则相应的交叉熵损失函数数学方式为：

$$H(p, q) = -\frac{1}{N} \sum_{i=1}^N [y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)] \quad (4.4)$$

当真实分布 $p(x)$ 非常接近概率分布 $q(x)$ 时，损失函数接近于零。该损失函数是在训练过程中导出的，用作评估模型性能的指标。最小化交叉熵损失有助于提高模型的测试精度。

4.5.2 AdamW 优化器

在卷积神经网络中，参数特性对模型的整体质量至关重要。因此，参数优化成为提升模型表现的关键途径。传统的梯度下降法用于搜索参数空间的最优解，而该方法已发展为随机梯度下降（SGD），该算法通过固定的学习率同步更新模型中所有权重。然而，这种不变的学习率可能不适合广泛的数据训练和复杂、高精度的模型，导致学习效率不理想。因此，在本章中，我们选择了一种新的方法：AdamW 算法^[64]，它有望进一步完善模型的结构并解决这些限制。

AdamW（Adam with Weight Decay Fix）是一种用于神经网络训练的优化算法，它是对标准 Adam 优化算法的改进，旨在更好地处理权重衰减（weight decay）的问题。权重衰减是一种正则化技术，用于防止模型过拟合。以下是 AdamW 算法的流程：

（1）初始化参数：

初始化模型参数（权重）： θ 初始化为适当的随机值。

初始化优化算法的动量（momentum）变量 m 和 v 初始化为零。

设置学习率 lr 、权重衰减率 wd （通常称为 λ ）、 β_1 和 β_2 这些超参数。

（2）迭代更新参数：

对每个迭代步骤 t ，计算损失函数关于模型参数 θ 的梯度 g_t 。

(3) 更新一阶和二阶矩估计：

计算一阶矩估 m_t 和二阶矩估 v_t ，它们分别表示梯度的移动平均值和梯度平方的移动平均值。

$$m_t = \beta_1 * m_{t-1} + (1 - \beta_1) * g_t \quad (4.5)$$

$$v_t = \beta_2 * v_{t-1} + (1 - \beta_2) * (g_t)^2 \quad (4.6)$$

(4) 修正偏差 (Bias Correction)：

由于初始化时 m 和 v 均为零，它们会有偏差，因此需要修正：

$$m_{t_hat} = m_t / (1 - \beta_1^t) \quad (4.7)$$

$$v_{t_hat} = v_t / (1 - \beta_2^t) \quad (4.8)$$

(5) 更新参数：

根据修正后的一阶和二阶矩估计，更新模型参数 θ ：

$$\theta_t = \theta_{t-1} - lr * m_{t_hat} / (\sqrt{v_{t_hat}} + \varepsilon) \quad (4.9)$$

这里， ε 是一个很小的常数，通常用于防止分母为零的情况。

(6) 应用权重衰减：

与标准 Adam 不同的是，AdamW 将权重衰减应用在参数更新之前，而不是将它包含在学习率中。所以，最终的参数更新为：

$$\theta_t = \theta_t - lr * wd * \theta_t \quad (4.10)$$

(7) 重复上述步骤直到达到预定的迭代次数或满足停止条件。

AdamW 的改进之一是将权重衰减单独应用于模型参数，而不是与学习率一起处理。这有助于更好地控制权重衰减的效果，防止模型的权重变得过大。它通常在训练神经网络时表现良好，可以提高模型的泛化性能。

4.6 ECA-DS-ResNet50 融合网络模型构建

本研究通过对原始网络的一系列创新改进，提出了专门针对高原夏菜质量分级的 ECA-DS-ResNet50 网络模型，其结构展示于图 4.4。首先，作为网络的主干，ResNet50 通过融入 ECA（高效通道注意力）机制，显著增强了对图像特征的捕捉能力。其次，鉴于 ResNet50 中 7x7 卷积核和池化层的高计算成本，本模型采

用深度可分离卷积来替换其 3×3 标准卷积。这一改变不仅继承了 MobileNetV2 中 DW（深度卷积）和 PW（逐点卷积）的高效机制，还有效降低了整体的计算量。此外，模型通过在每次卷积后加入池化层来进一步削减参数量，优化了网络架构。这种改进显著降低了模型的时间和空间复杂度，有效提高了运算效率。最后，通过引入能进行多尺度融合的 ASPP（空洞空间金字塔池化）模块，模型在增强图像特征抽取能力的同时，也提升了对多尺度信息的处理能力。

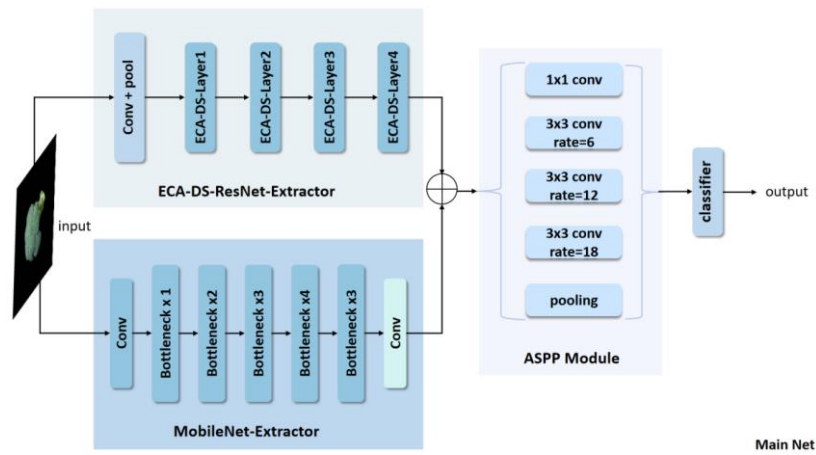


图 4.4 改进网络总体框架

本网络模型精巧地融合了三个主要部分：ECA-DS-ResNet-Extractor、MobileNet-Extractor 以及 ASPP 模块。在处理流程中，图像首先分别通过 ECA-DS-ResNet-Extractor 和 MobileNet-Extractor 模块进行特征提取。这两个模块各自提取的特征图随后进行逐元素相加融合，其中 MobileNet-Extractor 在最后一个 Bottleneck 之后加入了一个 1×1 的卷积层，用于实现维度的扩展。融合后的特征图进一步送入 ASPP 模块，该模块通过不同尺度的感受野捕捉图像的多尺度信息，有效提升图像分类的准确性和鲁棒性，在本实验中 ASPP 模块只使用了图 4.3 中 ASPP 结构图的中间三个卷积层。最终，通过一个全连接层的分类器，输出最终的结果。

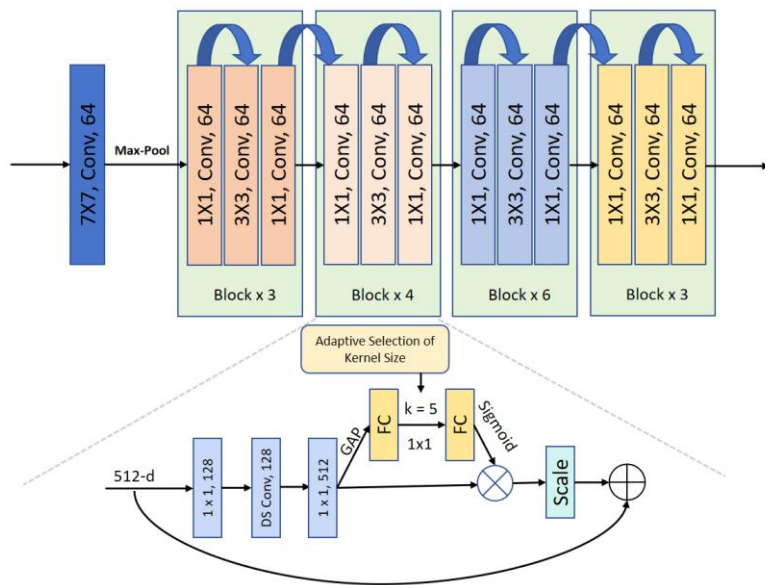
在模型运行过程中，输入的图片会同时经由 ECA-DS-ResNet-Extractor 和 MobileNet-Extractor 两个分支进行特征提取，紧接着进行特征融合，这一过程中融合了注意力机制。融合后的特征随后输入至 ASPP 模块进行进一步的多尺度特征融合，并最终通过分类器得到分类结果。

本研究提出的融合模型结合了多种先进技术,并在 ResNet50 和 MobileNetV2 的基础上进行了创新性的改进和整合。这种融合模型不仅吸收了 ResNet50 和 MobileNetV2 各自的优势,而且通过高效的特征提取、信息融合及多尺度感知机制,显著提高了图像分类任务的整体性能和效率。

下面将逐一详细介绍图 4.4 改进网络后的三个主要部分:

(1) ECA-DS-ResNet-Extractor 部分

在本研究改进的模型中, ECA-DS-ResNet-Extractor 的详细内部结构展示于下图 4.5。在先前提及的图 4.4 中, ECA-DS-Layer1-4 的每一层分别对应图 4.5 中的四个模块。这些模块中, 每一个均由 3、4、6、3 个 Block 组成, 而每个 Block 都是由 1x1, 3x3, 1x1 的卷积层构建。每个 Block 内部的详细结构展示在图 4.5 的虚线框下方。与传统的 1x1, 3x3, 1x1 卷积层不同, 此处将 3x3 的普通卷积替换为了 DS (深度可分离) 卷积。经过三层卷积处理后, 数据流入 ECA (高效通道注意力) 模块, 该模块包含两个分支。其中一个分支执行全局平均池化 (GAP), 随后数据通过两个全连接层, 并经过 Sigmoid 函数处理。处理后的结果与另一分支的直接输出相乘, 得到的 Scale 结果与最初的输入相加, 从而完成对特征的加强和优化。



ECA-DS-Extractor

图 4.5 ECA-DS-ResNet-Extractor 结构图

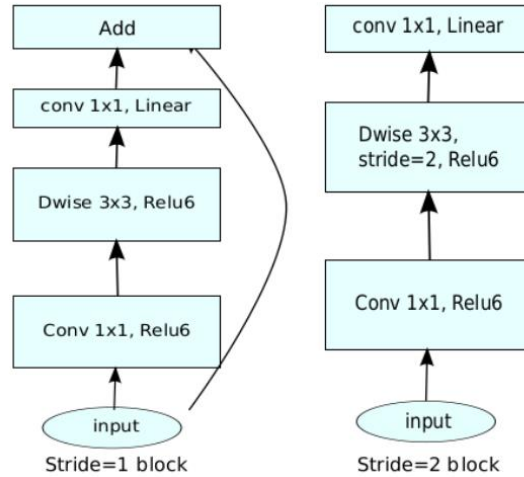


图 4.6 Bottleneck 内部图

(2) MobileNet-Extractor 部分

在本研究所改进的模型中，MobileNet-Extractor 的结构如图 4.4 所示。具体到每个 Bottleneck 块的内部结构，可参见图 4.6。值得注意的是，Bottleneck 块的结构会根据不同的步长 (stride) 有所变化。其主要差异表现在，当步长为 1 时，该模块的设计包括将原始输入与模块处理后的结果进行相加。这一特点使得网络在保持特征图尺寸不变的同时，有效地融合了不同层次的特征信息，从而增强了模型的特征提取和学习能力。通过这种精心设计的结构，MobileNet-Extractor 在本模型中发挥着关键的作用，优化了整体网络性能。

(3) ASPP 部分

ASPP 模块的详细结构已在第四章第四节进行了介绍。在本研究中，我们仅使用图 4.3 所示的 ASPP 模块中的中间三个卷积层来捕获多尺度信息。ASPP 模块被安置在网络的后端，用于将前面两个模块提取的特征图进行逐元素相加，然后送入 ASPP 模块进行多尺度信息的融合。这种设计策略显著增强了网络对图像特征的提取能力，有效地提升了图像的分辨率、识别精度以及处理复杂场景的能力。最后，通过一个全连接层的分类器来输出最终的分类结果，确保了模型的高效性和准确性。

ASPP 模块通常包括以下几个步骤：

(1) 多尺度空间卷积：ASPP 模块使用多个不同采样率（或称为膨胀率）的空洞卷积来扩展感受野。通过在卷积层中引入不同的采样率，可以有效地捕获不

同尺度的特征信息，从而提高模型的感知范围。

(2) 空间卷积金字塔池化：ASPP 模块还采用了空洞卷积金字塔池化的思想。即在不同的采样率下进行空洞卷积操作后，对得到的特征图进行池化操作。这样获得不同尺度下的池化特征，进一步增强了模型对不同尺度物体的感知能力。

(3) 特征融合：ASPP 模块通常会对多个尺度的特征进行融合，以产生更具有丰富语义信息的特征表示。这可以通过简单的级联或并行操作来实现。

4.7 本章小结

本章介绍了高原夏菜质量分级的方法，首先对原网络 ResNet50 的基本结构进行介绍，之后对改进模型的具体改进方法进行阐述，首先，作为网络的主干，ResNet50 通过融入 ECA（高效通道注意力）机制。其次，鉴于 ResNet50 中 7x7 卷积核和池化层的高计算成本，本模型采用深度可分离卷积来替换其 3x3 标准卷积。此外，模型在每层卷积后加入了池化层，进一步减少了参数数量，优化了网络结构，并且，加入 AdamW 优化算法，提高网络模型的准确率。最后，通过引入能进行多尺度融合的 ASPP（空洞空间金字塔池化）模块，模型在增强图像特征抽取能力的同时，也提升了对多尺度信息的处理能力。

5 实验结果与分析

5.1 高原夏菜数据集划分

在第三章中，我们成功地构建了一个有效的高原夏菜数据集。该数据集包括四种不同类型的蔬菜：结球甘蓝、娃娃菜、菜花和西兰花。每种蔬菜类型进一步分为三个质量等级：特级、一级和二级，最终形成 12 个不同的类别。最初，我们收集了每个类别 200 张图像，经过广泛的数据扩充和额外的摄影，扩展到每个类别 360 张图像，最终收集了 4320 张图像。为了增强数据集的多样性和稳健性，每个蔬菜类别都进行了随机几何变换，包括 90 度旋转、水平翻转和垂直翻转，每个图像都进行随机的其中一种几何变换。此外，对一组图像进行了人工修改包括椒盐噪声或高斯噪声，使总数据集达到近 5000 张图像。为了进行有效的训练和评估，数据集被精心划分为三个子集：训练集、验证集和测试集，比例为 7:2:1，确保了模型训练、性能验证和最终测试的平衡方法。

5.2 实验

5.2.1 实验内容

在使用本实验的高原夏菜数据集进行蔬菜质量等级识别任务之前，已经在大型数据集（ImageNet）上预训练过模型，因此它已具备一定的区分蔬菜质量等级的能力。

预训练模型在深度学习和机器学习领域扮演着非常重要的角色，尤其是在需要处理大量复杂数据的任务中。预训练模型的主要作用包括：

（1）加速训练过程：使用预训练模型可以显著减少训练时间。

（2）提高性能：对于数据量较小的任务，从头开始训练模型可能不会得到良好的性能，因为模型可能没有足够的数据来学习复杂的特征。预训练模型因为已经在大型数据集上训练，所以能够捕捉更丰富的特征信息，提高最终任务的准确率。

（3）克服数据不足的问题：在某些领域，如医学图像处理，获取大量标注

数据可能非常困难和昂贵。预训练模型可以在其他领域的大型数据集上进行训练，然后迁移到特定领域，减少对大量标注数据的需求。

(4) 特征提取：预训练模型可以作为强大的特征提取器。在一些应用中，可以直接使用预训练模型的输出作为特征，输入到简单的机器学习模型中进行分类或其他任务。

综上所述，预训练模型在提高训练效率、增强模型性能、解决数据不足问题以及作为迁移学习的基础等方面发挥着关键作用。

为全面评估改进模型的效能，本节对模型的训练成果进行了深入的分析。具体而言，我们对模型在执行蔬菜分类、蔬菜分级以及融合蔬菜分级和分类这三个任务的结果进行了详尽的审查：

(1) 蔬菜类别分类：第一部分重点考察该模型准确分类蔬菜类别的能力。评估作为了解模型基本分类强度的基线。

(2) 对不同质量的蔬菜进行分级：第二部分介绍了该模型在区分不同质量等级的蔬菜方面的有效程度，对四种类型的高原夏菜进行了原始和改进模型的训练。对准确率、F1 值、参数量、中值推理时间（inference time (median)）和每秒帧数（FPS）等关键性能指标进行了仔细的比较和分析，以衡量模型在质量分级中的有效性。接着与经典卷积神经网络模型比较分析，最后，以一系列消融实验结束。这些实验旨在横向比较数据集上改进模型的每个变体的训练结果，从而验证每个单独改进模块的贡献和有效性。这一全面的分析不仅展示了改进后模型的功能，还深入了解了对其性能贡献最大的具体增强功能。

(3) 针对 12 个不同类别的蔬菜，本研究同时进行了精准的分类与分级处理：在提升效率方面，一次性处理多个蔬菜类别，节省了时间和资源，避免了对每种蔬菜单独进行分类和分级的重复工作。此实验不仅提升了模型处理效率和准确性，而且还增强了模型的泛化能力和市场适应性，为蔬菜行业的发展提供了强有力的技术支持。

5.2.2 实验环境及参数设置

本实验训练通过 AutoDL 平台，该平台是一款云 GPU 深度学习环境出租平台，资源丰富且非常高效。所用的服务器配备了 GPU :RTX 3090(24GB) * 1 ；

CPU:16 vCPU Intel(R) Xeon(R) Platinum 8350C CPU @2.60GHz; 具有 Cuda 11.3 的加速框架, 80G 内存。深度学习配置的环境为 PyTorch 1.11.0 Python 3.8 (ubuntu20.04), 使用其中的深度学习工具箱来设计和实现高原夏菜图像等级识别。详细参数设置如表 5.1 所示。

表 5.1 实验参数设置

模型参数名称	详情
Batch size	32
Epoch	75
Ema_rate	0.999
Learning_rate	1e-4
lr_decay_type	cosineAnnWarm
Optimizer	AdamW
Weight_decay	0.1

5.2.3 评价指标

深度学习模型的功效最好通过使用各种指标量化和比较其性能来确定。在这项研究中, 我们主要通过准确率和 F_1 值来评估和测试改进的 ResNet50 网络在对高原夏菜质量进行分类方面的能力。此外, 使用模型参数的大小来衡量算法的复杂性, 同时使用 inference time (median) 中值推理时间和每秒帧数 (FPS) 来评估算法的时间复杂性。这些措施有助于对原始模型和改进模型进行全面比较, 也有助于确定改进模型的任何局限性。

准确率直观反映了模型将样本 (包括正类与负类) 正确判定的比例, 是衡量模型性能的基本指标。召回率则用于衡量模型对于正类样本的识别效果, 指出在所有真实正类样本中, 模型能够正确识别的份额。鉴于准确率与召回率往往呈现出一定的对立性, 寻找二者之间的最佳平衡点成为常见的做法。在此背景下, F_1 分数, 作为准确率与召回率的加权调和平均, 成为一个重要的综合评价指标。 F_1

分数接近 1 表示网络模型的性能更加平衡和有效,有效地解决了准确性和召回率之间的内在矛盾。

$$Accuracy = \frac{TP + TN}{N} \quad (5.1)$$

$$F_1 = \frac{2TP}{N + TP - TN} \quad (5.2)$$

其中, TP 为真正例样本数,即数据集中就是正样本,预测结果也为正样本; TN 为真负例样本数,即数据集中就是负样本,预测结果也为负样本; N 为数据集样本总数。

“inference time (median)” 是中值推理时间,一种衡量机器学习模型或系统在执行推理(即模型对新输入数据进行处理并给出结果的过程)时性能的评价指标。具体来说,它指的是:

inference time: 推理时间是模型从接收输入(如一张图像、一段语音等)到产生输出(如分类标签、检测结果等)所需的时间。它是衡量模型处理速度的直接指标。

Median: 中位数是一种统计度量,表示在一系列数据中处于中间位置的值。在这里,中位数推理时间是指在多次推理操作中,所有推理时间从小到大排列后位于正中间的那个时间值。它是一个比平均值更鲁棒的统计指标,因为它不会受到极端值的影响。

综合来看,“inference time (median)” 提供了一个关于模型在实际运行中推理速度的可靠度量,特别是在需要快速响应的应用场景(如实时数据处理、在线服务等)中,这一指标尤为重要。较低的中位数推理时间意味着模型能够更快地处理输入并给出结果,这在许多实时或近实时的应用中是非常重要的性能指标。

FPS 指“每秒帧数”(Frames Per Second),是一个用来衡量图像处理速度的重要参数。表示系统或设备每秒能处理或显示大约多少帧图像。**FPS** 越高,表明图像处理速度更快。

5.3 蔬菜类别分类

经过对原始 ResNet 模型的改进,我们在高原夏菜的训练和验证数据集上对

该模型进行了训练，并对特级、一级、二级这三个不同等级的蔬菜进行了分类。实验结果详见表 5.2。

表 5.2 三种级别高原夏菜分类性能对比

蔬菜类别	ResNet50/%		改进 ResNet50/%	
	准确率	F1	准确率	F1
特级	100.00	100.00	100.00	100.00
一级	99.78	99.64	100.00	100.00
二级	100.00	100.00	100.00	100.00

该表展示了在特级、一级和二级这三个不同级别的四种高原夏菜数据集上，改进后的 ECA-DS-ResNet50 模型与原始 ResNet50 模型的准确率和 F1 值变化情况。从表中数据可以明显看出，改进后的模型与原始网络模型在识别四种高原夏菜方面均展现出了良好的性能，能够充分区分这四种不同类型的蔬菜，此实验表明了模型准确分类蔬菜类别的能力，可以评估作为了解模型基本分类强度的基线。这一结果不仅验证了对原始 ResNet 模型改进的有效性，也展示了在不同质量等级蔬菜分类任务中的卓越表现。

5.4 不同质量蔬菜间分级

5.4.1 与原网络模型性能对比

在本研究中，我们利用改进后的 ResNet 模型对高原夏菜的训练集和验证集进行了训练，并对四种不同的蔬菜进行了质量等级的分类。实验结果详细列出在表 5.3 中。该表格展示了改进的 ECA-DS-ResNet50 模型与原始 ResNet50 模型在四种高原夏菜数据集上的准确率和 F1 值的变化情况。从表格数据中可以观察到，改进后的模型在四种高原夏菜的识别上均表现出色。特别是在菜花和娃娃菜上，其准确率和 F1 值明显高于其他两种蔬菜。这主要得益于菜花和娃娃菜在外观、大小和缺陷程度上的显著区别，使得它们在拍照采集过程中就能被有效区分，从而更容易进行质量分级识别，因此获得了较高的准确率。然而，结球甘蓝和西兰

花的准确率与 F1 值并不像前两者那样显著，这可能与第三章中提到的蔬菜预处理过程有关。结球甘蓝和西兰花在花球的外观和机械损伤程度上差异较小，导致它们的分级准确率不如菜花和娃娃菜。总体而言，改进模型在以上四种高原夏菜的识别中明显优于原始模型，充分证明了模型改进的有效性。

表 5.3 四种高原夏菜质量分级性能对比

蔬菜种类	ResNet50/%		改进 ResNet50/%	
	准确率	F1	准确率	F1
结球甘蓝	87.64	87.35	90.50	90.32
西兰花	90.10	88.87	93.67	92.54
菜花	98.45	98.45	99.30	99.25
娃娃菜	98.51	98.43	99.46	99.12

同时，准确率和 F1 值几乎相等，可能表明以下几点：

(1) 类别平衡：这种情况更可能出现在类别较为平衡的数据集上，因为在类别不平衡的情况下，即使准确率很高，F1 值也可能因为低召回率而较低。

(2) 模型性能均衡：表示模型在精确率和召回率之间取得了较好的平衡，没有过度偏向于其中一个指标。

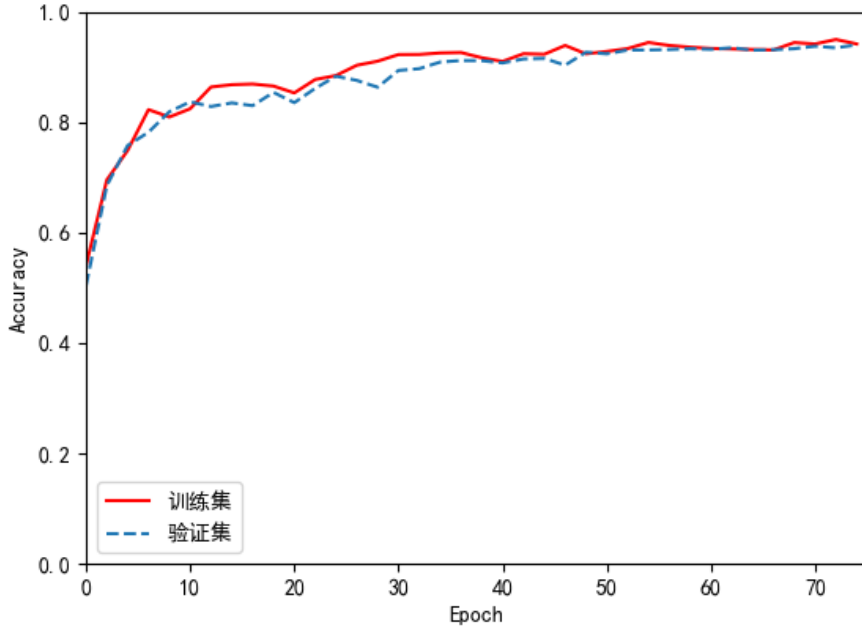
(3) 无极端表现：模型没有在某一个类别上表现极端差或极端好，这通常会导导致准确率和 F1 值之间的显著差异。

总的来说，准确率与 F1 值相等通常意味着模型在精确度和召回率方面取得了平衡的性能表现，尤其是在类别相对均衡的数据集中。

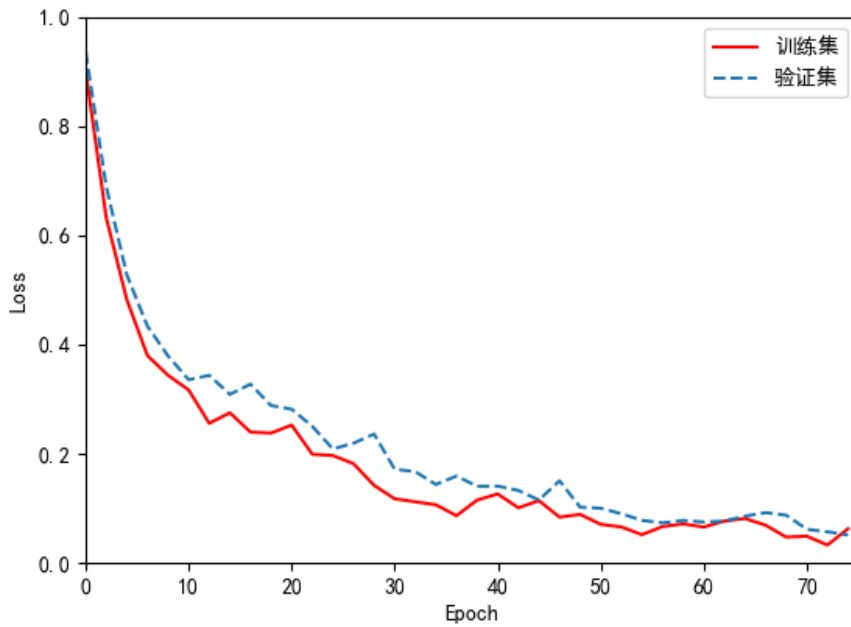
考虑到结球甘蓝和西兰花在本实验中的改进前后准确率提升表现明显优于其他两种蔬菜，并且能更有效地展示本实验在解决质量相近蔬菜分级问题上的成效，我们决定在后续的对比实验中主要使用这两种蔬菜作为研究对象。

在实验中，我们对改进后的 ResNet 模型在高原夏菜的训练集和验证集上进行了深入训练。为了更直观地呈现模型的性能，我们采用了可视化的方法来展示实验结果。以西兰花为例，其模型训练过程中的准确率和损失变化曲线在图 5.1 中有详细展示。通过这种可视化表示，我们不仅能清晰地观察到模型在训练和测

试过程中的表现，还能直观地理解模型对特定蔬菜分级的准确性和效率。这为评估模型在处理质量相近的蔬菜分级问题上的有效性提供了有力的证据。



(a) 网络训练 Accuracy 对比图



(b) 网络训练 Loss 对比图

图 5.1 ECA-DS-ResNet50 模型在西兰花上 Accuracy 与 Loss 曲线图

从上图中可以观察到，当模型迭代次数达到大约 40 次时，改进的 ResNet 模型的准确率和损失率都开始趋于平稳。这表明准确率已经趋于稳定，损失率也基本收敛，说明模型已经达到了饱和状态，此时的识别准确率已达到最优，高达 93.67%。

值得一提的是，在我们的训练集和验证集中，模型的准确率并不是从零开始的。这主要归因于我们在将模型应用于本次蔬菜质量等级识别任务之前，已经采用了在大型数据集（ImageNet）上预训练模型。因此，模型在一开始就已经具备了识别不同蔬菜质量等级的基本能力。这就解释了为什么模型能够在初始阶段就达到一定的准确率，而准确率曲线并非起始于 y 轴的零点。

这一点在评估模型的初始性能以及训练过程中的进步时尤其重要。对于数据量相对较小的任务来说，如果模型从头开始训练，可能由于数据不足而无法学习到复杂的特征，从而难以达到良好的性能。然而，由于预训练模型已经在大规模数据集上进行了训练，它能够捕捉到更多的特征信息，从而提高了针对特定任务的准确率。同时，模型一开始就具备识别蔬菜质量等级的能力，也大大减少了训练所需的时间。

我们对改进后的模型在结球甘蓝和西兰花数据集上进行了训练，并将其与单一 ResNet50 模型的训练结果进行了对比。为了全面评估模型性能，我们采用了准确率、模型参数量、中值推理时间（inference time (median)）和每秒帧数（FPS）作为评价指标。相关结果详见下表 5.4。

表 5.4 两种高原夏菜改进 ResNet50 模型前后的各指标对比

参数	结球甘蓝		西兰花	
	ResNet50	改进的 ResNet50	ResNet50	改进的 ResNet50
准确率	87.64%	90.50%	90.10%	93.67%
模型参数	23.51M	23.94M	23.51M	23.94M
inference time (median)	6.8512 ms	6.0822 ms	6.8780 ms	6.1777 ms
FPS	145.96	164.41	145.39	161.87

正如表 5.4 所示，在两种高原夏菜的识别任务上，改进后的模型均展现出了

良好的表现。在模型参数量几乎不变的情况下，结球甘蓝和西兰花的识别准确率分别提升了 2.86%和 3.37%。此外，从表中可以看出，对于结球甘蓝，改进后的 ResNet50 模型的中值推理时间为 6.0822 毫秒，而原始网络模型为 6.8512 毫秒。较低的中值推理时间意味着模型能够更快地处理输入数据并给出结果。同时，结球甘蓝在改进后的 ResNet50 模型中的 FPS 值达到了 164.41，即系统或设备每秒能处理约 164 帧图像，而原始网络模型的处理速度约为每秒 145 帧。这一显著的提高表明，改进后的模型在图像处理速度方面有了明显的优势，充分证明了模型改进的有效性。

表 5.5 结球甘蓝在每一个蔬菜质量分级中的各指标对比

蔬菜类别	准确率		F1 值	
	ResNet50	改进的 ResNet50	ResNet50	改进的 ResNet50
特级	93%	95%	95%	95%
一级	86%	90%	88%	90%
二级	83%	86%	79%	84%

从表 5.5 中可以清楚地看到，在执行质量分级任务时，改进后的 ResNet50 模型在每个蔬菜等级的识别过程中均实现了不同程度的性能提升。这种全面的改进使得整体的识别效果得到了显著提升，从而使模型的总体准确率得到了显著增长。这一结果不仅证明了改进措施的有效性，也展示了模型在处理更复杂和细化的分类任务时的强大能力。

改进前后的网络模型分别在特级、一级和二级的数据集上准确率和 F1 值逐渐降低，这可能与不同级别选取的蔬菜差异性有关，特级选取的蔬菜差异性大，模型更容易区分和识别，在识别特级蔬菜时表现更好，识别准确率也高。总体而言，改进后的 ResNet50 模型通过在各个级别上的细致优化，成功提高了对蔬菜质量等级的识别精度，从而在整体上显著提升了模型的实用价值和准确性。

5.4.2 与经典深度卷积神经网络模型的对比

将改进的模型与传统卷积神经网络进行比较评估。此比较旨在验证新模型带来的进步和改进。本文接下来选择在西兰花数据集上对模型进行训练。具体来说，我们在西兰花的验证数据集上对改进 ResNet50 模型与其他几种模型——ConvNeXt_Small、DenseNet121、CBAM_ResNet50 以及 EfficientNet_B5——进行了横向比较，并对比较结果进行了可视化展示，详见图 5.2。

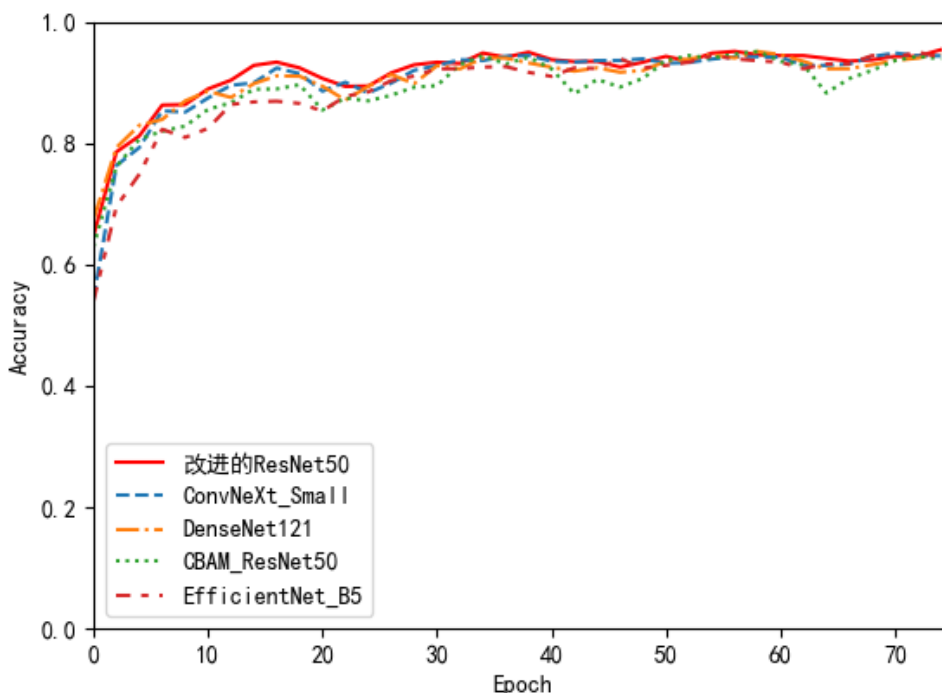


图 5.2 不同模型准确率对比图

从图 5.2 中可以清楚地看到，随着迭代次数的增加，五种模型的准确率整体上呈现出上升的趋势，最终逐渐趋于稳定。特别是在迭代次数达到大约 30 次时，改进后的 ResNet 模型在早期阶段的准确率普遍高于其他模型，显示出其优秀的性能。当迭代次数增加到大约 50 次时，这五种模型的准确率均达到一个平稳的状态，并且彼此之间的差异变得不那么显著，但总体上来看改进后的 ResNet 模型在实验结果方面依然占优势。这一现象表明，随着足够的训练，不同的模型在特定数据集上的表现趋于一致。各模型在西兰花验证集上的具体识别结果详细列于表 5.6 中。这些结果为我们提供了关于不同模型在实际应用中性能的宝贵信息，

有助于我们在未来的应用和研究中作出更加合理的模型选择和调整。

表 5.6 西兰花各模型实验结果对比表

模型	准确率/%	参数量/M	inference time (median)/ms	FPS
cbam_resnet50	91.78	26.03	12.4428	80.37
EfficientNet_B5	92.45	28.35	11.8039	84.72
ConvNeXt_Small	92.56	49.46	14.1894	70.48
DenseNet121	92.32	6.96	9.6174	103.98
改进 ResNet50	93.67	23.94	6.8780	145.39

根据表 5.6 的数据，我们可以看到本研究提出的方法在西兰花蔬菜图像验证集上展现了最佳的分级准确率提升效果，达到了 93.67%。改进后的 ResNet 模型在验证集上的准确率相比于 CBAM_ResNet50、EfficientNet_B5、ConvNeXt_Small、DenseNet121 分别提升了 1.89%、1.22%、1.11%、1.35%。与前三种模型相比，我们的模型在快速识别图像特征的同时，在模型参数量上具有明显优势。对于 DenseNet121 网络而言，我们的模型主要在模型识别图像时间方面表现更佳。本研究所改进的模型通过将深度可分离卷积、ECA（高效通道注意力）模块、ASPP 模块与 ResNet 的优点相结合，旨在减少计算量和避免过拟合，同时尽可能保留输入图像的关键特征。与以上四种经典卷积神经网络模型相比，本文提出的方法在提高分类准确率的同时，也有效缩短了模型在识别蔬菜质量分级时所需的时间，展现了优异的综合性能。

5.4.3 消融实验对比

为了深入探究改进模型的有效性，本研究设计了一系列的消融实验，对比了在西兰花验证集上应用以下五种不同方式的训练结果：

方式 1：单独使用 ResNet50 模型在西兰花数据集上进行训练。

方式 2：优化原始网络，将 ResNet50 中的标准卷积替换为深度可分离卷积。

方式 3：在方式 1 的基础上，引入 ECA（高效通道注意力）模块。

方式 4: 在方式 2 的基础上, 加入 ECA 模块。

方式 5: 基于方式 4, 将融合后的特征输入至 ASPP 模块进行进一步的多尺度特征融合, 通过分类器得到最终的分类结果, 构成本文提出的改进模型。

通过这些消融实验, 我们能够更清晰地理解各种改进措施对模型性能的具体影响。实验结果展现在下表 5.7 中, 通过对比这些不同配置下的模型性能, 我们可以更准确地评估本文所提出改进模型的有效性和实用价值。

表 5.7 消融实验结果

改进策略	准确率/%	参数量/M	FPS
ResNet50	90.10	23.51	145.39
DS-ResNet50	89.63	22.78	150.67
ECA-ResNet50	92.78	23.65	143.54
ECA-DS-ResNet50	91.86	23.12	146.87
ECA-DS-ResNet50+ASPP	93.67	23.94	161.87

如上表所示, 以上五种消融实验的结果对比如下:

方式 1 与方式 2 对比: 通过将 ResNet50 中的标准卷积替换为深度可分离卷积, 模型的 FPS 值得到了提升。这是因为深度可分离卷积相较于传统卷积, 减少了参数数量和计算复杂度, 使得模型更轻量 and 快速。尽管提高了模型的处理速度, 但也导致了一定程度的表达能力损失, 准确率略有下降。

方式 1 与方式 3 对比: 加入 ECA 模块后, 模型学习了通道间的关系, 从而增强了特征提取能力, 但增加了计算负担。这导致处理每帧图像所需的时间增加, FPS 值降低, 即每秒处理的帧数减少。然而, 准确率方面则有了显著提升。

方式 1、2、3、4 综合比较: 加入 ECA 注意力机制部分抵消了深度可分离卷积带来的速度提升, 使得模型的 FPS 值有所提高, 但提升幅度不及仅使用深度可分离卷积时显著。综合来看, 同时加入 ECA 注意力机制和深度可分离卷积后, 模型的 FPS 略有提高, 同时准确率也得到了提升。

方式 4 与方式 5 比较: 将 ASPP 模块加入并进行多尺度融合特征提取, 证明是一种有效的策略。这表明了对原始网络模型的改进是有效且可行的。

综上所述，通过这些消融实验，我们不仅验证了各种改进方式对模型性能的具体影响，而且展示了在提高处理速度和准确率方面取得的平衡。这些结果为未来模型的优化提供了有价值的参考。

5.5 12 类别蔬菜分级与分类

在交易市场上，对高原夏菜进行质量分类往往需要同时处理多种不同的蔬菜。因此，本研究同时探讨了高原夏菜在同时进行分级与分类方面的综合识别效果，旨在为高原夏菜市场的定价销售提供有价值的参考。实验结果详见下表 5.8，该表格展示了在包括特级、一级和二级共 12 个不同蔬菜类别的数据集上，改进后的 ECA-DS-ResNet50 模型与原始 ResNet50 模型在准确率和 F1 值方面的变化情况。

表 5.8 12 分类准确率与 F1 指标对比

模型	准确率/%	F1
ResNet50	94.78	94.51
改进 ResNet50	95.39	95.03

从表格中可以清晰看到，无论是改进后的模型还是原始网络模型，都在同时进行 12 个类别的高原夏菜精准分类与分级处理方面展现了出色的性能，并能有效区分这些不同类别的蔬菜。这一成果不仅证明了对原始 ResNet 模型所进行改进的有效性，而且也体现了模型在处理不同质量和等级蔬菜分类任务时的优异表现。通过这项研究，我们期望为未来高原夏菜的市场策略和销售决策提供科学的数据支持。

同时，对 12 个不同类别的蔬菜同时进行分类与分级的做法具有以下作用：

(1) 提高效率：一次性处理多个蔬菜类别，节省了时间和资源，避免了对每种蔬菜单独进行分类和分级的重复工作。

(2) 增加准确性：通过集成不同蔬菜的特征，模型可以学习到更丰富的信息，从而提高对每个类别蔬菜的分类和分级准确性。

(3) 强化模型泛化能力：同时处理多种类别的蔬菜能够使模型在面对多样

化数据时表现更加稳定，增强其泛化能力。

5.6 本章小结

本章的核心研究内容是对改进后的模型 ECA-DS-ResNet50 在高原夏菜图像数据集上的训练和应用进行了深入探讨。研究过程包括几个关键步骤：

(1) 蔬菜类别分类：我们首先对四种不同的蔬菜类别进行分类，并将改进模型与原始模型进行对比，以测试改进模型在蔬菜分类方面的效果，考察该模型准确分类蔬菜类别的能力，评估作为了解模型基本分类强度的基线。

(2) 高原夏菜分级处理：随后，我们对四种高原夏菜数据集进行了分级处理。特别地，在西兰花数据集上，我们将改进的模型与 CBAM_ResNet50、EfficientNet_B5、ConvNeXt_Small 和 DenseNet121 进行了横向比较，并进行了详细的消融实验。

(3) 同时进行分类与分级：最后，我们同时进行了高原夏菜的分级与分类工作，对 12 种不同类别的蔬菜进行了分类实验，以验证改进模型的可行性和有效性。通过这一系列的研究步骤，我们全面评估了 ECA-DS-ResNet50 模型在处理高原夏菜图像数据方面的性能，旨在为高原夏菜的自动化分类和分级提供一个更高效、更准确的解决方案。

6 总结与展望

6.1 总结

高原夏菜是甘肃省的支柱性特色产业,虽然规范生产过程方面取得了一些进展,但其销售过程的标准化还存在滞后问题。蔬菜行业的许多公司通常在没有标准化程序的情况下使用冷藏和运输蔬菜,这往往导致在最终市场直接销售。鉴于当前甘肃省高原夏菜数据集的缺乏以及缺少对蔬菜进行质量等级分类的问题,本论文构建了高原夏菜数据集,并提出了一种基于深度学习的高原夏菜质量分级模型,以解决上述问题。具体工作总结如下:

(1) 构建高原夏菜数据集:该数据集包括结球甘蓝、娃娃菜、菜花、西兰花四种蔬菜。鉴于蔬菜分类领域中公开且高质量的数据集较少,特别是高原夏菜数据集更为罕见,本研究建立了一套实验室采集装置,成功构建了高原夏菜数据集。这一数据集不仅用于本研究,还可进一步公开发布,以促进甘肃省乃至全国蔬菜分类领域的发展。

(2) 数据集扩充与预处理:针对原始数据集数量不足的问题,为了确保采集的高原夏菜数据集能在模型训练时取得良好的分级效果,本文首先使用 U^2Net 网络去除采集数据集的背景,减少背景特征对训练模型的干扰。原始数据集每个类别只有 200 张图像,共 2400 张,通过扩充拍摄,每个类别增加到 360 张,共 4320 张,然后对图像进行几何变换(翻转、旋转等),最后通过加入噪声将数据总量扩充至 5000 张。

(3) 提出高原夏菜质量分级模型:考虑到目前高原夏菜质量分级主要依靠传统人工方式,本文提出了一种新的多尺度融合 ECA-DS-ResNet50 模型,专门用于高原夏菜的质量分级,提出的这种融合模型结合了多种技术,在 ResNet50 和 MobileNetV2 的基础上进行了改进和融合。首先,ResNet50 作为主干网络,通过引入 ECA(Efficient Channel Attention)注意力机制和多尺度特征提取的 ASPP 模块,进一步增强了对图像特征的抽取能力。其次,针对 ResNet50 中 7×7 卷积核和池化层的计算成本较高的特点,将其 3×3 标准卷积替换成深度可分离卷积,整合 MobileNetV2 的 DW(Depthwise Convolution)和 PW(Pointwise Convolution)卷积机制取代 ResNet50 的常规卷积,有效降低了计算量。同时,在每层卷积后

引入池化层,进一步减少参数量,优化网络结构,大幅减少了时间和空间复杂度,有效提升模型的运算速度。

6.2 展望

本文深入研究了高原夏菜的质量分级方法。所提出的方法在现有的菜花和娃娃菜数据集上取得了良好的效果,能够准确识别不同等级的蔬菜产品,对甘肃省的蔬菜产业具有重要价值。对于结球甘蓝和西兰花数据集,模型的准确率有了显著提升。然而,当前模型在某些方面仍存在不足,需要进一步深入研究和改进。

(1) 数据集的全面性不够: 鉴于高原夏菜种类繁多,本文仅选用了四种进行研究,可能具有一定的局限性。我们希望进一步扩充高原夏菜数据集,包括数量、种类及蔬菜的高质量缺陷等,以增强数据集的强度并提高模型的普适性。

(2) 特征提取的难度: 由于各类高原夏菜之间存在显著差异,特别是在结球甘蓝和西兰花模型中提取特征时难以区分。因此,对于难以区分质量等级的蔬菜,如何更好地提取表面特征也是高原夏菜质量分级方法的难点之一,需要进一步优化和改进。

(3) 分级算法持续改进: 目前在蔬菜分级研究方面仍然缺乏。尽管本文提出的模型方法已取得良好效果,后续研究仍可以利用其他经典卷积神经网络改进网络模型,从而提高分级准确性,使蔬菜图像分级更加精准。

参考文献

- [1] 王雅萍. 兰州新区西岔镇高原夏菜产业发展现状与对策研究[D]. 兰州大学, 2020.
- [2] 徐长春, 郑戈, 聂善明. 高原夏菜产业发展现状、存在问题及对策建议—基于甘肃省榆中县的调研[J]. 中国农业资源与区划, 2016,37(02):52-56.
- [3] 王茂森. 计算机视觉技术在农产品质量检测中的运用探究[J]. 数字通信世界, 2022(06):129-131.
- [4] Dubey R S, Jalal S A. Fruit and vegetable recognition by fusing colour and texture features of the image using machine learning[J]. Int. J. of Applied Pattern Recognition, 2015, 2 (2): 160-181.
- [5] 程春燕. 基于机器视觉的脐橙品质自动检测分类技术研究[D]. 江西理工大学, 2018.
- [6] 陶华伟, 赵力, 奚吉等. 基于颜色及纹理特征的果蔬种类识别方法[J]. 农业工程学报, 2014, 30(16): 305-311.
- [7] Dubey R S, Jalal A. Robust Approach for Fruit and Vegetable Classification [J]. Procedia Engineering, 2012, 38 (C): 3449-3453.
- [8] Grinblat L G, Uzal C L, Larese G M, et al. Deep learning for plant identification using vein morphological patterns [J]. Computers and Electronics in Agriculture, 2016, 127 418-424.
- [9] Pan L, Sun Y, Xiao H, et al. Hyperspectral imaging with different illumination patterns for the hollowness classification of white radish [J]. Postharvest Biology and Technology, 2017, 126 40-49.
- [10] 郑远攀, 李广阳, 李晔. 深度学习在图像识别中的应用研究综述[J]. 计算机工程与应用, 2019, 55 (12): 20-36.
- [11] Bhargava A, Bansal A. Fruits and vegetables quality evaluation using computer vision: A review [J]. Journal of King Saud University - Computer and Information Sciences, 2018, 33 (3): 243-257.

- [12] Bolle R M , Connell J H , Haas N , et al. VeggieVision: A Produce Recognition System[C]// IEEE Workshop on Applications of Computer Vision. IEEE Computer Society, 1996: 244-251.
- [13] 刘禾, 汪懋华. 水果果形判别人工神经网络专家系统的研究[J]. 农业工程学报, 1996,(01):171-176.
- [14] Shebiah N R . Fruit Recognition using Color and Texture Features [J]. Journal of Emerging Trends in Computing and Information Sciences, 2010, 1 (2): 80.
- [15] Rocha ,Anderson ,Hauagge , et al. Automatic fruit and vegetable classification from images [J]. Computers and Electronics in Agriculture, 2010, 70 (1): 96-104.
- [16] 霍志林, 吴洪潭, 华亨等. 灰度共生矩阵在蔬菜种类识别中的应用研究[J]. 中国计量学院学报, 2015, 26 (01): 105-109.
- [17] 张亚静, Sakae Shibusawa, 李民赞. 基于机器视觉的番茄内部品质预测[J]. 农业学报, 2010, 26(S2): 366-370.
- [18] Zhang Y , Wang S , Ji G , et al. Fruit classification using computer vision and feedforward neural network [J]. Journal of Food Engineering, 2014, 143 167-177.
- [19] Madgi M, Danti A, Anami B S. Mean and Range Color Features Based Identification of Common Indian Leafy Vegetables[J]. International Journal of Signal Processing Image Processing & Pattern Recognition, 2012, 151-160.
- [20] 姜伟, 倪舟, 汪丰. 蔬菜自动分类识别系统的设计及实现[J]. 北京农业, 2014(27):188+196.
- [21] A. M G , G. E K , D. O S , et al. Classification of Fruit Flies by Gender in Images Using Smartphones and the YOLOv4-Tiny Neural Network [J]. Mathematics, 2022, 10 (3): 295-295.
- [22] Wan P , Toudeshki A , Tan H , et al. A methodology for fresh tomato maturity detection using computer vision [J]. Computers and Electronics in Agriculture, 2018, 146 43-50.
- [23] 芦范. 基于改进神经网络算法的蔬菜图像识别 [J]. 食品与机械, 2020, 36(02):146-150.

- [24] Jordan I M ,Mitchell M T . Machine learning: Trends, perspectives, and prospects [J]. Science, 2015, 349 (6245): 255-260.
- [25] Gu J ,Wang Z ,Kuen J , et al. Recent advances in convolutional neural networks [J]. Pattern Recognition, 2018, 77 354-377.
- [26] Schmidhuber J . Deep learning in neural networks: An overview [J]. Neural Networks, 2015, 61 85-117.
- [27] Shi J ,Li Z ,Zhu T , et al. Defect Detection of Industry Wood Veneer Based on NAS and Multi-Channel Mask R-CNN [J]. Sensors, 2020, 20 (16): 20164398.
- [28] Yabin Y ,Ying L ,Jiawei C , et al. Detection Method for Bolted Connection Looseness at Small Angles of Timber Structures based on Deep Learning [J]. Sensors, 2021, 21 (9): 3106-3106.
- [29] Simonyan K , Zisserman A . Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014:1-14.
- [30] Szegedy C ,0015 L W ,Jia Y , et al. Going Deeper with Convolutions[J]. IEEE Conference on Computer Vision & Pattern Recognition, 2015:1-9.
- [31] He K , Zhang X , Ren S , et al. Deep Residual Learning for Image Recognition[J]. Computer Science, 2015: 770-778.
- [32] Li Z ,Li F ,Zhu L , et al. Vegetable Recognition and Classification Based on Improved VGG Deep Learning Network Model [J]. International Journal of Computational Intelligence Systems, 2020, 13 (1): 559-564.
- [33] 马越.基于深度学习的果蔬检测与分类方法的研究与实现[D].北京邮电大学,2018.
- [34] Gill S H ,Khalaf I O ,Alotaibi Y , et al. Fruit Image Classification Using Deep Learning [J]. Computers, Materials & Continua, 2022, 71 (3): 5135-5150.
- [35] Pennington A J ,Fisher A R . Classification of fruits and vegetables [J]. Journal of Food Composition and Analysis, 2008, 22 (Suppl.1): S23-S31.
- [36] Raikar M M ,M S M ,Kuchanur C , et al. Classification and Grading of Okra-ladies finger using Deep Learning [J]. Procedia Computer Science, 2020, 171 (C): 2380-2389.

- [37] 许学斌,赵雨晴,路龙宾,张佳达.基于 LBP 的粒子群优化混合核函数 ELM 的蔬菜水果图像分类方法[J].机械设计与研究, 2021 , 37(4):15-20, 25.
- [38] Jayakumari R ,Nidamanuri R R ,Ramiya M A . Object-level classification of vegetable crops in 3D LiDAR point cloud using deep learning convolutional neural networks [J]. Precision Agriculture, 2021,22(5): 1617–1633.
- [39] Steinbrener J ,Posch K ,Leitner R . Hyperspectral fruit and vegetable classification using convolutional neural networks [J]. Computers and Electronics in Agriculture, 2019, 162 364-372.
- [40] Ling Z ,Zhenbo L ,Chen L , et al. High performance vegetable classification from images based on AlexNet deep learning model [J]. International Journal of Agricultural and Biological Engineering, 2018, 11 (4): 190-196.
- [41] Ma Y ,Ding X . Real-Time Multi-View Face Detection and Pose Estimation Based on Cost-Sensitive AdaBoost [J]. Tsinghua Science & Technology, 2005, 10 (2): 152-157.
- [42] Tan M P ,Sebastien L ,Francois M . Attribute Profiles on Derived Textural Features for Highly Textured Optical Image Classification [J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15 (7): 1125-1129.
- [43] Naranjo-Torres J, Mora M, Hernández-García R, et al. A review of convolutional neural network applied to fruit image processing[J]. Applied Sciences, 2020, 10(10): 3443.
- [44] Lu D, Weng Q. A survey of image classification methods and techniques for improving classification performance[J]. International journal of Remote sensing, 2007, 28(5): 823-870.
- [45] 张志龙.基于计算机视觉的水果分级方法研究[D].贵州财经大学,2020.
- [46] Kai Z , Zuo W , Chen Y , et al. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising[J]. IEEE Transactions on Image Processing, 2016, 26(7): 3142-3155.
- [47] Xiang Q, Wang X, Li R, et al. Fruit Image Classification Based on MobileNetV2 with Transfer Learning Technique[C]//Proceedings of the 3rd international

- conference on computer science and application engineering. 2019: 1-7.
- [48] Xu B ,Wang N ,Chen T , et al. Empirical Evaluation of Rectified Activations in Convolutional Network. [J]. CoRR, 2015.
- [49] Yonis G . Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique [J]. Sustainability, 2023, 15 (3): 1906-1906.
- [50] Xiaojun J ,Jun C ,Yong C . Weed Identification Using Deep Learning and Image Processing in Vegetable Plantation [J]. IEEE ACCESS, 2021, 9 10940-10950.
- [51] Hameed K ,Chai D ,Rassau A . A comprehensive review of fruit and vegetable classification techniques [J]. Image and Vision Computing, 2018, 80 24-44.
- [52] Wenjun W ,Chao S .Deep learning-based detection and condition classification of bridge steel bearings[J].Automation in Construction,2023,156.
- [53] Migayo M D ,Kaijage S ,Swetala S , et al.Automated Optimization-Based Deep Learning Models for Image Classification Tasks[J].Computers,2023,12(9).
- [54] Raheel S .Fruit-classification model resilience under adversarial attack[J].SN Applied Sciences,2021,4(1).
- [55] Yi C ,Jun B ,Chao K . Application of machine vision and convolutional neural networks in discriminating tobacco leaf maturity on mobile devices [J]. Smart Agricultural Technology, 2023, 5.
- [56] Deqin X ,Jianzhao F ,Tanyu L , et al. Classification and recognition scheme for vegetable pests based on the BOF-SVM model [J]. International Journal of Agricultural and Biological Engineering, 2018, 11 (3): 190-196.
- [57] Mukhriddin M ,Azamjon M ,Jinsoo C . Improved Classification Approach for Fruits and Vegetables Freshness Based on Deep Learning [J]. Sensors, 2022, 22 (21): 8192-8192.
- [58] M. J R ,Fernando D ,Pablo R , et al. Explainable machine learning for performance anomaly detection and classification in mobile networks [J]. Computer Communications, 2023, 200 113-131.
- [59] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Springer. European Conference on Computer Vision. Munich:

- Springer, 2018: 3-19.
- [60] Jie, Shen, Samuel, et al. Squeeze-and-Excitation Networks.[J]. IEEE transactions on pattern analysis and machine intelligence, 2019.42: 7132–7141.
- [61] Qin X, Zhang Z, Huang C, et al. U2-Net: Going Deeper with Nested U-Structure for Salient Object Detection [J] . Pattern Recognition,2020,106: 107-404.
- [62] T/GSQA 012-2020,甘肃省质量协会团体标准—高原夏菜等级划分[S].甘肃: 甘肃省质量协会, 2020.
- [63] Q. W ,B. W ,P. Z , et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [J]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020, 11531-11539.
- [64] Guiwen J . Security Detection Design for Laboratory Networks Based on Enhanced LSTM and AdamW Algorithms [J]. International Journal of Information Technologies and Systems Approach (IJITSA), 2023, 16 (2): 1-13.

攻读硕士学位期间从事的科研工作及取得的成果

参与课题：

甘肃省自然科学基金项目“深度学习在高原夏菜质量分级中的应用研究”
(21JR1RA283)

发表论文：

- [1] 郭泓, 尚庆生, 赵薇, 韩运龙. 基于贝叶斯算法的中文新闻标题分类研究[J]. 现代信息科技, 2023,7(23):89-92.
- [2] 巩晴, 尚庆生, 郭泓, 韩运龙. 芒果图像的改进 EfficientNet 模型分级[J]. 宜宾学院学报, 2022,22(12):1-5.

致谢

行于至此，落笔为终。从 2021 年的秋天到 2024 年的夏天，是时候告别了。从对第一年的热切期待到现在的不情愿和怀旧，这标志着我生命中的一个重要篇章。这段旅程中的经历可能不会被铭记一生，但毫无疑问，它们将被我永远珍惜，用一生来感激。

盛行千里，不忘师恩。首先，向我的导师尚庆生老师表示最深切的感谢。你的理解、宽容和指导，无论是在学术上还是在日常生活中，都是无价的。从最初的论文选题到一丝不苟的修改定稿，都有您孜孜不倦的耐心和支持指引着我。我们无数次的交流、讨论、无数文件的传阅，创造了充满感激和灵感的记忆，让我获得了以前未知的知识和技能。感谢管理科学与工程专业的所有老师们，你们无私的教导加深了我对所学领域的了解，祝愿老师们在职业和个人生活中取得成功和幸福。

慈母手中线，游子身上衣。我要感谢家人对我生活上的照顾，感谢你们支持我的学术努力，并给予我坚定不移的精神鼓励。多年来，你们无条件的爱护和坚定的支持，以及那些细致入微的关怀，共同构筑了超过二十年的温馨记忆。对我而言，家始终是最安心的避风港，是面对挑战与压力时，我最渴望归依的所在。我只希望父母能够健康快乐，长寿安康。

珍惜相逢，无需言欠。对于我的室友及实验室中的朋友们，我深表感激。初次相遇时的惊喜与喜悦，随着时间的流逝依旧温暖人心。在疫情期间，我们共同度过的时光、面对困难时的相互支持与鼓励，都铸就了无价的回忆。在这茫茫人海中能够遇见你们，实属幸运。我坚信，每一次的告别，都是为了下一次更美好的相逢。愿你们的未来光明璀璨，如繁花绽放。

以梦为马，不负韶华。最后，我想感谢我自己。感谢自己面对机会时从不怯懦；面对困难时从不畏惧；面对磨难时从不妥协；面对任务时从不迟疑。我坚信所有经历都是成长，希望以后始终保持对生活的热爱，不忘初心，砥砺前行，不负家人，不负自己。