

分类号 TP391.1  
UDC

密级  
编号 10741



## 硕士学位论文

论文题目 基于深度学习的高原夏菜质量分级研究

研究生姓名: 巩晴

指导教师姓名、职称: 尚庆生 副教授

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: 2023年6月6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： 巩晴 签字日期： 2023.5.20

导师签名： 尚庆生 签字日期： 2023.5.20

## 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： 巩晴 签字日期： 2023.5.20

导师签名： 尚庆生 签字日期： 2023.5.20

# **Research on quality grading of plateau summer vegetables based on deep learning**

**Candidate : Gong Qing**

**Supervisor: Shang Qingsheng**

## 摘 要

随着社会的进步与发展,人们对蔬菜、水果等生活必需品的质量越来越重视,需求量也逐渐增大,但由于种种诸如销售环节无标准化、从事蔬菜经营企业对蔬菜质量分级的意识较弱等原因,使蔬菜售卖时品质与价格不能对等,不能做到优质优价。当前,高原夏菜已成为甘肃省的支柱性特色产业,但其依旧采用的是传统的人力分拣或者普通机械分拣装置等方式,这些低效的方式不仅会消耗大量的人力、物力与成本,并且产出的蔬菜质量不能得到保证,从而造成消极影响。随着深度学习热度的不断上涨,研究基于深度学习的高原夏菜质量分级方法对甘肃省的高原夏菜销售也具有重要的意义和广泛的应用价值。因此,本文针对采用传统分级方法所出现的问题,提出了利用深度学习研究高原夏菜质量分级的方法,主要内容如下:

(1) 构建高原夏菜数据集。针对目前高原夏菜数据集缺少的问题,建立了一套高原夏菜图像采集装置,采集了结球甘蓝、娃娃菜、菜花、西兰花四种蔬菜作为原始数据集,共计 2400 张。

(2) 扩充原始数据集。为保证采集的高原夏菜数据集在模型训练时能有良好的分级效果,本文首先对图像进行粗分割,排除背景干扰;接着使用数据增强的方法对原始图像进行扩充,使其更有利于后续卷积神经网络模型的训练。

(3) 提出了针对高原夏菜质量分级的多尺度融合 CA-Ghost-EfficientNet 模型。首先构建出轻量化模型,将 EfficientNet 网络模型中的第一层卷积层替换为 Ghost 层,同时,在网络的最后一层前嵌入 CBAM 轻量化注意力模块,使网络模型训练速度加快的同时,也能让网络更加关注细微的特征,检测和定位局部有用的信息,使得相似物种的分析更加精确。接着将网络中 MBConv 结构中的注意力机制 SE 模块替换为 CA 模块,使网络能够同时保留特征的长期依赖关系与精准的位置信息。最后改进网络的优化算法,使用 RAdam 算法并结合多尺度融合算法,使网络能够保留数据集图像的更多特征,同时也避免了网络陷入局部最优的现象。

本文通过将预处理之后的高原夏菜数据集,在改进的 EfficientNet 模型与经典神经模型上进行训练,并进行消融对比实验,实验结果表明,本文所提方法的准确率与模型参数量都明显优于其他网络与未改进网络,证明了改进网络模

型的有效性与可行性。

**关键词：**高原夏菜 图像质量分级 EfficientNet 网络 注意力模块

## Abstract

With the progress and development of society, people pay more and more attention to the quality of vegetables, fruits and other daily necessities, and the demand has gradually increased, but due to various reasons such as the lack of standardization in the sales link and the weak awareness of vegetable quality grading in vegetable trading enterprises, the quality and price of vegetables cannot be equal, and high quality and price cannot be achieved. At present, plateau summer vegetables have become a pillar characteristic industry in Gansu Province, but they still use traditional manual sorting or ordinary mechanical sorting devices, which will not only consume a lot of manpower, material resources and costs, but also the quality of the vegetables produced cannot be guaranteed, resulting in negative impact. With the rising popularity of deep learning, the study of the quality grading method of plateau summer vegetables based on deep learning is also of great significance and wide application value for the sales of plateau summer vegetables in Gansu Province. Therefore, in view of the problems arising from the traditional grading method, this paper proposes a method to study the quality grading of plateau summer vegetables by deep learning, the main contents are as follows:

(1) Build a dataset of plateau summer vegetables. In view of the current lack of plateau summer vegetable dataset, a set of plateau summer

vegetable image acquisition device was established, and four vegetables, kale, baby cabbage, cauliflower and broccoli, were collected as the original data set, with a total of 2400 pictures.

(2) Enrich the original dataset. In order to ensure that the collected plateau summer vegetable dataset can have a good grading effect during model training, this paper first coarsely segmented the image to eliminate background interference. Then, the data augmentation method is used to augment the original image to make it more conducive to the subsequent training of convolutional neural network models.

(3) A multi-scale fusion CA-Ghost-EfficientNet model for the quality grading of plateau summer vegetables is proposed. Firstly, a lightweight model is constructed, the first convolutional layer in the EfficientNet network model is replaced by the Ghost layer, and at the same time, the CBAM lightweight attention module is embedded before the last layer of the network, which accelerates the training speed of the network model and allows the network to pay more attention to subtle features, detect and locate locally useful information, and make the analysis of similar species more accurate. Then, the attention mechanism SE module in the MBConv structure in the network is replaced with the CA module, so that the network can retain the long-term dependence of features and accurate location information at the same time. Finally, the optimization algorithm of the network is improved, and the RAdam algorithm is combined with

the multi-scale fusion algorithm, so that the network can retain more features of the dataset image, and also avoid the phenomenon that the network falls into local optimum.

In this paper, the improved EfficientNet model and classical neural model are trained on the preprocessed plateau summer vegetable dataset, and the ablation comparison experiment is carried out, and the experimental results show that the accuracy and number of model parameters of the proposed method are significantly better than other networks and unimproved networks, which proves the effectiveness and feasibility of the improved network model.

**Keywords:** Highland Natsu; Image quality grading; EfficientNet Network; 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.....	2
1.2.2 基于深度学习的蔬菜图像分类.....	4
1.2.3 基于 EfficientNet 网络的图像分类.....	5
1.3 研究内容和创新点.....	6
1.3.1 研究内容.....	6
1.3.2 研究框架.....	7
1.3.3 创新点.....	8
1.4 本章小结.....	9
2 相关理论研究.....	10
2.1 深度学习概述.....	10
2.2 卷积神经网络的基本概念.....	11
2.3 几种典型的卷积神经网络.....	17
2.3.1 VGGNet 网络简介.....	17
2.3.2 ResNet 网络简介.....	17
2.3.3 DenseNet 网络简介.....	19
2.4 注意力机制模块.....	21
2.4.1 SENet 模块.....	21
2.4.2 CA 模块.....	22
2.5 优化算法.....	23
2.5.1 SGD 优化算法.....	23
2.5.2 Adam 优化算法.....	24
2.6 深度学习框架.....	25
2.7 本章小结.....	25
3 高原夏菜图像数据集的构建及预处理.....	26

3.1 高原夏菜图像数据集的构建.....	26
3.1.1 图像采集.....	26
3.1.2 数据集存在问题.....	27
3.2 数据预处理.....	28
3.2.1 高原夏菜图像分割背景移除处理.....	28
3.2.2 高原夏菜图像空间几何变换操作.....	30
3.3 蔬菜等级划分原则.....	31
3.4 本章小结.....	32
<b>4 基于深度学习高原夏菜质量分级模型构建 .....</b>	<b>33</b>
4.1 EfficientNet 的网络结构.....	33
4.1.1 基本结构.....	33
4.1.2 网络参数.....	34
4.2 轻量化模型构建.....	35
4.2.1 替换轻量化模块.....	35
4.2.2 嵌入轻量注意力模块.....	37
4.3 替换的注意力机制模块.....	38
4.4 网络超参数的设置.....	39
4.4.1 优化算法.....	39
4.4.2 损失函数.....	40
4.5 CA-Ghost-EfficientNet 多尺度融合网络模型构建 .....	40
4.6 本章小结.....	43
<b>5 高原夏菜质量分级的实验结果与分析 .....</b>	<b>44</b>
5.1 算法流程.....	44
5.2 高原夏菜数据集划分.....	45
5.3 实验环境与评价指标.....	45
5.3.1 实验环境及参数设置.....	45
5.3.2 评价指标.....	46
5.4 结果与分析.....	47
5.4.1 与原网络模型性能对比.....	47

5.4.2 与经典深度卷积神经网络模型性能对比.....	49
5.4.3 消融实验对比.....	51
5.5 本章小结.....	52
<b>6 总结与展望.....</b>	<b>53</b>
6.1 工作总结.....	53
6.2 工作展望.....	54
<b>参考文献.....</b>	<b>55</b>
<b>攻读硕士学位期间从事的科研工作及取得的成果 .....</b>	<b>55</b>
<b>致谢.....</b>	<b>62</b>

# 1 引言

## 1.1 研究背景和意义

在我国的国民经济中，农业始终占据着非常重要的地位。随着经济全球化的不断发展，农产品的种类与交易量都在不断增加，销售终端也迎来了历史性的变化<sup>[1]</sup>。高原夏菜作为甘肃省的支柱性特色产业，产品具有优异的品质，是农民脱贫致富的重要产业，在二十多年的发展过程中，其能够有效地解决东南部城市的蔬菜出现供不应求的问题，发展势头良好，得到各级政府相关业务部门政策支持<sup>[2]</sup>。

近年来，菜农们自发地对农业种植结构进行调整，增加了蔬菜作物的种植面积，目前在一些地方，高原夏菜已经成为人们收入的主要来源和当地的支柱产业。当前高原夏菜生产环节的标准化已经逐步实施，但销售环节的标准化仍然滞后，经营蔬菜行业公司大多都将蔬菜进行冷藏存储并统一运输，未直接面向终端市场，对产品进行分级处理的意识较弱，多采用传统的分级分拣方式，比如依靠人力进行分拣<sup>[3]</sup>，不仅会耗费巨大的人力成本，并且产出的蔬菜品质良莠不齐，平均品质被拉低，致使本来能畅销、能卖出好价钱的蔬菜也以较低的价格被收购、包装，直接影响了总体销售价格，不宜进行大规模生产。

与传统的人工检测、识别分类技术相比，采用计算机视觉进行图像识别、检测、分类不仅能够提高效率，同时也能提高准确度。当前，计算机视觉技术被普遍地应用在蔬菜水果的分类、植物与农作物的病虫害识别、药片的整缺度识别等方面中，其能够快速定位识别检测中所需要的特征，实现更加高效、经济的提取。国外学者利用计算机视觉技术对农产品的外观品质检测研究较早，成果颇丰，主要集中于对谷物、干果、水果、鸡蛋等的检测。国内计算机视觉技术在农产品上的应用起步较晚，但也取得了一些成果。近年来随着深度学习技术在图像识别领域的巨大突破，以 VGGNet、GoogleNet、ResNet 等为代表的卷积神经网络模型，不仅在大规模计算机视觉挑战赛取得非常好的成绩，而且也有很多的学者将这些模型应用在蔬菜水果的识别分类、农作物病害识别等领域，都取得了较好的识别效果，这也为蔬菜图像的识别方法提供了新思路 and 理论可行性。

因此，为了减少对高原夏菜质量等级分类时所消耗的人力、物力与成本，本文提出了基于深度学习的高原夏菜质量分级方法，建立高原夏菜分级图像资源库，构建出的高原夏菜数据集可填补当前领域的空白，不仅可供本文的研究，还可进一步发布公开；提出针对高原夏菜质量分级的多尺度融合 CA-Ghost-EfficientNet 模型，不仅能够节省人力物力，减少人工成本，同时也能够提高蔬菜分级性能，加快蔬菜分级的速度。本文研究基于深度学习的高原夏菜质量分级，预期形成一套快速、高效、性能优异的蔬菜分级方法，丰富农产品分级这一研究领域的理论及应用，研究成果可用于设计制造蔬菜自动识别和分级流水线。

## 1.2 研究现状

国外在 20 世纪 80 年代就已经集中研究果蔬的品质检测与分类，随着计算机视觉的发展，将计算机技术应用在果蔬分类领域上的研究也越来越多。当前，国外在蔬菜识别领域上的研究系统已经较为成熟，并且在超市里得到了应用，成功地售卖出蔬菜水果。但国内的蔬菜分类大多依靠人工，耗时费力效率低，与国外还存在很大的差距，随着科技的发展进步以及众多学者对蔬菜水果识别的研究，在不久的将来，我国在蔬菜水果识别领域必将发展迅速，更加高效可靠。

果蔬的图像分类技术流程主要分为四个步骤：首先将图像输入进网络模型中，接着对输入的图像进行预处理，使网络能够提取到更多的图像特征，最后进行图像的分类。从上述的四个步骤出发，果蔬的分类经历了基于机器学习的图像分类研究<sup>[4-8]</sup>、基于深度学习的图像分类研究<sup>[9-11]</sup>，随着计算机视觉的发展，利用计算机视觉对果蔬进行颜色分拣分级已成为保持产品质量、提高产品价值的主要途径<sup>[12-16]</sup>。

### 1.2.1 基于机器学习的蔬菜图像分类

基于传统图像处理的蔬菜图像分类通过对蔬菜图像的预处理，进而进行特征选择，通过颜色特征、纹理特征、几何特征等对蔬菜进行分类与分级。

1996 年，Bolle<sup>[17]</sup>对蔬菜的颜色、纹理等特征进行提取，从而进行蔬菜图像的分类，但该系统提取特征时容易受到外界光的干扰，影响识别的准确度。同年国内，刘禾<sup>[18]</sup>设计了一套果形自动判别系统，该系统基于传统的神经网络，其数

数据集采用了两种不同种类的苹果, 通过使用该网络模型将苹果的果形判别准确度提升至 80% 左右。2010 年, Rocha 等人<sup>[19]</sup>描述了利用直方图、颜色和形状描述符合无监督学习方法从图像中自动进行水果和蔬菜分类。同年, 李金伟<sup>[20]</sup>提出了一种快速灰度截取分割阈值法提取马铃薯表面深色部分和十色模型进行缺陷检测。同年, 张亚静等<sup>[21]</sup>首次开发了番茄图像采集机器视觉系统, 利用机器视觉方法, 通过番茄表面的颜色、纹理特征和 BP 神经网络从“定性”和“定量”两方面预测了番茄品质。实验结果表明, 利用机器视觉方法对番茄生长阶段进行预测较为可行。2012 年, Danti 等人<sup>[22]</sup>为了形成特征向量, 他们首先对图像进行裁剪和调整大小, 然后提取 HSV 图像的色调和饱和度通道的平均值和范围, 利用 BPNN 分类器对 10 类叶菜进行分类, 分类成功率为 96.40%。同年, Suresha 等人<sup>[23]</sup>使用分水岭分割提取感兴趣的区域作为预处理和决策树分类器进行训练和分类, 利用 RGB 颜色空间中的纹理测度, 在包含 8 种不同蔬菜的数据集上获得了 95% 的分类准确率。Dubey 等人<sup>[24]</sup>在 2015 年将图像经过分割后提取不同的颜色和纹理特征, 并将它们结合起来, 实验证明, 将两者结合起来比单独的颜色和纹理特征具有更好的识别效果。同年, Madgi 等人<sup>[25]</sup>提出了一种基于 RGB 颜色和局部二值模式纹理特征的蔬菜分类方法。王水华等人<sup>[26]</sup>提出了一种基于适应度尺度混沌人工蜂群算法和前馈神经网络的方法, 该网络提取颜色、形状和纹理作为特征, 达到了相当好的准确度。熊俊涛等人<sup>[27]</sup>在 2017 年提出以主成分向量作为分类特征, 分析设计了 SVM 多分类系统, 改进了分类判别函数, 整体样本分类识别率达到 96.6%。2018 年, Wan 等人<sup>[28]</sup>提出了一种将特征颜色值与反向传播神经网络分类技术相结合的番茄成熟度检测方法, 将番茄表面最大内切圆作为颜色特征提取区域, 随后划分为同心圆半径不同的五个特征颜色子区域, 最后通过三层 BP 神经网络对两个品种不同的番茄进行分类, 准确率达到 99.31%。人工感官评价合理化分析往往是乏味和耗时的, 黄兴义等人<sup>[29]</sup>在 2019 年提出了基于多感官数据融合的 BPNN 模型, 该模型大大提高了检测精度, 采用机器学习无需任何化学预处理就能对菠菜新鲜度进行快速无损检测, 节省人力。2020 年, 芦范<sup>[30]</sup>通过优化径向基函数神经网络、量子遗传算法删除了冗余权重和神经元获得了优化后的神经网络模型, 该模型对蔬菜图像的形状、纹理、颜色特征识别率均有提升, 优于其他算法。

综上,基于传统图像处理的蔬菜分类方法大多数都会在对图像进行预处理后,对颜色、纹理、形状等特征提取从而对蔬菜进行检测及分类,线性分类器和 KNN 分类器在传统的分类中占有重要的地位,由于分类器需要提取大量的特征才能使训练达到较优的效果,因此会导致在运行中内存占用大,计算时间长,限制了方法的发展和准确性。虽然采用机器学习的方法进行蔬菜图像识别与分类,可以提高分类的准确度,但该种方法在多种蔬菜识别分类中不能得到很好的应用。

### 1.2.2 基于深度学习的蔬菜图像分类

深度学习的概念<sup>[30-33]</sup>起源于人工神经网络,在特征学习和表达方面表现出了优异的性能,其通过结合低层特征形成更抽象的高层特征,从而发现数据的分布特征,提高图像的识别准确率。

2019 年曾平等<sup>[34]</sup>通过改进 LeNet-5 模型对水果图像进行分类,不仅提高了识别准确率,而且也避免了传统人工提取特征的繁琐过程,提高了效率。2020 年, Raikar 等人<sup>[35]</sup>对秋葵手指分类分级,对比 AlexNet、GoogLeNet、ResNet50 三种模型并使用迁移学习训练网络,虽然 ResNet50 消耗了最大的训练时间,但其准确率远远高于其他两种模型。同年,薛刚等人<sup>[36]</sup>提出了一种基于注意的卷积自编码器(CAE-ADN)密集连接卷积网络,该网络使用卷积自编码器对图像进行预训练,使用基于注意的 DenseNet 提取图像特征,可以大大提高水果分拣的效率,从而降低生鲜供应链、工厂、超市等的成本。李等人<sup>[37]</sup>为提高蔬菜的自动识别和分类精度,提出改进 VGG 网络模型来训练蔬菜图像数据集,将前两个全连接层(VGG-M)的输出特征结合起来,并在网络中加入批量归一化层。经过实验验证,在测试数据集上对分类识别的正确率高达 96.5%。

2021 年,付月生等人<sup>[38]</sup>对 GoogLeNet 进行优化,减少卷积内核数量,调整 Inception 的结构,并通过引入 Swish 激活函数和 DropBlock 层,使 GoogLeNet 的测试准确率提高了 2%,改进后的模型在准确率、召回率等方面都优于其他模型。同年,许学斌等人<sup>[39]</sup>提出了 LBP-PSO-KELM 的分类方法以解决实时高精度对蔬菜水果进行分类的问题,该方法在 Fruits-360 数据集上进行实验,不仅减少了模型计算时间,同时提高模型准确率,达到 98.8658%; Gill 等人<sup>[40]</sup>利用卷积神经网络、递归神经网络和长短时记忆建立了一个多模型水果图像识别系统, CNN 通

过不同的卷积层提取图像特征,利用 RNN 对不同特征进行标记,最后利用 LSTM 对提取的最优特征进行分类,实验证明,此分类技术优于单独使用 CNN、RNN、RNN-CNN 的图像分类技术。Jayakumari 等人<sup>[41]</sup>提出了一种从 3D 角度对作物进行语义分割的深度卷积神经网络,对点云中的作物植物对象进行高效的特征学习和基于对象的分割,实验结果表明,茄子和白菜的植株水平作物识别精度最高达 90%左右。

此外,注意力机制近年来得到了广泛的应用。过去,深度学习模型的焦点是整个输入图像,而注意力机制的焦点是输入图像中的目标区域,将焦点从整体转移到局部。Li 等人<sup>[42]</sup>使用 ResNet18、ResNet34 和 VGG16 模型添加残块,对苹果病害进行识别和分类。实验结果表明,网络层数较少的 ResNet18 具有最好的识别能力。Son 等人<sup>[43]</sup>提出了一种结合 FS-SubNet 和 SAC-SubNet 的叶片识别点注意机制,FS-SubNet 采用新的训练数据库,并将子网与 SAC-SubNet 连接,实现了苹果叶片病害的识别。实验结果表明,病害识别正确率达到 89.4%,提取了更多的鉴别特征和语义点信息,提高了苹果叶片病害识别正确率。

### 1.2.3 基于 EfficientNet 网络的图像分类

EfficientNet 是一种新的卷积神经网络,对比之前的卷积神经网络,该网络对网络的深度、宽度、分辨率设置固定的尺度缩放系数统一进行缩放,以其高参数效率和速度<sup>[44]</sup>而闻名。

2021 年,王振亚等人<sup>[45]</sup>针对不合格药片检测问题,采用简化后的 EfficientNet 网络,并将网络中的 SENet 模块替换为 ECA 模块,避免了因降维所导致的副作用,该改进模型对药片分类的准确率达到 98.93%,使不合格药片可以更加高效地被检测出来。关胜琪等人<sup>[46]</sup>提出了一种利用改进的 GAN 和高效 Net 进行带钢缺陷分类的算法,构建标签反卷积网络,将图像标签逐层解卷积得到条件掩码,叠加到生成器和判别器中形成 Mask - CGAN,对模式搜索生成对抗网络进行改进,实验显示 Mask - CGAN 能够生成真实的生活图像,解决了深度学习中样本不足的问题,改进的高效 Net 能够准确高效地对带钢缺陷进行分类。



2022 年，光金正等人<sup>[47]</sup>针对植物图像的分类，对 EfficientNet 网络中原激活函数 Swish 与激活函数 Mish、ReLU 进行比较，发现相较于前两者，激活函数 Mish 的图像更加平滑，且在小于零的情况下允许神经网络中存在更小的负梯度，同时采用迁移学习，其分类准确率为 97.2%，较改进前的网络准确率提高了 0.4%。王芳等人<sup>[48]</sup>针对垃圾数据集采用 U<sup>2</sup>-Net 物体显著性检测模型与 EfficientNet 迁移学习结合构建模型，该模型的识别精度达到 94.2%，垃圾分类效率得到了显著的提升；Jaisakthi 等人<sup>[49]</sup>使用了基于迁移学习技术的 EfficientNet 架构，同时使用了 Ranger 优化器，对皮肤镜图像中的皮肤病变进行分类，通过 Ranger 优化器对 EfficientNet - B6 进行优化微调，获得了 96.81% 的准确率。姜天宇等人<sup>[50]</sup>提出了基于 EfficientNet 的木薯叶病变自动分类模型，引入了标签平滑和随机丢弃节点，同时采用双稳态逻辑损失进行预测，改进后的模型在木薯叶病变公共数据集上的准确率达到 89.66%。帖军等人<sup>[51]</sup>将 EfficientNet 中的 SENet 核心模块替换为多分支卷积网络制 SKNet，用以识别番茄叶片病害，实验结果表明，该模型的平均准确率可达到 99.64%，且在自然场景中，准确率较原模型也提高 3.81%。

综上，在关于 EfficientNet 网络模型对图像分类中，大多数会从激活函数、注意力机制模块、优化器等方面对原始网络进行改进，同时引入迁移学习，将权重预先在公开数据集上训练好，随后引入到改进模型中，对参数进行微调，提升网络的学习性能。

## 1.3 研究内容和创新点

### 1.3.1 研究内容

根据以上国内外研究现状，在参考众多对于蔬菜分类的文献后，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，本文对高原夏菜质量分级的研究内容主要体现在以下几个方面：

(1) 高原夏菜数据集的构建。针对目前缺乏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，通过采集娃娃菜、西兰花、结球甘蓝和菜花四种高原夏菜的分级图像数据，经过预处理和标注，构建出蔬菜分类、分级标

准数据集。

(2) 数据集预处理。对原始数据集进行扩充，以达到模型训练的要求，从而得到好的模型训练结果。首先对图像进行粗分割，排除背景干扰；接着使用数据增强的方法对原始图像进行处理，实现样本增强，有助于避免模型过拟合。

(3) 高原夏菜质量分级模型构建。针对本论文中自建的高原夏菜数据集，提出了一种结合 CA 注意力机制、Ghost 模块、CBAM 模块和多尺度融合的 EfficientNet 分级模型。将改进模型与经典的卷积神经网络在自建数据集上进行训练，并进行消融实验，对比模型之间的准确率与参数，实验结果表明，改进模型对高原夏菜质量分级具有良好的效果。

### 1.3.2 研究框架

本文共有六个章节，各个章节的内容安排如下，其组织结构如图 1.1 所示：

第一章：引言，详细阐述了甘肃省高原夏菜的主要背景及对其进行质量分级的意义，主要介绍了基于机器学习与深度学习的蔬菜图像分类研究现状，以及基于 EfficientNet 网络的图像分类国内外研究现状，并提出了本文的研究内容及论文创新点；

第二章：相关理论研究，主要介绍了蔬菜图像分类分级时用到的理论基础，包括深度学习、卷积神经网络的基本概念、神经网络架构、注意力机制模块、优化算法和深度学习框架。

第三章：高原夏菜图像数据集的构建及其预处理，主要介绍了蔬菜图像的采集、预处理、分级原则等，构建了高原夏菜分级数据集。

第四章：基于深度学习的高原夏菜模型构建。首先对原网络 EfficientNet 的基本结构进行了介绍，接着对改进 EfficientNet 网络的注意力机制模块、优化算法展开叙述，并构建出用于本文实验所用数据集多尺度融合的 CA-Ghost-EfficientNet 模型。

第五章：高原夏菜质量分级的实验结果与分析。将改进的模块在高原夏菜数据集进行训练，通过准确率、参数量以及  $F_1$  值来验证改进模块的有效性，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。对本文的研究内容进行总结，并提出下

一步展望。

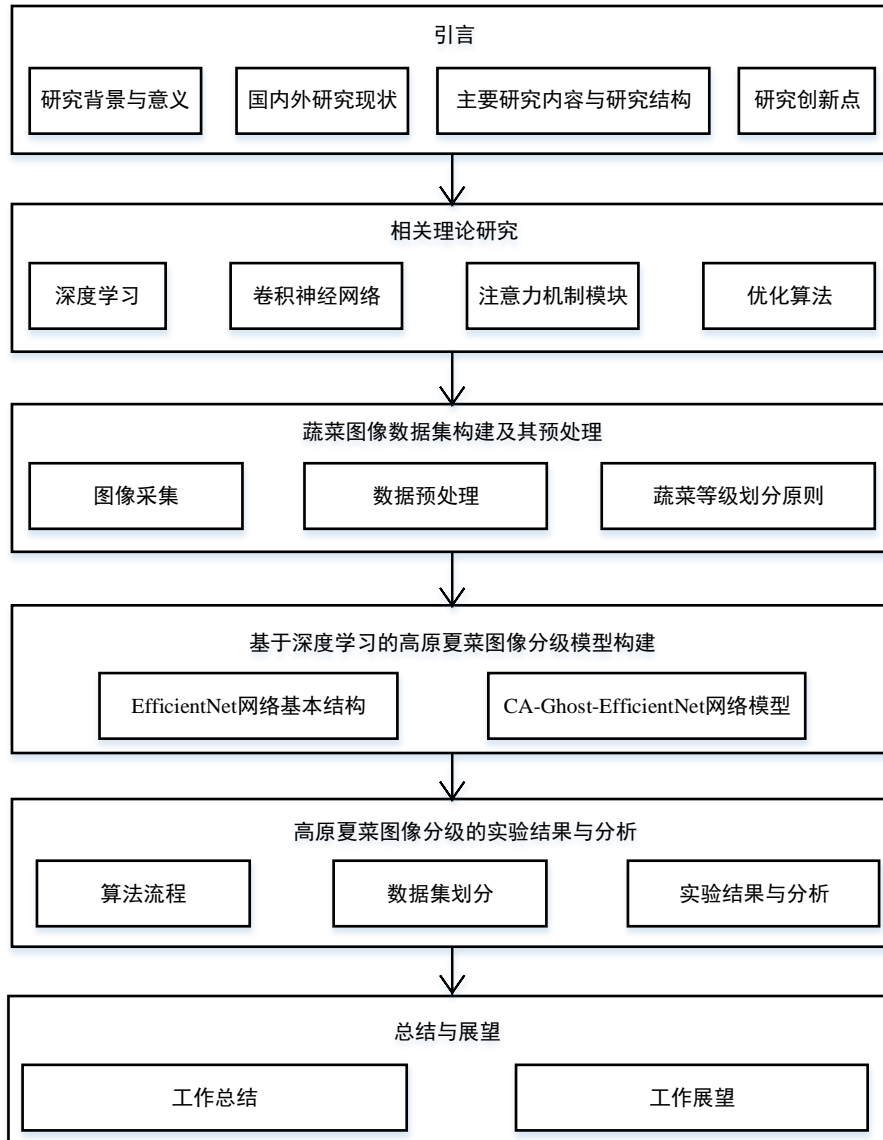


图 1.1 研究组织结构

### 1.3.3 创新点

(1) 构建了高原夏菜分级数据集。针对高原夏菜质量分级的需求，构建出用于高原夏菜质量分级研究的数据集，同时采用数据增强的方法将原始数据集图像进行扩充，以达到网络模型学习训练时的要求。

(2) 提出了一种高原夏菜分级算法。结合 CA 注意力机制、Ghost 模块、

CBAM 模块和多尺度融合的 EfficientNet 网络,提出针对高原夏菜质量分级方法。改进的网络模型使得网络更加关注细微的特征,能够检测和定位局部有用的信息,同时也使得相似物种的分析更加精确,为蔬菜分类与分级研究提供新方法和思路。

## 1.4 本章小结

首先,本章阐明了对蔬菜分类进行研究的背景以及意义;然后对前人在基于机器学习、深度学习的蔬菜分类领域,以及使用 EfficientNet 网络进行图像分类所做的研究工作进行了总结,有助于为下一步蔬菜数据集的构建和模型的调优找准切入点和着力点,提供研究思路和依据;最后对文章的主要研究内容、结构安排与创新点进行介绍。

## 2 相关理论研究

本文提出的高原夏菜质量分级模型基于卷积神经网络的深度神经网络模型，采用深度学习对图像进行分类是近年来蔬菜图像分类分级中最为常用的方法。本章将分别介绍蔬菜图像分类分级时用到的理论基础，包括深度学习、卷积神经网络的基本概念、神经网络架构、注意力机制模块、优化算法和 Pytorch 深度学习框架。

### 2.1 深度学习概述

人工智能、机器学习与深度学习分别属于继承和发展的关系，人工智能能够将人的思想、行为动作通过计算机传达出来，使计算机具有拟人化的特点，机器学习与深度学习是通过将大量的数据输入进网络模型中，使其在训练中提高完成任务的能力，该技术现已广泛地应用在语音、疾病、图像识别与检测中。

深度学习实质上是通过大量的数据进行训练，并在模型中添加多个隐藏层，使网络学习到更多特征，从而使网络模型的分类准确率得到提高。如图 2.1 所示，浅层神经网络结构相较于深层神经网络的结构较为简单，中间的隐藏层一般仅有 1 层或 2 层的非线性转换层，但在样本数据集非常少、遇到比较复杂的函数时，性能表现却不佳，在遇到相对复杂的图像分类问题时，其网络的泛化能力也不能完全体现。深度神经网络与传统的浅层神经网络的不同在于<sup>[52]</sup>：

(1) 更加注重模型结构的深度。隐藏层的层数多于两层，其中包含了大量相互连接的神经元，层数越多拟合函数越强，从而对输入图像的特征进行提取，使其划分不同类型的数据；

(2) 明确突出了表示学习的重要性。深度学习的学习能力强于浅层神经网络，能够学习到数据更多的特征，在模型训练的过程中，每一个神经元之间的权值是通过不断调整直至调整到最优的参数，使整个模型的训练结果达到最优，从而使网络的性能得到提高，可以将复杂的函数用较少的参数表示<sup>[53]</sup>。

(3) 深度学习在提高网络准确率时，并不需要运用更加复杂的方法，在训练网络模型时，通过训练大量的数据集，提高网络的性能，使准确率提升，且迁移学习也增强了网络的适应性，易于转换。

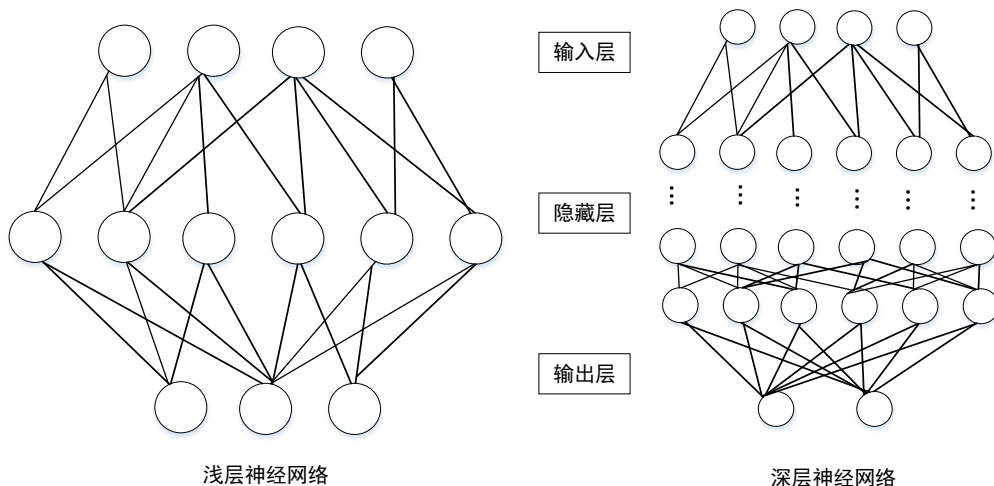


图 2.1 浅层与深层神经网络结构

## 2.2 卷积神经网络的基本概念

传统的卷积神经网络（Convolutional Neural Network, CNN）属于端到端结构，它的连接方式为逐层连接，这种连接方式使网络中每一层只响应前一层的神经元，并且能够自动提取图像的特征信息，成功地应用在了各个领域。一个深度卷积神经网络模型中，包含了卷积层、池化层以及全连接层，以经典的数字识别卷积神经网络 LeNet-5<sup>[54]</sup>为例，其结构示意图如图 2.2 所示，其首先将图像输入进网络模型中，经过两次卷积之后进入到池化层，最后经过全连接层对结果进行输出。

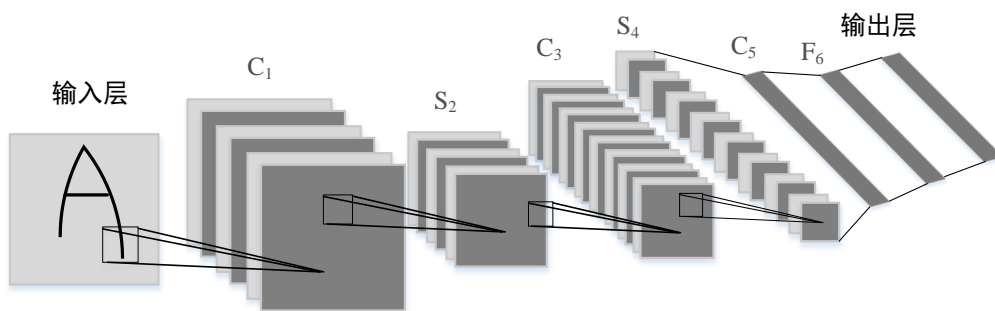


图 2.2 LeNet-5 模型结构

### (1) 卷积层

卷积层是由众多的卷积单元所组成，并通过反向传播算法去优化卷积单元中的参数。随着卷积层数的增加，通过卷积层所提取到的特征也越来越多，网络也

能够学习到更加复杂的特征。卷积操作简而言之就是对输入的图像进行特征提取，其可以看作是将两个矩阵进行相乘，即将卷积核所对应的输入样本区域与卷积核进行相乘，从而得出输出特征图，即卷积核与输入样本上与之相对应的感受野区域，通过内积运算，即可得到输出特征图的神经元的值，之后通过移动感受野，就可得到输出特征图。

关于卷积的操作结构示意图如图 2.3 所示，最底下一层为输入样本，其大小为  $5 \times 5$ ，中间一层为卷积核，其大小为  $3 \times 3$ ，通过该卷积核对输入样本进行卷积操作，得到最上面一层的输出特征图，其大小与卷积核的大小一样，但小于输入样本的大小。如果想要保持输出特征图的大小不变，则需要对输入样本的高和宽两侧填充元素，通常为 0 元素，填充的大小为  $P = (F - 1) / 2$ ， $F$  对应卷积核的大小。

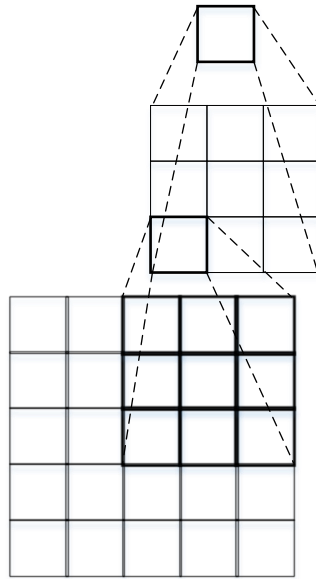


图 2.3 卷积操作示意图

如图 2.4 所示，卷积层  $l$  的第  $j$  单元的输出值  $a_j^l$  的计算为公式 2.1，图中为一个  $3 \times 3$  的卷积核，通过进行内积运算，从而得出输出特征图上神经元的值。卷积操作是为了使网络模型可以更加全方位地提取到目标特征，对边缘与复杂的特征保持同等重视，随着层级的增加，提取的特征也越来越复杂。

$$a_j^l = f \left( b_j^l + \sum_{i \in M_j^l} a_j^{l-1} * k_{ij}^l \right) \quad (2.1)$$

其中  $a_j^l$  为卷积层  $l$  的第  $j$  单元的输出值,  $b_j^l$  为偏置,  $M_j^l$  为对应感受野区域输入特征图,  $k$  为卷积核。

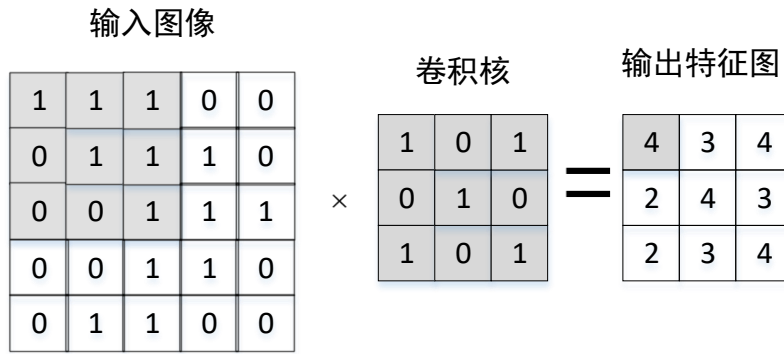


图 2.4 卷积运算示意图

### (2) 批归一化层

通常情况下,在浅层神经网络中,对数据进行归一化预处理就已经很有效了,但对于深层神经网络,即使在输入数据时已经进行归一化的操作,但在神经网络中存在内部协变量偏移,即在每次神经网络中的前一层更新参数时,每一个隐藏单元的输入分布都会发生变化,这使得网络模型在训练结果输出时因参数的更新导致不够稳定,很难训练出有利于任务的模型,同时也会降低模型训练时的速度。

BN (Batch Normalization) 层<sup>[55]</sup>的提出是为了解决模型在训练时所出现的各种问题,其也被称为批归一化层。在深度模型的训练过程中,批量归一化利用小批量上的均值和标准差,不中断地调节神经网络中间输出值,从而使整个神经网络各层中间的输出值能更平稳。批归一化层一般插到卷积层与全连接层的后面,非线性处理之前,当使用批量梯度下降时,为使每一维成为标准的高斯分布(均值为 0, 方差为 1),在批量的维度下对前一层的输出值进行归一化,即

$$\hat{X}_i^t = \frac{X_i^{t-1} - E(X^{t-1})}{\sqrt{D(X^{t-1}) + \varepsilon}} \quad (2.2)$$

其中,  $X_i^{t-1}$  代表前一层输出批量中的第  $i$  个,设置  $\varepsilon$  是为了避免出现零除现象出现。



为了能使批归一化层工作在激活的非线性区,需要在对其进行缩放和移位处理,如公式 2.3 所示,  $\beta, \gamma$  为模型中学习参数,使用其进行缩放和移动归一化分布;  $h$  为神经网络输出的前一层。其具体的运算流程公式如下表 2.1 所示:

$$BN(h; \gamma, \beta) = \beta + \gamma \frac{h - \hat{E}(h)}{\sqrt{Var + \varepsilon}} \quad (2.3)$$

表 2.1 BN 层算法实现步骤

输入: 批处理 输入  $x: B = \{x_1, \dots, m\}$

输出: 规范化后的网络响应  $\{y_i = BN_{\gamma, \beta}(x_i)\}$

1:  $\mu_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i$  //计算批处理时数据的均值

2:  $\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_b)^2$  //计算批处理时数据的方差

3:  $\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}}$  //进行规范化处理

4:  $y_i \leftarrow \gamma \hat{x}_i + \beta = BN_{\gamma, \beta}(x_i)$  //进行尺度变换、偏移操作

5: 返回学习的参数  $\beta$  和  $\gamma$

### (3) 激活函数层

激活函数层是为了帮助卷积神经网络学习所输入数据中的复杂模式,其核心是通过对卷积层的输出,进行简单的非线性映射,从而使整个网络模型的复杂性增加,同时提升了整个网络的非线性刻画能力<sup>[56]</sup>。如果不使用激活函数,不做非线性操作,在多个层输入后依然进行的是线性变换,模型的适应能力不好,表征能力差。

激活函数能够让网络具有非线性的特点,提高网络学习复杂特征的能力。常见的激活函数大都连续可微,通常被分为两类,第一类被称为饱和函数,其中包括 Sigmoid 函数、Tanh 函数等;第二类被称为非饱和函数,包括 ReLU 函数,以及通过它变换得到的 Leaky ReLU 函数等。

以上两种饱和函数的数学定义如公式 2.4, 2.5 所示, Sigmoid 函数通常会出现在二分类的场景中,其值域在 0 与 1 之间,一般作用于输出层; Tanh 函数是

一个奇函数，函数的图像关于原点对称，其值域在-1 到 1 之间，通常用于隐藏层。虽 Tanh 函数优于 Sigmoid 函数，但两者都需计算指数，复杂度高，而且两者的饱和神经元依然会引起梯度消失现象。

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2.4)$$

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

在深度学习中，经常使用的非饱和函数是 ReLU 函数，该函数可以有效地解决训练时出现的梯度消失现象，是因为该函数可以在  $x > 0$  时保持梯度不衰减。ReLU 函数的计算效率高，在实际应用中收敛速度相较上述两个饱和函数快很多，其定义如下。

$$f(x) = \max(0, x) \quad (2.6)$$

ReLU 函数虽可以缓解过拟合现象的出现，但当  $x < 0$  无梯度时，会导致权重无法更新。Leaky ReLU 函数的提出避免了 ReLU 可能出现的神经元不被激活的现象，其数学表达式如 2.7 所示，其中  $\alpha$  一般取值为 0.01。

$$\sigma(x) = \max(\alpha x, x) \quad (2.7)$$

#### (4) 池化层

在卷积网络中，卷积层之后一般为特征提取的池化层，将经过卷积层所得到的特征结果进行整合，从而得到新的特征输出，也就是对特征图进行特征压缩。例如，对一个区域大小为  $2 \times 2$  的特征图进行池化操作，其经过池化后所得到的新的特征图大小为之前大小的  $1/4$ 。池化层的作用分为以下三点：第一，能够将主要特征保存下来，同时减少模型的计算量，使其更方便优化，防止过拟合现象的产生；第二，保障特征的不变性，保证特征不因预处理时的操作产生变化；第三，进行特征降维，池化操作简单来讲就是通过改变原有维度的大小，使模型保留更多的特征，同时通过前向传播，减小特征图进入到下一层时的大小，从而减少模型参数个数与计算量。常用的池化操作有平均池化和最大值池化。

如下图所示，平均池化的前向传播是将卷积层的输出不重复的进行平均值下采样，从而得到平均池化操作后的值。其反向传播则与前向传播相反，将一个值

去平均分到四个格子里，以保证池化前后的梯度之和保持不变。

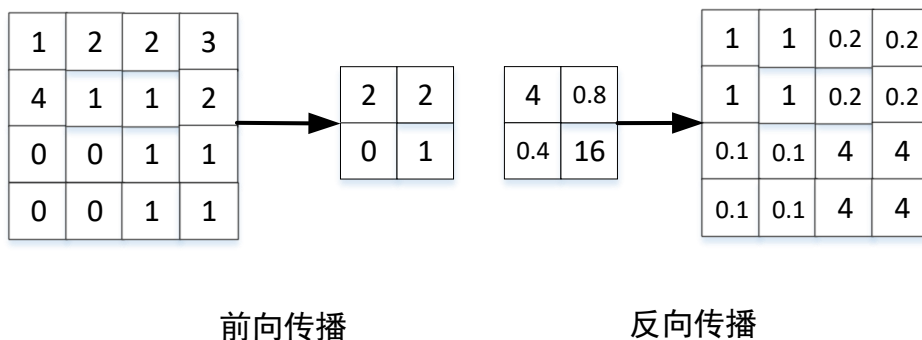


图 2.5 平均池化过程

如下图所示，是最大值池化的两种过程，其前向传播是通过上一层的输出进行不重复的下采样，而其反向传播则是只关注区域内的最大值，将残差传递到该最大值的位置，其他的区域置零。

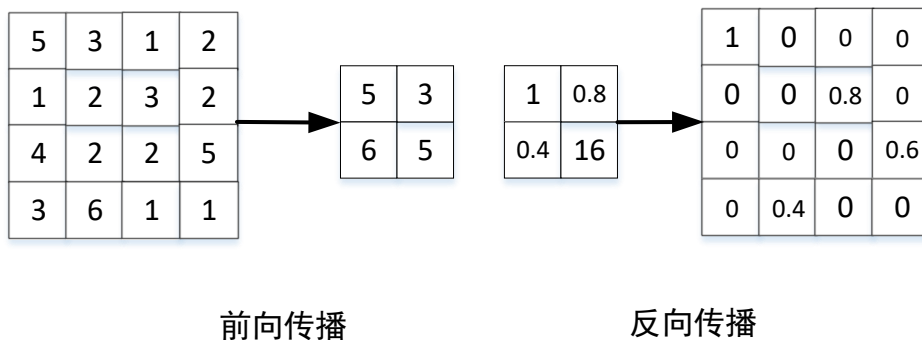


图 2.6 最大值池化过程

### (5) 全连接层

卷积神经网络的最后一层一般为全连接层，其每个结点都与上层的结点相连，故而成为全连接层（Full Connected Layers, FC），该层将上一层即池化层输出的特征图整合转化成一维的特征向量，将所形成的特征向量传递到输出层，实现了图像的分级分类与端到端的学习过程。

## 2.3 几种典型的卷积神经网络

### 2.3.1 VGGNet 网络简介

VGGNet 是牛津大学与 Google DeepMind 公司的研究员一起合作开发出的深度卷积神经网络，应用于图片分类。该网络的卷积层与其他网络有一个明显区别，其特征图的空间分辨率与通道数是反比关系，前一个单调递减，后一个单调递增，从而可以让输入的数据图像可以得到更好的转换。该网络将卷积层进行重复堆叠，使其构成卷积层组，提高感受野范围，能够增强网络模型的学习与特征表达能力。以 VGGNet16 网络为例，其结构图如图 2.7 所示。

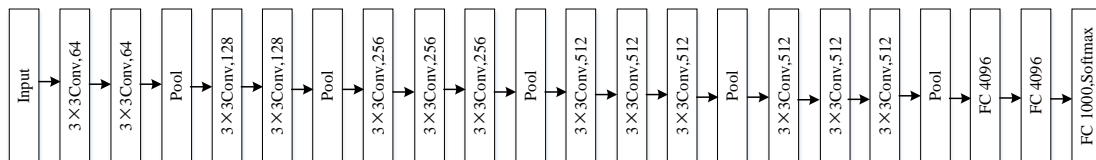


图 2.7 VGGNet16 网络结构图

从上图可以看出该网络总共有二十一个阶段，其中共有十三个卷积层阶段，五个最大池化层阶段，以及三个全连接层阶段。该网络采用了卷积组来扩大卷积核的感受野范围，不仅能够减少参数的数量，同时也能够增强非线性变化，使特征的提取能力增强。同时，VGGNet 网络也通过增加或减少网络的层数，改变网络的结构，使网络能够适应不同的模型训练需求。

### 2.3.2 ResNet 网络简介

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以为使分类的准确率得到提高，获得更好的学习效果，但将浅层网络进行直接堆叠，形成深层网络，并不能提高分类准确率，反而会出现梯度消失、梯度爆炸等问题，不能提高模型的优化效果。虽然可以通过标准初始化和正则化解决上述所出现的问题，使其保证网络在反向传播时采用随机梯度下降，但在遇到更深的网络时，效果却更差了。

针对上述情况, Kaiming He 等人提出了深度残差网络 ResNet<sup>[57]</sup>, 该网络相较于其他的卷积神经网络引入了恒等映射, 通过计算残差以解决层数过多引起的退化问题, 其模型核心是通过建立出上一层与下一层之间的“短路连接”, 有助于训练出更深的网络, 残差网络示意图见图 2.8 所示。残差网络中的残差块可以表示为  $H(x) = F(x) + x$ , 其中  $H(x)$  为网络输入  $x$  的期望输出, 当  $F(x) = 0$  时,  $H(x) = x$ 。残差网络通过在基准网络(VGGNet)中插入一个 shortcut connections 来表示, 即“恒等快捷连接”, 使网络能够跳跃层级, 同时使网络的参数量与计算量保持不变, 并且整个网络模型依然可以进行端到端的训练。

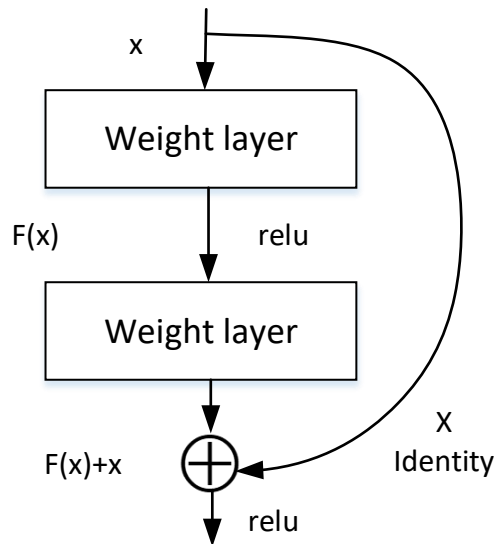


图 2.8 残差网络示意图

ResNet 网络依旧沿用了 VGGNet 的  $3 \times 3$  卷积层设计, 每个残差块具有两个  $3 \times 3$  卷积层, 周期性地滤波器数目加倍并且使用步幅为 2 的空间下采样。残差网络中残差块有两种设计, 如图 2.9 所示, 对待不同的网络层数选择相应的模块, 网络的参数减少了很多, 训练也就相对容易一些。残差网络是由一系列残差块组成的, 每个残差块定义如下:

$$y = F(x, \{W_i\}) + x \quad (2.8)$$

其中  $x, y$  分别为残差块的输入、输出向量;  $F(x, \{W_i\})$  为要训练的残差映射,

且上式中  $x$  和  $F$  的维度需要一致。

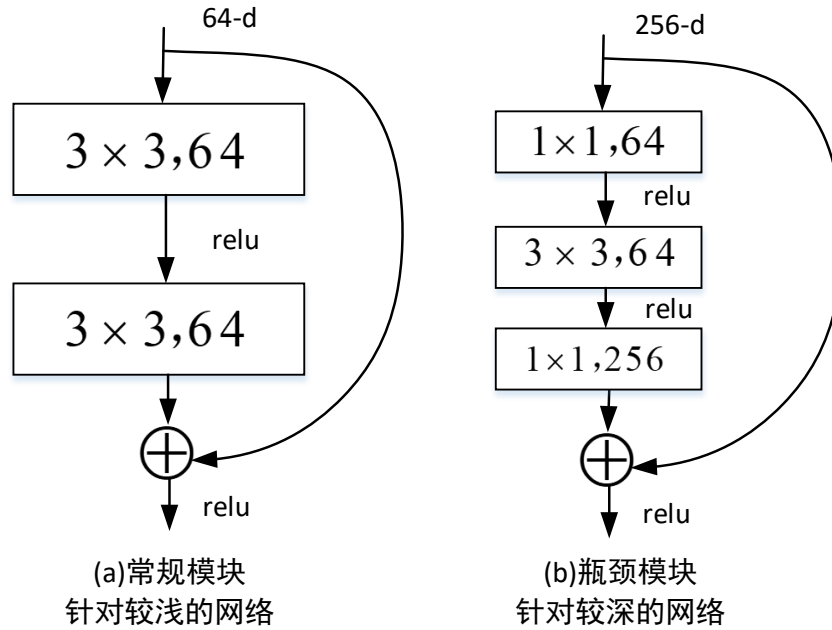


图 2.9 基础残差模块

### 2.3.3 DenseNet 网络简介

相比于 ResNet，DenseNet 模型独立于网络结构的深化和拓宽，提出密集连接机制，使其实现层级之间的直接连接，从而去提高网络性能<sup>[58]</sup>，简而言之就是每个层去接收所有前一层的特征图，因此，所有图层都可以直接连接具有匹配的特征图大小到其他图层。DenseNet 开发了一个完整的连接模块即密集块，由多个完全连接的层即密集层组成。密集层的结构如图 2.10 所示，首先对输入进行批量归一化，在处理激活函数后，通过  $1 \times 1$  卷积调整通道数量，并应用  $3 \times 3$  卷积，以实现再次批量归一化后的特征降维。

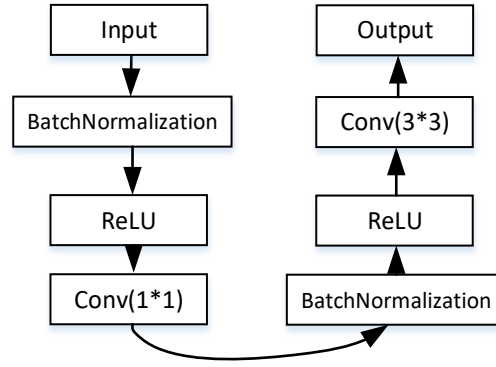


图 2.10 密集区块结构图

密集块使其内部的每一层最终都能从所有层的输出中获得梯度，这使得特征传递函数更加有效，减轻了梯度的消失。密集区块的表达式如下：

$$X_l = H_l([X_0, X_1, \dots, X_{l-1}]) \quad (2.9)$$

其中  $X_l$  是卷积层的输出， $H_l$  是由卷积层实现的非线性函数。DenseNet 在激活功能之前对输入数据执行批量标准化。批处理归一化层处理的数据满足正态分布，可以有效避免数据分布的偏差，远离激活函数的导数饱和区域。批量归一化层可以加快训练过程并有效地减少消失梯度。主要计算过程如下：

$$\hat{x}^{(k)} = \frac{x^{(k)} - E(x^{(k)})}{\sqrt{\text{Var}(x^{(k)})}} \quad (2.10)$$

当卷积核的数量较少时，DenseNet 仍能很好地工作，即网络在每次卷积过程中增加了  $1 \times 1$  卷积运算，这不仅减少了参数，还集成了每个信道的特性。在密集块之后添加过渡层，该结构还执行  $1 \times 1$  卷积运算以融合密集块中所有层的输出，以进一步减少网络参数。

DenseNet 为诸如医学诊断之类的分类问题提供了可接受的结果，该网络成功的主要原因是它比一些常见的网络更深，并且其优化更容易。DenseNet 具有以下几个优点，缓解了消失的梯度问题，加强特征传播，鼓励特征重用，并且现有网络相比参数更低。

## 2.4 注意力机制模块

### 2.4.1 SENet 模块

SENet(Squeeze-and-Excitation Networks)<sup>[59]</sup> 由 Hu 等人提出, 在最后一届的 ImageNet2017 竞赛分类任务获得第一名。卷积通常侧重于空间尺度信息的融合。通过引入注意力机制, SENet 专注于不同通道之间的联系, 以便了解每个通道特征的重要性。SENet 网络的创新点是关注信道之间的关系, 希望模型可以自动学习到不同信道特征的重要程度。SENet 的结构如图 2.11 所示, 包含两个操作: 挤压(Squeeze)和激励(Excitation), 可以通过挤压操作获得特征图中每个轨迹的全局特征; 通过激励可以了解通道之间的关系, 并可以获得不同通道之间的权重。

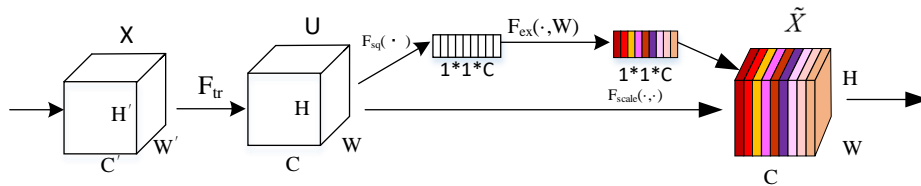


图 2.11 SE 模块

挤压操作, 通常将一个通道上整个空间特征编码为一个全局特征, 采用全局平均池化操作来实现:

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j), z \in R^c \quad (2.11)$$

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:

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \text{ReLU}(W_1 z)) \quad (2.12)$$

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 模块的输入特征层被全局合并, 之后进行两次完全连接。第一次完全



连接的神经元数量较少，第二次完全连接神经元数量与输入特征层相同。在两个完整连接之间设置 ReLU 层，随后执行另一个 sigmoid 以将值固定在 0-1 之间。此时，获得输入特征层的每个通道的权重值。最后，可以通过将该权重值乘以原始输入特征层来获得加权特征层。

## 2.4.2 CA 模块

SE 模块计算通道注意力需要通过 2D 全局池化，并在低计算成本下取得了优异的表现，但该模块并未考虑位置信息的重要性。CBAM 模块获取位置信息时采用减少通道维度的方法，但这种方法在长距离依赖关系中并不能获得很好的效果。CA 模块<sup>[60]</sup> (Coordinate Attention) 是一种高效的轻量级注意力模块，该模块的提出使网络在保持低计算成本的同时，也能够更大的区域上进行注意力，可以任意插入各个卷积神经网络中，提高了网络模型的特征表达能力。该模块的提出，有效地解决了的问题，能够分别沿着水平坐标和垂直坐标两个方向，将其的输入特征利用两个一维的全局池化操作，使用池化核的两个空间范围对其进行编码，聚合为一对方向感知特征图，这一对方向感知特征图里存在着沿着不同空间方向特征的长距离依赖与精确位置信息。

坐标注意力机制的编码操作如图 2.12 所示，首先，先将全局池化进行分解，沿着两个方向聚合特征，转化为一对一维的特征编码，分解的数学公式如 2.13 所示，其中  $H, W$  分别代表特征图的高度及宽度， $c$  则代表通道数。随后，对其进行拼接并使用卷积核大小为  $1 \times 1$  的卷积核进行变换操作，计算公式如 2.14 所示，图 2.12 中的  $r$  为缩放因子。接着，沿着空间的维度将  $f$  分成两个单独的张量，并经过卷积核大小为  $1 \times 1$  对其进行卷积操作，输出得到最后的特征图  $g^h, g^w$ 。最后，注意力坐标模块的输出如公式 2.15 所示。

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (2.13)$$

$$f = \delta(F_1[z^h, z^w]) \quad (2.14)$$

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (2.15)$$

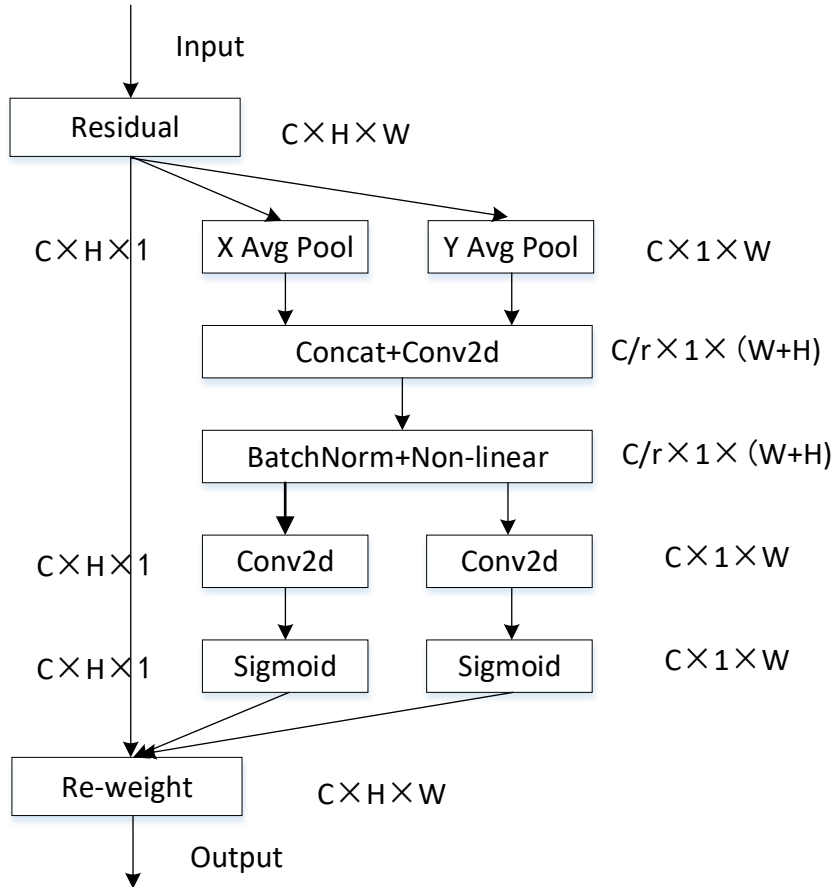


图 2.12 坐标注意力机制

## 2.5 优化算法

### 2.5.1 SGD 优化算法

在机器学习中经常会出现一个问题，只要大的数据集才能够得到比较好的泛化，但这样会增加计算成本，也会增加训练的时间。随机梯度下降法也被称为SGD (Stochastic Gradient Desent) [61]，该优化算法的核心是期望用较小的样本近似估计，在每一次的迭代计算中，随机抽出一组样本，训练计算小样本的梯度，随后进行参数权重的更新，然后继续这个过程，随机抽取后计算梯度，而且在每一次的迭代过程中，都会随机打乱样本。其数学表达式如下，其中  $\eta$  为学习率， $g_t$  为梯度方向。

$$\begin{cases} g_t = \nabla_{\theta_{t-1}} f(\theta_{t-1}) \\ \Delta \theta_t = -\eta * g_t \end{cases} \quad (2.16)$$

随机梯度下降是最基本的优化方法，将样本每一次的随机打乱更新，能够加快训练速度，同时支持参数等的在线更新，但每一次更新只使用一个样本，会造成下降速度不快，相较其他的优化算法需要更多的计算成本，而且在收敛时会在沟壑两边持续震荡，容易出现局部最优现象，在沿着误差梯度向下学习的过程中，也很难确定误差最小值。

## 2.5.2 Adam 优化算法

自适应学习优化算法（Adaptive Moment Estimation, Adam）<sup>[62]</sup>对 SGD 优化算法进行扩展，该方法计算每一个参数或权重的自适应学习率，当在面对大量数据和参数时，会非常有效。该算法能够在处理稀疏梯度问题中占据优势，也能够维护每一个参数的学习速率，在处理非平稳目标时占据优势。同时该算法在均方根传递算法的基础上对梯度滑动平均及偏差纠正进行改进，其算法实现步骤如下表所示。

表 2.2 Adam 算法实现步骤

---

确定参数学习率  $\alpha = 0.001$ ，一阶和二阶矩衰减系数  $\beta_1 = 0.9, \beta_2 = 0.999$  以及损失函数  $f(\theta)$ ，初始化各参数向量。

---

当参数  $\theta$  不收敛时：

- 1:  $t \leftarrow t + 1$  //更新步数
  - 2:  $g_t \leftarrow \nabla_{\theta} f_t(\theta_{t-1})$  //计算原目标函数对参数  $\theta$  的梯度
  - 3:  $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$  //计算梯度的一阶矩
  - 4:  $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$  //计算梯度的二阶矩
  - 5:  $\hat{m}_t \leftarrow m_t / (1 - \beta_1^t)$  //对一阶矩进行校正
  - 6:  $\hat{v}_t \leftarrow v_t / (1 - \beta_2^t)$  //对二阶矩进行校正
-

续表 2.2 Adam 算法实现步骤

---

7:  $\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \hat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon)$  //更新参数  $\theta_t$

Return  $\theta_t$

---

其中,  $t$  为步数,  $m_t, v_t$  分别为一阶矩、二阶矩, 即梯度  $g_t, g_t^2$  的期望,  $\beta_1^t, \beta_2^t$  分别是指  $\beta_1, \beta_2$  的  $t$  次方,  $\hat{m}_t, \hat{v}_t$  为偏置修正, 考虑到两个分别在零初始值情况下的向零偏置。在更新参数  $\theta_t$  时, 将  $\hat{m}_t$  视为更新梯度,  $\alpha / (\sqrt{\hat{v}_t} + \varepsilon)$  视为更新参数的学习率。

## 2.6 深度学习框架

随着机器学习、深度学习的近几年的爆炸式发展, 在书写代码时不可避免地会出现, 为了提高实验时的效率, 众多研究者们将所需的基础代码整合至一个框架内, 以供使用。开发有关深度学习的项目, 选择一个契合的框架是至关重要的, 不但可以减少工作量, 还能节省时间。开源深度学习框架随着各大公司高校的加入也变得越来越, 目前最为流行的深度学习框架有 Tensorflow、Pytorch、Keras、Caffe、Torch 等<sup>[63]</sup>。从 GitHub 网站上, 根据各大框架的讨论度以及各大公司、顶级会议、众多比赛的选择来看, Tensorflow 和 Pytorch 的使用人数以及讨论热度等已经远远超过其他深度学习框架。

本文中实验所用的深度学习框架为 Pytorch, 与深度学习框架 TensorFlow 不同, PyTorch 库使用动态更新的图形进行操作, 灵活性较高, 速度较快。其训练神经网络的过程简洁直观, 其源码也比其他框架更易理解, 并且 Pytorch 库中包含着许多预先训练好的模型, 能够在强大的 GPU 加速基础上实现张量和动态神经网络, 能够简化算法的优化, 将更多的精力放在算法的实现上。

## 2.7 本章小结

本章首先阐述了深度学习的基本概念, 然后对卷积神经网络中的五个基本结构进行了介绍, 随后对注意力机制以及优化算法进行阐述, 最后对深度学习的框

架做出介绍。

### 3 高原夏菜图像数据集的构建及预处理

目前图像分类分级的数据集主要分为两大类：公开数据集与自制数据集，而针对高原夏菜质量分级的公开数据集却几乎没有。因此，为解决目前高原夏菜图像分类分级中数据集缺少的问题，计划搭建一个用于高原夏菜分类分级的标准数据集，一方面通过设备采集原始高原夏菜图像，另一方面通过图像增强方式扩充高原夏菜图像数据集。结合原始数据与扩充数据，将其共同应用于模型进行训练。

#### 3.1 高原夏菜图像数据集的构建

##### 3.1.1 图像采集

本实验所使用的数据集来源为高原夏菜蔬菜交易场所，确保了实验数据真实可靠。深度学习的研究需要数据做支撑，而具有高度清晰且优质的图像才能确保神经网络模型训练的准确性。因此，为了得到更优质的高原夏菜数据集，本文搭建了实验室环境下的计算机采集图像硬件。

由于手机自带闪光灯能提供稳定的光源，且拍出的图像清晰、分辨率高，足以满足模型训练的数据集要求，因此本实验使用手机作为主要采集设备，采集设备还包括计算机、拍照支架等，如图 3.1 所示。采集流程主要为以下两个步骤：第一步，使用手机拍摄高原夏菜图像，通过闪光灯的开关控制光源；第二步，将所拍摄好的高原夏菜数据集通过数据线传送到计算机上，以供后续的模式训练。

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，每个种类的高原夏菜拍摄 600 张，共采集的高原夏菜数据集为 2400 张。

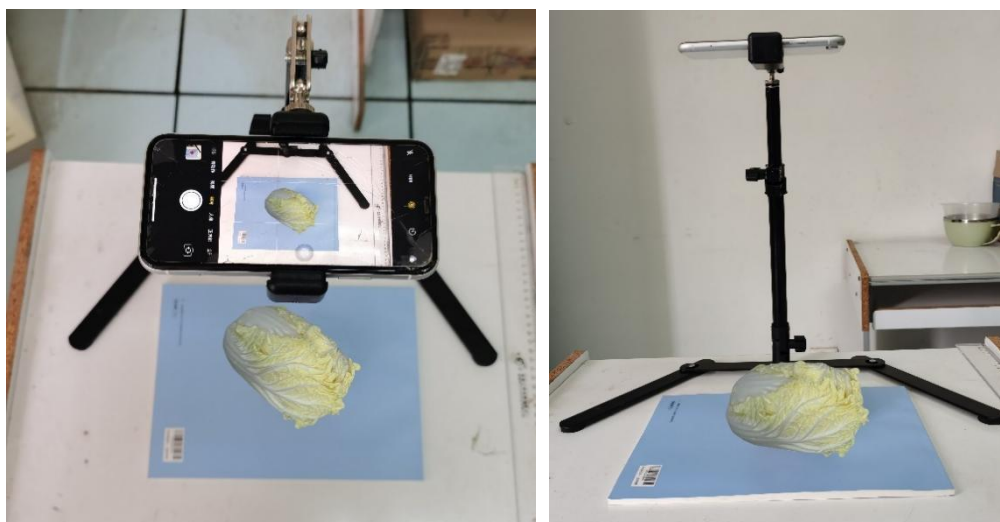


图 3.1 高原夏菜图像采集装置

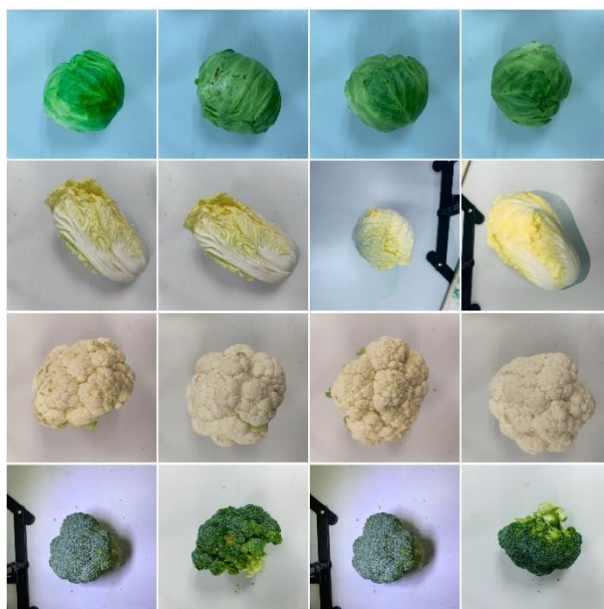


图 3.2 部分初始数据集

### 3.1.2 数据集存在问题

通过上述的拍摄所采集到的高原夏菜数据集会出现以下两个问题：

(1) 背景问题。如图 3.2 的娃娃菜与西兰花数据所示，不同品种在拍摄数据集时，并不能使背景保持统一，在后续提取特征时会造成干扰，影响分级准确性。

(2) 数据集较小。关于高原夏菜公开的数据集几乎没有，且该数据集为人工拍摄，拍摄的图像数据集较少，会导致模型产生过拟合的问题。

因此，为了使模型训练时可以得到一个好的结果，需要对所采集的数据集进

行处理，从而避免特征提取时所造成的干扰与过拟合问题。

## 3.2 数据预处理

由于在图像采集的过程中，会受到外界因素的影响，采集得到的数据集不够标准规范，会降低模型分级时的精度与准确度，因此，在对高原夏菜质量分级之前，需要对图像进行预处理，首先将高原夏菜的背景进行移除，剔除影响后续特征提取的无用信息；随后为了提高模型的泛化能力，将原始数据集进行扩充，以实现此目标。

### 3.2.1 高原夏菜图像分割背景移除处理

为了保留更多特征信息、去除干扰和提高识别率，预处理步骤常常需要对图像进行去噪和分割增强处理。由于拍摄时高原夏菜背景会包含拍照支架等设备因素，会影响到后续特征的提取，使其模型准确率及分级结果受到影响，因此本文采用 U<sup>2</sup>-Net<sup>[64]</sup>模型对目标特征的背景进行剔除，避免由背景所造成的干扰影响后续的分级结果。

U<sup>2</sup>-Net 网络结构是由 Qin 等人提出，其网络结构如图 3.3 所示。它的主要架构是一个 U-Net，就像编码解码一样，每个阶段都由 U 型残差块（RSU）组成，该网络将两个 U-Net 结构进行嵌套堆叠，共包含十一个阶段的 U 型结构，每一个阶段内都采用该模块抓取到多尺度的特征，其主体部分由阶段编码器与显著图融合模块组中，损失函数可定义为

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

其中， $M$  为层数， $l_{side}^m$  ( $M = 6$ ) 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， $\omega_{side}^m$  为对应各个损失函数的权重， $\omega_{fuse}$  为最终输出的预测图对应的权重。

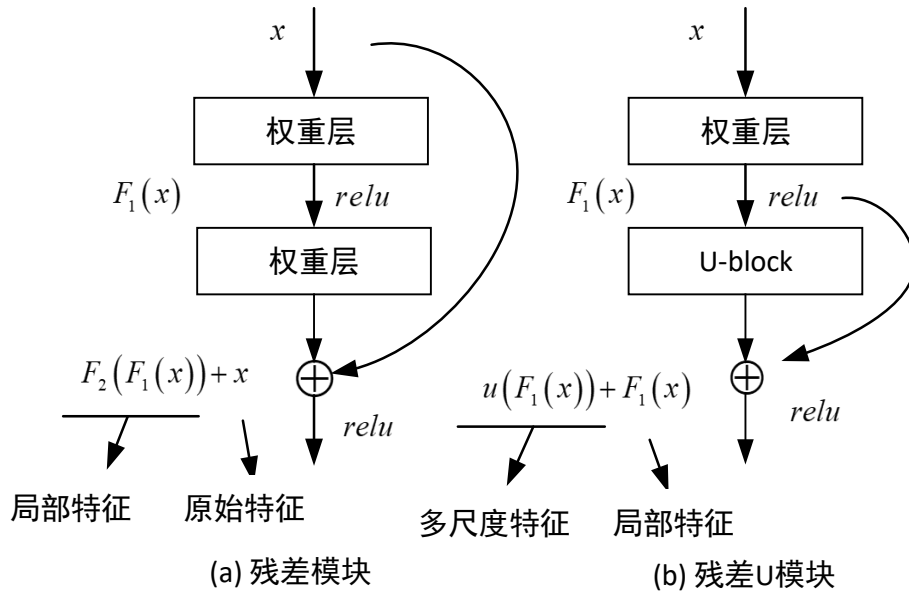


图 3.3 残差模块的网络结构对比

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：

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

其中， $(r,c)$ 是像素的坐标， $(H,W)$ 是图像的高度和宽度， $P_{G(r,c)}$ 是高原夏菜初始的像素值， $P_{s(r,c)}$ 是分割的背景移除图像的像素值。

结合检测结果，将原始图像中物体前景分离，进而进行背景移除，从而使数据集背景一致，处理后结果如图 3.4 所示。



图 3.4 高原夏菜移除背景对比图



### 3.2.2 高原夏菜图像空间几何变换操作

图像空间几何变换是图像增强的一种方法，主要是指通过旋转、平移、翻转等操作对样本数据进行处理，以达到扩充样本数据集的目的。以上述移除背景后的娃娃菜为例，通过旋转、翻转对数据集进行扩充。

(1) 旋转：将图像从原点开始，顺时针地旋转，角度设定为 $\theta$ ，图像中各像素点的坐标变换公式如 3.3 所示。

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} \quad (3.3)$$

公式 3.3 中， $x_0, y_0$  表示旋转前各像素点的横纵坐标， $x$  和  $y$  分别代表图像经过顺时针旋转后所得到的横纵坐标。本实验取 $\theta=6^\circ$ ，即将裁剪后的数据集中的原图每隔 $6^\circ$ 将原图像旋转一次，每张原图可得 86 张不同旋转角度的图像，如图 3.5 所示，为旋转后的部分数据图像，其中左起第一个图示为原图，第二个到第四个为旋转所得图像。



图 3.5 图像旋转示例图

(2) 翻转：在高原夏菜图像数据集中选取部分图像沿竖直、水平两方向进行翻转，翻转图示如 3.6 所示，其中，图中从左到右分别为原图、水平翻转、垂直翻转。



图 3.6 图像翻转示例图

经过上述采用的两种几何变换方式，每个品种的高原夏菜图像都扩增至 900 张，整体数据集扩增至 3600 张。

### 3.3 蔬菜等级划分原则

2020 年甘肃省质量协会颁布了《高原夏菜等级划分标准》(T/GSQA 012-2020)<sup>[65]</sup>对高原夏菜进行等级划分，成为了高原夏菜质量等级判断的权威标准。因此，本文在对高原夏菜质量分级的研究过程中，全程参考此标准，以保证分级结果的可靠性，分级标准如表 3.1 所示。

该标准按照高原夏菜外观的紧实度、大小、鲜嫩度、病虫害感染等情况特点，将每一类高原夏菜分为 3 个等级，即特级蔬菜、一级蔬菜和二级蔬菜，其分类的详细依据主要根据蔬菜大小外观：是否紧实、外体是否完整；表面是否有缺陷；是否有病虫害损伤、机械损伤；花型是否紧实洁白等。

因此，本文将四种高原夏菜的外观结实度、鲜嫩度与有无病虫害感染情况作为研究对象，将原图像与预处理后的图像数据结合后进行统一划分，每类高原夏菜都分为以下三个等级：特级产品、一级产品、二级产品。同时，在该数据集进行模型训练前，需要按照等级对数据进行命名，本文通过数字来代表具体蔬菜等级，具体为：0 代表特级产品，1 代表一级产品，2 代表二级产品。每个种类的高原夏菜命名形式用该蔬菜的大写英文字母所表示，以结球甘蓝为例，其命名形式为“JQGL.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

### 3.4 本章小结

本章主要讲述了如何构建高原夏菜图像数据集，以及图像处理、分级原则。图像采集环节时，为了保证训练数据的统一性，采用统一的高度、光源进行采集，以避免过多外界因素对模型训练结果所造成的干扰；同时，为避免模型在训练时出现过拟合的现象，本论文中采用数据增强的方式对数据集进行扩充，并运用  $U^2Net$  网络进行背景移除，从而提升模型准确率；高原夏菜等级的划分依据甘肃省质量协会团体标准，以蔬菜的大小、质量、颜色、外观等对蔬菜进行层次类别划分。

## 4 基于深度学习高原夏菜质量分级模型构建

通过上一章对高原夏菜数据集的采集与预处理后,需要将高原夏菜图像输入进分级模型中进行等级分类。因此,本章提出了用于高原夏菜质量分级的模型 CA-Ghost-EfficientNet,首先对原网络 EfficientNet 的基本结构进行介绍,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 EfficientNet的网络结构

在传统神经网络中,一般采取增加网络的深度、改变特征提取的层数以及提高输入图像的分辨率三者其一使模型的准确率、效率得到更好的提升,但网络深度加深的同时梯度消失的问题也将出现,图像分辨率的提高也会加大模型的计算量,准确率也随着分辨率的增加而下降。而 Tan 等人提出 EfficientNet 模型<sup>[66]</sup>,是针对图像的深度、宽度及分辨率 3 个维度进行调整,使模型的准确率得到提高。

#### 4.1.1 基本结构

EfficientNet-b0 网络的架构如表 4.1 所示,从表中可以看到该网络总共分为了 9 个阶段,其中第一个阶段的卷积核大小  $3 \times 3$ ,步距为 2 的卷积层,包含 BN (Batch Normalization) 与 Swish 激活函数,第二个到第八个阶段是重复堆叠的 MBConv 结构,第九个阶段是由一个卷积核大小为  $1 \times 1$  的卷积层、一个平均池化层以及一个全连接层所组成。

EfficientNet 网络的深度、宽度和分辨率与其他单一调整其中一个的网络不同,它会设置固定的尺度缩放系数统一进行缩放,从而使得网络在三维环境下的性能更好。其复合比例系数的计算公式如 4.1 所示:

$$\begin{cases} \text{depth} : d = \alpha^\phi \\ \text{width} : w = \beta^\phi \\ \text{resolution} : r = \gamma^\phi \end{cases}, \quad \text{s.t.} \begin{cases} \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \\ \alpha \geq 1, \beta \geq 1, \gamma \geq 1 \end{cases} \quad (4.1)$$

其中:  $d, w, r$  三个因子分别是缩放网络对深度、宽度和分辨率的缩放系数; $\alpha, \beta, \gamma$  三个因子都是由一个很小范围的网络所搜索得到的常量,决定了具体的

分配资源情况； $\phi$  因子是一个特定的系数，通过扩大与减小该系数，将网络的深度、宽度以及输入图像的分辨率进行统一的调节。

表 4.1 EfficientNet-b0 网络的架构

Stage i	Operator $\hat{F}_i$	Resolution $\hat{H}_i \times \hat{W}_i$	#Channels $\hat{C}_i$	#Layers $\hat{L}_i$
1	Conv3×3	224×224	32	1
2	MBCConv1,k3×3	112×112	16	1
3	MBCConv6,k3×3	112×112	24	2
4	MBCConv6,k5×5	56×56	40	2
5	MBCConv6,k3×3	28×28	80	3
6	MBCConv6,k5×5	14×14	112	3
7	MBCConv6,k5×5	14×14	192	4
8	MBCConv6,k3×3	7×7	320	1
9	Conv1×1&Pooling&FC	7×7	1280	1

该模型的主干网络使用了 MobileNet V2 网络中的 MBCConv 结构，如图 4.1 所示，该结构包含了两个卷积核大小为 1×1 的普通卷积层、一个 k×k 的深度转换卷（Depthwise Conv），其中 k×k 有两种结构，分别为 3×3 与 5×5，一个 SE（Squeeze-and-Excitation）模块以及一个 Dropout 层。其中，两个卷积层有不同的作用，分别为升维与降维。

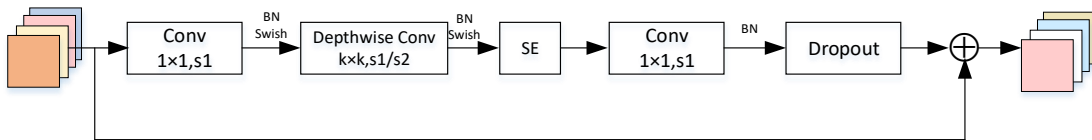


图 4.1 MBCConv 结构图

### 4.1.2 网络参数

将公式 4.1 中  $\alpha$ 、 $\beta$ 、 $\gamma$  固定为常数，放大不同  $\phi$  的基线网络 EfficientNet-b0，得到 EfficientNet-b0 ~ EfficientNet-b7 模型，其最合适的输入参数如下表 4.2 所示，其中 Input\_size 为训练网络时输入图像的大小，Width\_coefficient、

Depth\_coefficient 为宽度、深度两个维度上的倍率因子，Drop\_coefficient 为 MBConv 结构中随机失活层使用的失活率的大小，Dropout\_rate 为全连接层之前的随机失活层的失活率大小。

表 4.2 EfficientNet b0-b7 网络参数

Model	Input_size	Width_coefficient	Depth_coefficient	Drop_coefficient	Dropout_rate
B0	224×224	1.0	1.0	0.2	0.2
B1	240×240	1.0	1.1	0.2	0.2
B2	260×260	1.1	1.2	0.2	0.3
B3	300×300	1.2	1.4	0.2	0.3
B4	380×380	1.4	1.8	0.2	0.4
B5	456×456	1.6	2.2	0.2	0.4
B6	528×528	1.8	2.6	0.2	0.5
B7	600×600	2.0	3.1	0.2	0.5

## 4.2 轻量化模型构建

本小节对原网络模型进行初次改进，构建出了轻量化模型，替换并嵌入了轻量化模块，使改进的网络模型在训练时，能够在减少参数量的同时提升网络模型的准确率。

### 4.2.1 替换轻量化模块

原网络 EfficientNet 模型在训练生成特征图时，会伴随有大量的卷积运算，产生较多的运算量并占据较多的内存，而 Ghost 模块是一种即插即用的模块，且该模块所需的参数总数较少。因此为减少其运算量，本文结合此轻量化模块，将原网络中的第一层卷积替换为该模块，从而提高模型训练时的效率，以 EfficientNet-b0 结构为例，改进的网络模型具体结构如图 4.2 所示。

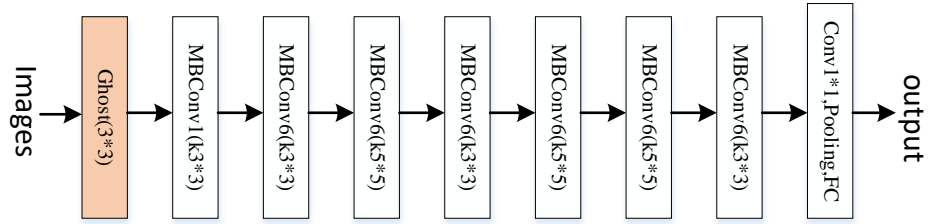


图 4.2 Ghost 层替换结构图

Ghost 模块<sup>[67]</sup>是一种新型轻量化模块，该模块可以通过一系列线性变换将固有的特征图生成出许多 Ghost 特征图，可以替换任何经典卷积神经网络中的卷积操作，从而提高卷积运算的效率。其卷积的过程如图 4.3 所示。

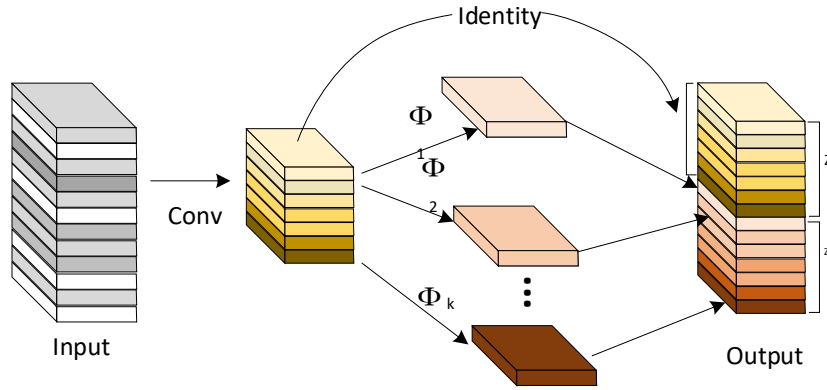


图 4.3 Ghost 模块的卷积过程

上图中，可以看到通过 Ghost 模块得到了两部分特征图，一部分为通过传统卷积所得到的特征图，用  $z$  表示，另一部分为通过线性映射所得到的特征图，用  $z'$  表示，在输出时，将所得到的两个特征图进行拼接，其具体的输出过程如下。

假设使用传统的卷积进行卷积运算得到固有特征图，卷积操作的数学表达式如下。其中  $X \in R^{c \times h \times w}$  为输入特征， $f \in R^{c \times k \times k \times n}$  为卷积核， $Y \in R^{h' \times w' \times n}$  为输出特征， $h', w', k \times k$  分别为输出特征图的高度、宽度以及卷积核的大小， $b$  为偏置项，这个操作中的运算量为  $n \cdot h' \cdot w' \cdot c \cdot k \cdot k$ 。

$$Y = X * f + b \tag{4.2}$$

因为在运算中，网络特征图的生成需要大量参数的参与，而较多的卷积核会

造成较大的计算量，基于上述的问题，Ghost 模块则仅利用一次卷积去生成  $m$  个原始特征图，计算公式如下。式中的  $f' \in R^{c \times k \times k \times m}$  为卷积核。

$$Y' = X * f' \quad (4.3)$$

之后将上述卷积操作所生成特征图，通过线性的映射，得到与之对应的  $n$  个特征图，其数学表达式如下。其中  $y'_i$  代表第  $i$  个的原始特征图， $y_{i,j}$  是指将第  $i$  个特征图进行线性变化得到第  $j$  个幻影特征图。

$$y_{i,j} = \phi_{i,j}(y'_i) \quad (4.4)$$

最后，将固有特征图与线性变换后得到的幻影特征图进行拼接，从而得到最终的输出结果。

#### 4.2.2 嵌入轻量注意力模块

CBAM (Convolutional Block Attention Module)<sup>[68]</sup>是一种轻量注意力机制模块，其处理特征时要经过两个模块，通道注意力模块和空间注意力模块，通道注意机制可以去除冗余的特征信息，空间注意机制可以去除无关的背景信息，其结构如图 4.4 所示。

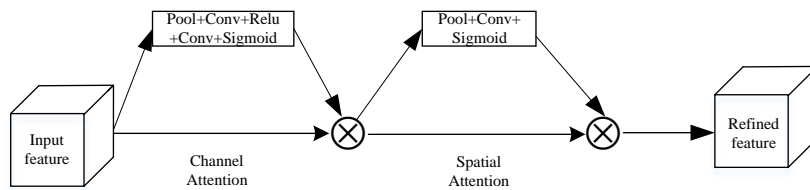


图 4.4 CBAM 模块

由于采集的高原夏菜由交易场所提供，人为已经对其进行加工处理，所以其自身蔬菜质量等级较高，不同等级的高原夏菜缺陷都有减少。因此，为了让原网络更加关注细微的特征，检测和定位局部有用的信息，使相似物种的分析更加精确，本论文在原网络的最后一层前添加了 CBAM 注意力模块，如图 4.5 所示，其将网络中最后一个卷积层的输出首先经过通道注意力模块，接着将该模块的输出作为空间注意模块的输入，最后对两个模块所得到的结果进行加权，从而得到



整个网络模块的输出特征。

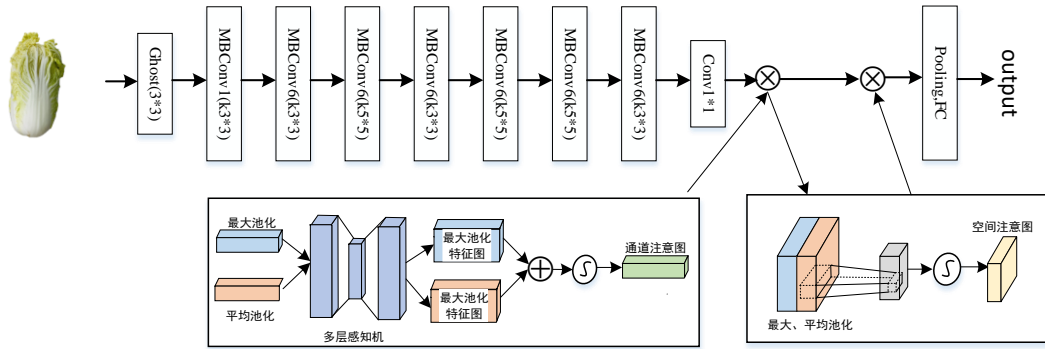


图 4.5 CBAM 模块添加结构图

通道注意模块传递输入的特征图，通过多层感知机所输出的特征执行逐元素求和操作，以生成最终的频道注意特征图。对其和输入特征图进行乘法运算，生成空间注意模块所需的输入特征，公式为：

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (4.5)$$

空间注意模块使用前面提到的通道注意输出的特征图作为这个的输入特征图。将该特征和该模块的输入相乘，以生成最终生成的特征图。公式如 2.7 所示：

$$M_s(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \quad (4.6)$$

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

### 4.3 替换的注意力机制模块

在 EfficientNet 网络中，每一个轻量翻转瓶颈卷积核(MBConv, Mobile Inverted Bottleneck Convolution) 中的注意力机制为 SE 模块，但该模块只考虑信道之间的信息编码，忽略了在蔬菜质量分级中缺陷的位置信息，缺陷的信息会直接影响到高原夏菜质量分级的结构。针对该问题，本论文中将该模块中的 SE 模块替换为 CA 模块，如图 4.6 所示。

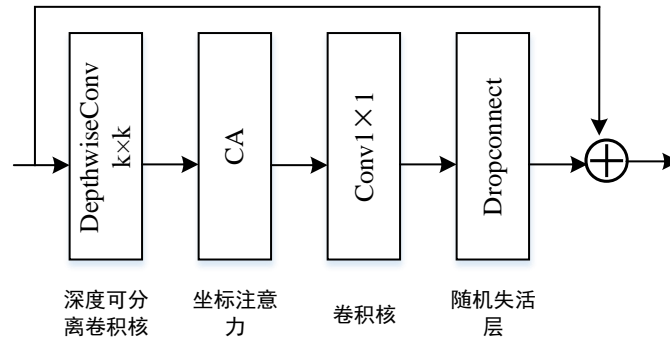


图 4.6 改进的轻量翻转瓶颈卷积核结构图

CA 模块可以保留精准的位置信息，同时也能够捕捉到长距离的依赖关系。可以使原 EfficientNet 网络在关注到高原夏菜图像的区域的同时，也能够确定数据集中缺陷的位置信息，从而提升分级的准确性。

## 4.4 网络超参数的设置

超参数的选择与设置会影响到网络的性能，对网络模型训练时收敛的速度、模型分级的准确率都会产生影响。本小节将与模型自身相关的超参数进行改进微调，并作简要说明。

### 4.4.1 优化算法

原网络的优化算法为随机梯度算法，但是，SGD 算法中各个参数的学习速度都是相同的，这就难以寻找出合适的学习率，同时该算法很容易出现局部优化的现象，导致了模型不够理想。

为了解决 SGD 算法所造成的问题，本论文拟采用 RAdam 优化算法<sup>[69]</sup>，用来提高模型的收敛速度。RAdam 优化算法通过引入修正因子  $r_t$ ，用来修正自适应学习率。具体算法的数学表达式如下，其中一阶矩  $m_t$ 、二阶矩  $v_t$ ，偏置修正  $\hat{m}_t$  的计算与自适应优化算法相同。

$$\begin{cases} \rho_\infty = \frac{2}{1-\beta_2} - 1 \\ \rho_t = \rho_\infty - 2t \cdot \beta_2^t (1-\beta_2^t) \end{cases} \quad (4.7)$$

当方差可解 ( $\rho_t > 4$ ) 时:

$$\begin{cases} l_t = \sqrt{(1 - \beta_2^t) / v_t} \\ r_t = \sqrt{\frac{(\rho_t - 4)(\rho_t - 2)\rho_\infty}{(\rho_\infty - 4)(\rho_\infty - 2)\rho_t}} \\ \theta_t = \theta_{t-1} - \alpha_1 r_t \hat{m}_t l_t \end{cases} \quad (4.8)$$

当方差可解 ( $\rho_t \leq 4$ ) 时:

$$\theta_t = \theta_{t-1} - \alpha_1 \hat{m}_t \quad (4.9)$$

其中,  $\rho$  为自由度,  $\alpha_t, l_t$  分别代表步长与自适应学习率,  $r_t$  为修正因子,  $\theta_t$  为结果参数。

与随机梯度算法、自适应学习算法相比, Radam 可以通过修正因子来修正自适应学习率, 避免出现局部最优的现象, 同时能够降低学习率方差。同时, RAdam 算法的鲁棒性强于上述两算法, 能够加速收敛并提高泛化能力。

#### 4.4.2 损失函数

在网络训练的过程中, 当只考虑正确标签位置的损失时, 不去考虑其他标签的损失, 这将导致该模型过分强调提高正确标签的概率, 从而会造成过度拟合。针对这个问题, 本文提出了一种基于标签平滑的交叉熵损失方法, 既考虑了数据集样本训练时正确标签位置的损失, 同时也考虑到其他轻微误差的标签位置的损失, 使模型在提高正确分类概率的同时也减少了错误的分类情况, 提高了模型的学习能力。Cross-Entropy 损失函数可以描述如下:

$$H(Lable, f(x)) = -\sum_x Lable \times \log(f(x)) \quad (4.10)$$

### 4.5 CA-Ghost-EfficientNet多尺度融合网络模型构建

通过以上所述对原网络的改进, 本文针对高原夏菜质量分级提出了 CA-Ghost-EfficientNet 网络模型, 结构如图 4.7 所示。首先构建轻量化模型, 将原有网络模型中的第一层替换为 Ghost 模块, 并在最后一层前嵌入轻量化 CBAM 模

块，使网络能够在提高分级准确率的同时减少模型参数量；接着引入 CA 模块替换掉原网络中的 SE 模块，使网络既能关注到高原夏菜的相关区域，也不丢失其表面的缺陷信息。

如图 4.7 所示，基于 CA-Ghost-EfficientNet 的高原夏菜质量分级的过程如下：

首先，对采集得到的高原夏菜图像进行预处理，通过数据增强等操作对原始数据集进行扩充并剔除不相干背景特征，接着将处理后的图像输入到 CA-Ghost-EfficientNet 网络中，最后，通过全连接层得到高原夏菜质量分级的结果。

该改进网络模型共有十一个阶段，首先输入图像通过 Ghost 卷积层，接下来第二个到第八个阶段是重复堆叠的 MBConv 结构，其中用 CA 模块替换掉 SE 模块，第九阶段是一个卷积核大小 $1\times 1$ 的卷积层，第十阶段添加了注意力机制模块 CBAM，将前一个卷积层的输出特征作为该嵌入模块的输入特征向量，最后进入全连接层，并通过 Softmax 分类层，根据蔬菜数据集将输出设置为 3 个类别。

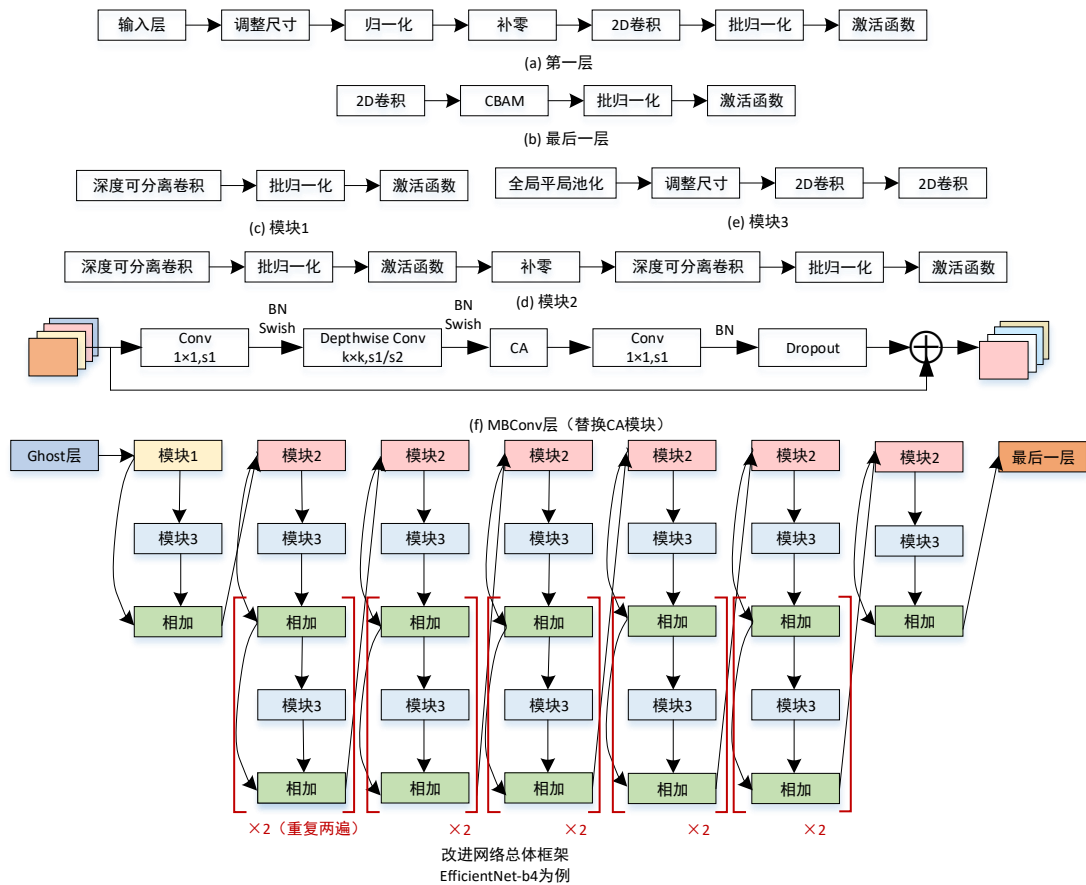


图 4.7 改进网络的总体框架

当训练 EfficientNet 单模型时，需要将图像大小调整到最合适的输入大小，虽然调整大小方法可以将原始图像转化为各种大小的数据集，但大小的变化不可避免地会影响图像的原始特征。同时，本文的高原夏菜数据集中，近一半的图像高度在[360,420]，大部分图像宽度在[493,564]。因此本文在以上的改进基础上，选择 EfficientNet-b4 和 EfficientNet-b6 两种模型进行多尺度融合，使网络实现特征保留最大化，该改进模型中，b4 模型能较好地提取高度维度图像的特征，b6 模型能较好地提取宽度维度图像的特征，其结构图如下。

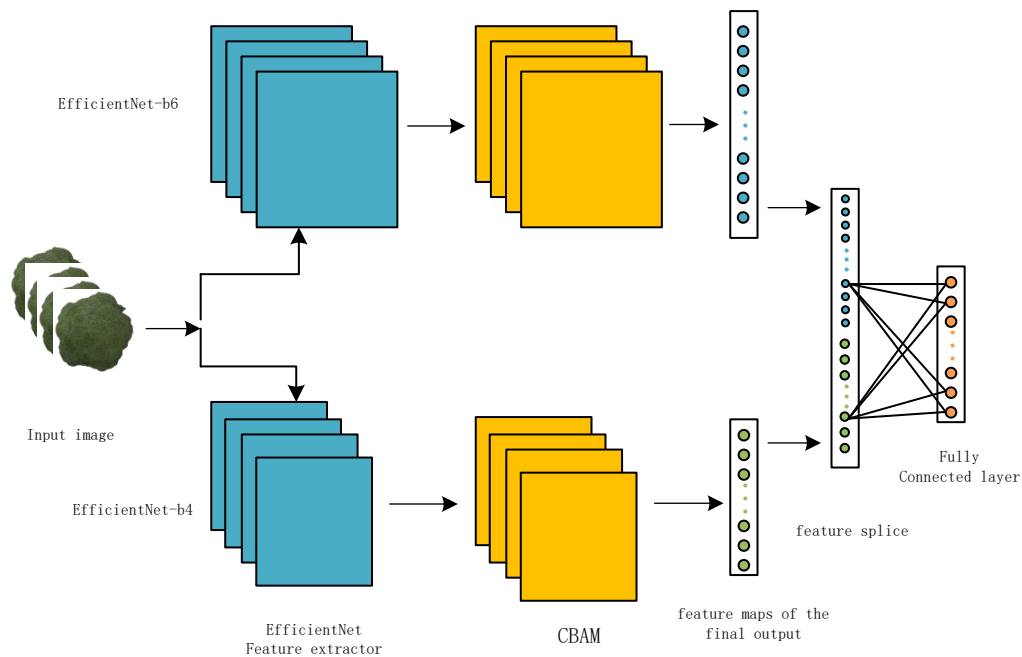


图 4.8 CA-Ghost-EfficientNet 多尺度融合模型

改进的网络与原网络模型相对比，CA-Ghost-EfficientNet 多尺度融合模型主要做了以下改进：

(1) 重新构建轻量翻转瓶颈卷积核模块。本文将 EfficientNet-b4 与 b6 网络模型中的 SE 模块替换为 CA 模块，具体结构如图 4.6 所示使网络能同时保留特征的长期依赖关系与精准的位置信息。

(2) 构建出轻量化模型。由于模型在进行高原夏菜质量分级时，会产生大量的计算量，并且会占据较大的内存，因此，本文将原网络中的第一层替换为 Ghost 层，从而可以使网络利用线性映射生成更多的特征图，减少计算量以及所

占内存，使模型的训练速度加快，同时，在网络最后一层前嵌入轻量化注意力模块，使网络对相似物种的分析更加准确，其总体结构图如 4.4、4.5 所示。

(3) 多尺度融合。本文结合多尺度融合算法，并融合 EfficientNetb0 ~ EfficientNet-b7 模型的网络特点，提出了一种基于 EfficientNet 的多尺度融合算法，其网络结构如图 4.8 所示，当高原夏菜图像（图中所示为西兰花）被输入到 EfficientNet-b4 和 EfficientNet-b6 两个基础网络提取相关特征时，它们将分别通过嵌入 CBAM 的 b4 和 b6 特征提取器。在这个过程中，两个网络结构的不同尺度输出将作为 CBAM 的输入特征，将经过 CBAM 的输出特征进行拼接，从而得出分级结果。

## 4.6 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 EfficientNet 的基本结构进行介绍，同时将该网络通过缩放因子所得到 b1-b7 的参数进行概括。之后对改进模型的具体改进方法进行阐述，先构建出轻量化模型，将原网络中的第一层的普通卷积层替换为 Ghost 层，同时，在网络最后一层前添加轻量化注意力模块 CBAM，从而使网络减少在特征提取过程中的运算量，并提高分级准确率；接着将 MBConv 层中的注意力机制模块 SE 用 CA 模块进行替换，使其在划分蔬菜等级的同时也能够注意到蔬菜上的缺陷，有助于蔬菜等级的分类；最后将随机梯度算法替换为 Radam 优化算法，提高网络模型的准确率。

## 5 高原夏菜质量分级的实验结果与分析

### 5.1 算法流程

本文提出的改进 EfficientNet 网络的高原夏菜质量分级具体流程如下图所示，其中， $\mu$ 为 epoch 的大小。具体算法的具体步骤如下：

(1) 将采集得到的四种高原夏菜经过预处理后得到的图像输入到改进后的 CA-Ghost-EfficientNet 网络模型中，对其进行特征提取。

(2) 通过网络设置的模型参数训练后，Softmax 分类器对质量分级结果进行预测，之后对输出的分级结果用 Cross-Entropy 损失函数计算准确率和损失值。

(3) 若训练轮次小于或等于所设置的大小时，通过优化算法更新网络权重，之后继续学习。

(4) 重复上述步骤，更新网络权重，直到训练轮次大于设置的大小 $\mu$ 时，得到分级结果。

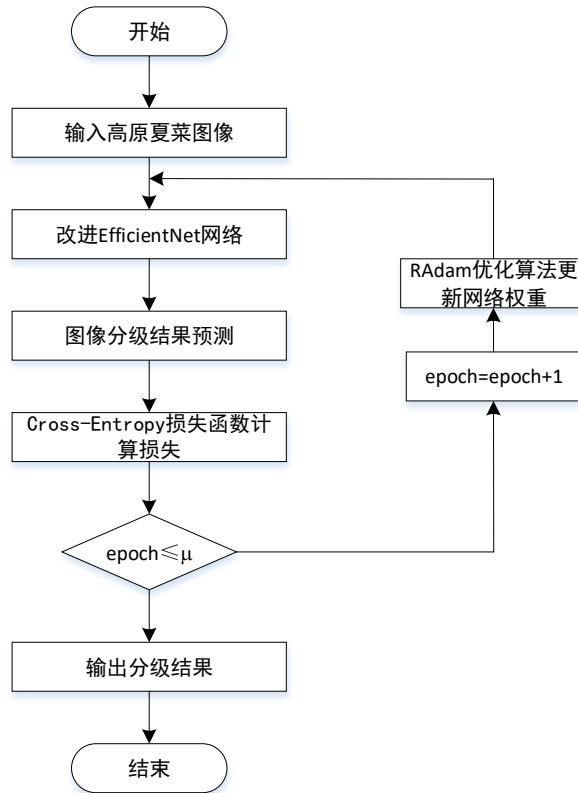


图 5.1 算法流程图

## 5.2 高原夏菜数据集划分

在进行模型训练时，数据集通常的划分方法有以下三种：留出法、交叉验证法与自助法。由于本实验所采集到的高原夏菜图像数量不充足，因此本文采用交叉验证的方法来划分数据集，从而提高模型的泛化能力。

实验前通过人工拍摄采集得到原始图像 2400 张，通过第三章的步骤，我们得到一个有效的高原夏菜数据集，共有 3600 张高原夏菜图片的数据集，其中每类蔬菜共 900 张，300 张为特级蔬菜，300 张为一级蔬菜，300 张为二级蔬菜。每类蔬菜随机选取 240 张作为训练，60 张作为测试。具体划分如下表所示。

表 5.1 数据集划分

	训练集（每类别蔬菜数量）	测试集（每类别蔬菜数量）
结球甘蓝	240	60
娃娃菜	240	60
菜花	240	60
西兰花	240	60

## 5.3 实验环境与评价指标

### 5.3.1 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:15 vCPU AMD EPYC 7543 32-Core Processor；具有 Cuda 11.3 的加速框架，80G 内存。深度学习配置的环境为 PyTorch 1.10.0、Python 3.8，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 5.2 所示。



表 5.2 实验参数设置

模型参数名称	详情
Batch size	16
Epoch	80
Initial learning rate	7.5e-5
Learning rate schedule	Cosine annealing
Class num	3
Optimizer	RAdam
Dropout	0.5

### 5.3.2 评价指标

深度学习模型的性能需要具体量化,通过不同的指标进行对比分析,从而验证模型的性能。本文对改进的 EfficientNet 网络的高原夏菜质量分级实验中通过准确率、F1 值进行评估和检测模型的性能。同时,本文用模型参数量的大小来验证算法的复杂度,用以对比原模型与改进模型,验证改进模型的有限性。

准确率代表着模型预测分级正确图像个数占总图像个数的比例;精确率代表预测正确的图像数占全部预测为正的的比例,召回率反映的是预测正确的图像数占所有实际为正的的比例,由于精确率与召回率是呈负相关的,通常情况下会出现矛盾,为了解决这一矛盾,则引入 F1 值作为综合评价指标,该值是通过两者进行加权调和平均得到的,F1 值越接近 1,网络模型的性能越好。

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

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

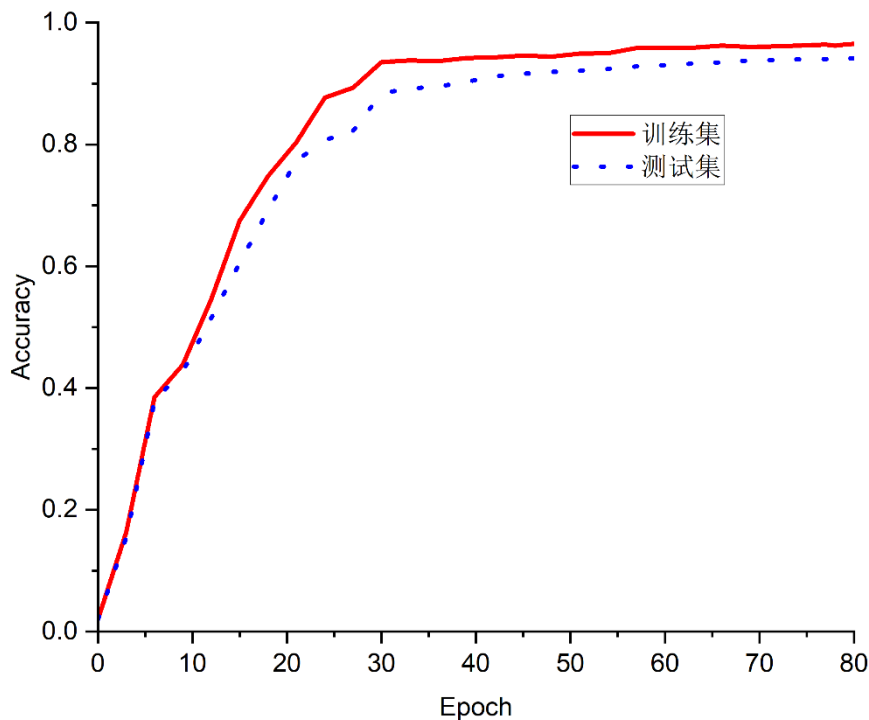
其中,  $TP$  为真正例样本数,即数据集中就是正样本,预测结果也为正样本; $TN$  为真负例样本数,即数据集中就是负样本,预测结果也为负样本; $N$  为数据集样本总数。

## 5.4 结果与分析

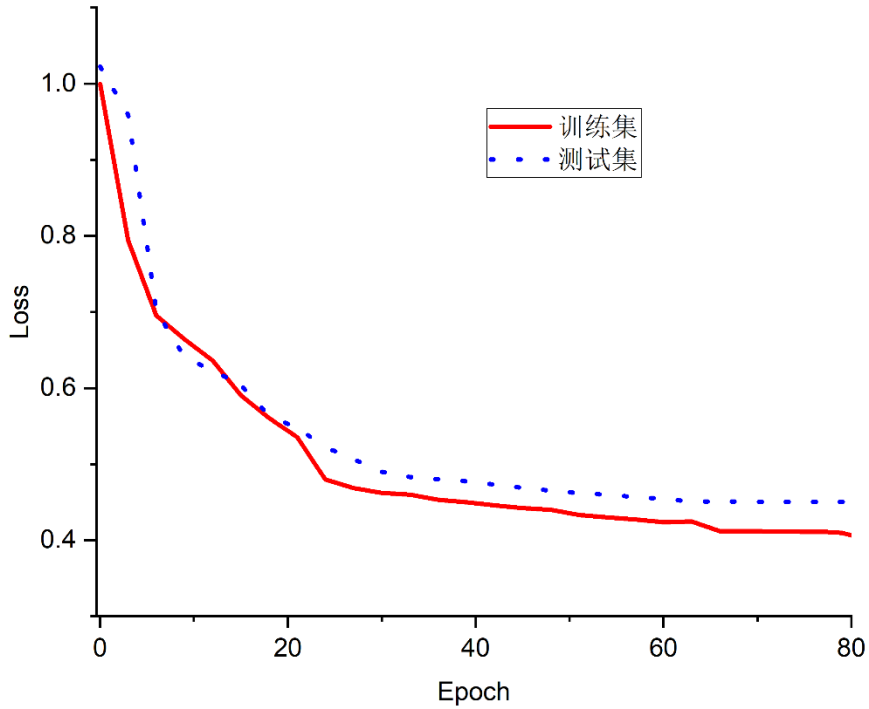
为验证改进模型的性能，本节对模型训练结果进行分析，分为三部分：第一部分是将四种高原夏菜输入进原网络模型与改进模型中进行训练，对比模型的准确率、F1 值与参数量并进行分析；第二部分将改进模型与经典的卷积神经网络模型进行对比分析；第三部分进行消融实验，对每一步改进的模型在数据集上的训练结果进行横向对比，从而来验证各个改进模块的有效性。

### 5.4.1 与原网络模型性能对比

对于 4.5 节提出的高原夏菜质量分级模型 CA-Ghost-EfficientNet 网络，采用 5.3 节中的试验设置进行模型训练。将改进的 EfficientNet 模型在高原夏菜的训练集和测试集上进行训练，并可视化显示，以结球甘蓝为例，其准确率及损失率曲线如图 5.2 所示。



(a) 网络训练准确率对比图



(b) 网络训练损失率对比图

图 5.2 CA-Ghost-EfficientNet 模型准确率与损失率曲线图

可以从上图中看到，当模型迭代次数在 30 次左右时，改进的 EfficientNet 模型的准确率与损失率都开始平缓。当迭代次数为 50 次左右时，准确率趋于稳定，损失率也基本收敛，证明此时模型已经达到饱和状态，识别的准确率也达到了最好，达到了 94.12%。

将改进模型在四种高原夏菜数据集上进行训练，并与在单模型 EfficientNet-b5 上训练的结果进行对比，用准确率与模型参数量进行评价，结果如下表 5.3、5.4 所示。

表 5.3 四种高原夏菜在改进 EfficientNet 模型的性能对比

蔬菜种类	准确率/%	F1/%	参数量/M
结球甘蓝	94.12	91.10	4.723
娃娃菜	92.06	87.97	4.736
菜花	86.54	81.20	5.367
西兰花	88.12	83.46	5.342

表 5.4 四种高原夏菜改进 EfficientNet 模型前后的准确率与 F1 值对比

蔬菜种类	EfficientNet-b5/%		改进 EfficientNet/%	
	准确率	F1	准确率	F1
结球甘蓝	90.93	88.75	94.12	91.10
娃娃菜	89.63	84.96	92.06	87.97
菜花	81.68	74.81	86.54	81.20
西兰花	82.43	75.51	88.12	83.46

如表 5.4 所示，是改进的 CA-Ghost-EfficientNet 和 EfficientNet-b0 模型在四种高原夏菜数据集上准确率与 F1 值的变化表格。可以看到改进模型在四种高原夏菜上都有较好的表现，结合两表，可以看到结球甘蓝、娃娃菜的准确率与 F1 值明显高于其他两种蔬菜，这与第三章所讲的蔬菜与预处理有关，结球甘蓝与娃娃菜的外观、大小、缺陷程度在拍照采集时就有较大的区分，可以很好地区分等级，但菜花与西兰花两种蔬菜呈现花状，且花球与机械损伤程度相差不大，因此分级准确率没有以上两种蔬菜高。但在以上四种蔬菜中，改进模型的准确度明显高于原模型，证明了改进模型的有效性。

#### 5.4.2 与经典深度卷积神经网络模型性能对比

从表 5.3 中可以看出，该改进模型在结球甘蓝表现相较于其他三种优异，因此下文选择在结球甘蓝数据集对模型进行训练。在结球甘蓝测试数据集下分别将改进的多尺度融合的 CA-Ghost-EfficientNet 模型与 VGGNet16、ResNet50、DenseNet169 模型进行横向对比，并进行可视化显示，其中超参数设置为：一阶和二阶矩衰减系数  $\beta_1 = 0.9, \beta_2 = 0.999$ ，运行结果如下图所示。

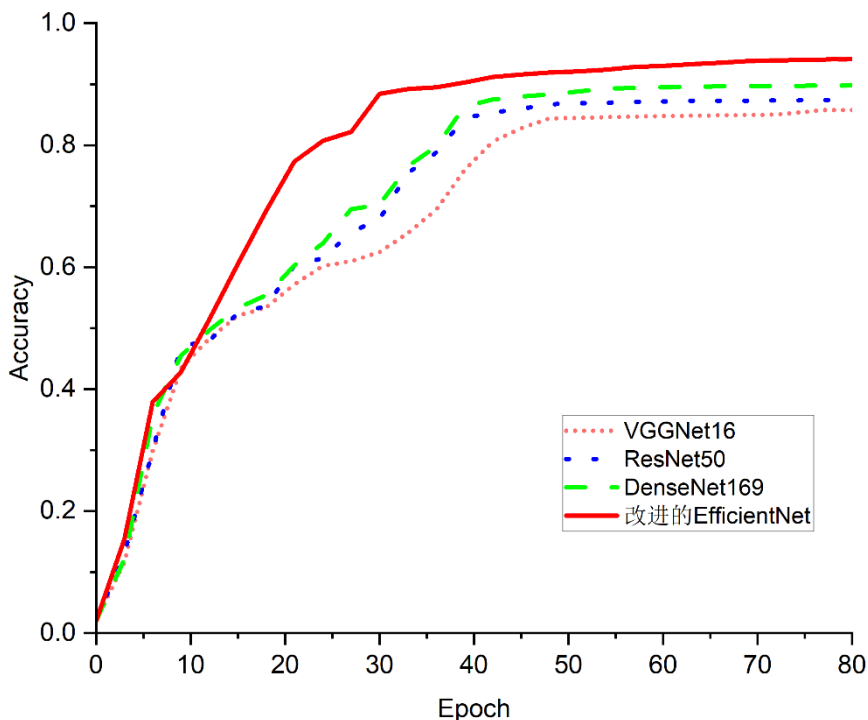


图 5.3 不同模型准确率对比图

通过图 5.3 可以看到，四种模型的准确率都随着迭代次数的增加而增加，直至趋于平缓。在迭代次数在 30 次左右时，可以看出改进的 EfficientNet 模型的准确率增长最快，收敛速度明显快于其他两种模型；当迭代次数为 50 次左右时，四种模型的准确率都趋于平缓，各模型在结球甘蓝测试集上的识别结果如表 5.5 所示。

表 5.5 各模型的实验结果对比

模型	准确率/%	F <sub>1</sub> /%	参数量/M
VGGNet16	85.78	80.12	46.7
ResNet50	87.45	82.37	42.3
DenseNet169	89.83	85.70	20.7
改进 EfficientNet	<b>94.12</b>	<b>91.10</b>	<b>4.7</b>

从表 5.5 中可以看出，本文所提到的方法在结球甘蓝蔬菜图像测试集上具有最佳的分级准确率，为 94.12%，所改进的 EfficientNet 模型在测试集上比

VGGNet16、ResNet50、DenseNet169 的准确率分别提高了 8.34%、6.67%、4.29%。相较于以上三种通过增加网络深度提高分类准确度的网络而言，本文所改进的模型则是将轻量化模块、CA 模块、CBAM 模块与 EfficientNet 的优点进行融合，在减少计算量、避免过拟合现象出现的同时，也尽可能保留了输入图像的特征。与以上三种经典的卷积神经网络相比，本文所提方法在增加准确率的同时，也减少了模型参数量。

### 5.4.3 消融实验对比

为进一步验证改进模型的有效性，本文进行消融实验，通过以下五种方式在结球甘蓝测试集上对比：

方式 1：使用 EfficientNet-b5 单模型在结球甘蓝数据集上进行训练；

方式 2：改进原网络的优化算法，使用 RAdam 优化器；

方式 3：在方式 2 的基础上，用 Ghost 轻量化模块替换原网络中的第一个卷积层，同时在网络最后一层添加轻量化模块 CBAM，从而构建出轻量化模型；

方式 4：在方式 3 的基础上，将 MBConv 结构中 SE 模块替换为 CA 模块；

方式 5：在方式 4 的基础上采用 EfficientNet-b4 与 b6 网络进行融合，形成本文所构建的多尺度融合 CA-Ghost-EfficientNet 模型。消融实验对比结果如下表所示。

表 5.6 消融实验结果

改进策略	准确率/%	参数量/M
EfficientNet-b5	90.93	7.642
EfficientNet-b5+RAdam	91.63	7.642
Ghost-EfficientNet-b5+RAdam +CBAM	92.78	4.737
CA-Ghost-EfficientNet+RAdam +CBAM	93.26	4.723
多尺度融合的 CA-Ghost-EfficientNet	94.12	4.723

以上五种消融实验的方式的结果如上表所示。通过对比方式 1 与方式 2 的结果可知，RAdam 优化器加速模型的收敛的同时模型的参数量也未有增加，准确

率提升到 91.63%，这证明本论文中提出的方式 2 的可行性。结合 1、2、3 的结果可以看出，构建的轻量化模型不仅参数量减少了，且模型的准确率也较之前有所提升，这证明替换的 Ghost 模块与嵌入的 CBAM 轻量化模块的有效性。进一步比较方式 3 与 4 的结果，可知替换的 CA 模块有更强的学习能力。另外，比较改进方式 4 和 5 的结果，可知将两种模型进行多尺度融合是有效的。因此，通过以上的对比，可知对原网络模型的改进有效和可行的。

从表 5.6 中可以看出，本文所提出的改进模型在准确率上相较于原模型提高了 3.19%，参数量约为改进前的 62%。将随机梯度算法替换为 RAdam 算法，模型的参数量不变，但准确度提高了 0.7%；将原网络的第一层替换为 Ghost 层，同时将 MBConv 模块中的注意力机制 SE 替换为 CA 模块，并嵌入 CBAM 模块，在准确率提高 1.84%的同时模型参数量降低了 2.919M；通过融合多尺度融合思想，可以有效地提高准确率并降低模型参数量。

## 5.5 本章小结

本章主要研究内容为改进模型 CA-Ghost-EfficientNet 的在高原夏菜图像数据集上训练。首先将改进模型在四种高原夏菜蔬菜数据集上进行训练，并与原模型进行对比，测试改进模型的分级效果。接着在结球甘蓝数据集上将该模型与 VGGNet16、ResNet50、DenseNet169 网络模型对比分析，最后进行消融实验，验证改进模型的可行性与有效性。

## 6 总结与展望

### 6.1 工作总结

高原夏菜“优质优价”的实现对于甘肃省农产业的发展至关重要，其不仅能够大大提高甘肃省蔬菜市场的竞争力，同时也能提高百姓的收入。针对当前甘肃省高原夏菜数据集缺少以及未进行质量等级分类的问题，本论文构建出高原夏菜数据集，并提出了基于深度学习的高原夏菜质量分级模型，以解决以上问题。具体工作总结如下：

(1) 构建高原夏菜图像数据集，其中包含结球甘蓝、娃娃菜、菜花、西兰花四种蔬菜。当前在蔬菜分类领域中公开的、高质量的数据集很少，高原夏菜数据集更是少之又少，因此本文建立了一套实验室采集装置，构建出高原夏菜数据集。该数据集不仅可供本文的研究，还可进一步发布公开，进而促使甘肃省乃至全国的蔬菜分类领域的发展。

(2) 针对原始数据集不足的情况，为保证采集的高原夏菜数据集在模型训练时能有良好的分级效果，本文首先通过  $U^2Net$  网络对采集的数据集背景进行移除，从而减少背景特征对训练模型的干扰。之后通过图像几何变换，包括图像翻转、旋转，对数据集进行扩充，扩充至 3600 张。

(3) 针对当前对于高原夏菜质量分级大多依旧采用传统人工的方式，本文提出了针对高原夏菜质量分级的多尺度融合的 CA-Ghost-EfficientNet 模型。首先将原网络中 SGD 算法替换为 RAdam 算法，避免出现局部最优的现象，也能加快模型收敛，模型最确率为 91.63%，相较于原模型提高了 0.7%。接着将网络中的第一层替换为 Ghost 层，同时在网络最后一层之前添加 CBAM 模块，在准确率提高 1.15% 的同时模型参数量降低了 2.905M。之后将 EfficientNet 网络模型中 MBConv 模块中的注意力机制 SE 模块替换为 CA 模块，并结合多尺度融合算法，使网络能同时保留特征的长期依赖关系与精准的位置信息，该改进模型的准确率达到 94.12%，为甘肃省高原夏菜质量分级的研究发展提供了一定的经验。



## 6.2 工作展望

本文针对高原夏菜质量分级方法进行了研究,所提出的方法在现有的结球甘蓝、娃娃菜数据集上有较好的效果,能够精确识别不同等级蔬菜产品,对甘肃省蔬菜产业有着很大的受用价值,但针对菜花、西兰花两种蔬菜以及数据集的样本数量,目前模型仍有一些不足之处,有待继续深入研究和完善。

(1) 高原夏菜数据集不够完善。由于高原夏菜种类繁多,本文只采用了四种进行研究,泛用性可能较低。后续希望对高原夏菜数据集进一步扩充,包括蔬菜的数量、品种以及高质量缺陷的添加,使数据集更加强壮,提高模型的泛用性。

(2) 由于高原夏菜各种类之间区别明显,菜花与西兰花两种蔬菜的外观呈现花状,且花球与机械损伤度相差不大,容易造成模型提取特征时不能很好区分,因此对于难以区分的蔬菜,如何更好地提取表面的特征也是高原夏菜质量分级方法的一个难点,还需要进一步地优化改进。

(3) 分级算法需要不断地改进。对于蔬菜分级的研究当前还是较为缺少,本文所提出的模型方法虽取得了良好的效果,但后续研究还可以采用其他经典的卷积神经网络,对网络模型进行改进,进而提高分级的准确度,从而使蔬菜图像的分级更加准确。

## 参考文献

- [1] 杨朔. 基于深度学习的蔬菜图像智能识别及应用研究[D]. 南昌大学.
- [2] 杨梦园. 乡村振兴背景下高原夏菜产业的发展研究——以兰州市榆中县为例[J]. 今日财富, 2020(10):43-44.
- [3] 黄爱珍. 高原夏菜发展问题及对策分析[J]. 农业开发与装备, 2019, No.213(09):7.
- [4] Shiv Ram Dubey and A.S. Jalal. Robust Approach for Fruit and Vegetable Classification[J]. Procedia Engineering, 2012, 38(C) : 3449-3453.
- [5] F. Kurtulmuş and H. Ünal. Discriminating rapeseed varieties using computer vision and machine learning[J]. Expert Systems With Applications, 2015, 42(4) : 1880-1891.
- [6] 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.
- [7] 陈艳军, 张俊雄, 李伟, 等. 基于机器视觉的苹果最大横切面直径分级方法[J]. 农业工程学报, 2012, 28(2):5. 2012, 28(2):284-288.
- [8] 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]. 农业工程报, 2014, 30(16):305-311.
- [9] José de Jesús Rubio. A method with neural networks for the classification of fruits and vegetables[J]. Soft Computing, 2017, 21(23) : 7207-7220.
- [10] D. Latha and M. Mohamed Sathik and Y. Jacob Vetha Raj. Multilevel approach of cbir techniques for vegetable classification using hybrid image features[J]. ICTACT Journal on Image and Video Processing, 2016, 6(3) : 1174-1179.
- [11] Leiqing Pan et al. Hyperspectral imaging with different illumination patterns for the hollowness classification of white radish[J]. Postharvest Biology and Technology, 2017, 126 : 40-49.
- [12] George M. Multiple Fruit and Vegetable Sorting System Using Machine Vision. [J]. International Journal of Advancements in Technology , 2015 , 6(1): 1-4.
- [13] Patel K K , Kar A , Jha S N , et al. Machine vision system: a tool for quality inspection of food and agricultural products[J]. Journal of Food Science &

- Technology, 2012, 49(2):123-141.
- [14] Xiao Q Niu W, Zhang H. Predicting fruit maturity stagedynamically based on fuzzy recognition and colorfeature[C]//2015 6th International Conference onSoftware Engineering and Service Science (ICSESS).IEEE, 2015: 944-948
- [15] Kondo N.Automation on fruit and vegetable grading system and food traceability[J]. Trends in Food Science & Technology, 2010, 21(3):145-152.
- [16] Sun D W . Inspecting pizza topping percentage and distribution by a computer vision method[J]. Journal of Food Engineering, 2000, 44(4):245-249.
- [17] 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.
- [18] 刘禾, 汪懋华.水果果形判别人工神经网络专家系统的研究. 农业工程学报, 1996,12 (1): 171-176.
- [19] Rocha A , Hauagge D C ,Wainer J ,et al. Automatic fruit and vegetable classification from images[J]. Computers & Electronics in Agriculture, 2010, 70(1):96-104.
- [20] Li J , Liao G , Jin J , et al. Method of potato external defects detection based on fast gray intercept threshold segmentation algorithm and ten-color model[J]. Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering, 2010, 26(10):236-242.
- [21] 张亚静,Sakae Shibusaw,等.基于机器视觉的番茄内部品质预测[J].农业工程学报, 2010 , 26(S2): 366-370.
- [22] 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.
- [23] Kumar S , Kumar S . Texture Features and Decision Trees based Vegetables Classification[J]. ijca proceedings on national conferecne on advanced computing & communications, 2012: 21-26.
- [24] Dubey S R , Jalal A S . Fruit and vegetable recognition by fusing colour and texture features of the image using machine learning[J]. Humanika, , 2015, 2(2) : 160-181.

- [25] Madgi, M., Danti, A., & Anami, B. Combined RGB colour and local binary pattern statistics features-based classification and identification of vegetable images [J]. *Int. J. of Applied Pattern Recognition*, 2015, 2(4) : 340-352.
- [26] Shuihua Wang et al. Fruit Classification by Wavelet-Entropy and Feedforward Neural Network Trained by Fitness-Scaled Chaotic ABC and Biogeography-Based Optimization[J]. *Entropy*, 2015, 17(8) : 5711-5728.
- [27] Juntao Xiong et al. Classification of Potato External Quality based on SVM and PCA[J]. *International Journal of Performability Engineering*, 2017, 17(4): 469-478.
- [28] Wan, P., Toudeshki, A., Tan, H., & Ehsani, R. (2018). A methodology for fresh tomato maturity detection using computer vision. *Computers and Electronics in Agriculture*, 146, 43–50.
- [29] Huang X , Yu S , Xu H , et al. Rapid and nondestructive detection of freshness quality of postharvest spinaches based on machine vision and electronic nose[J]. *Journal of Food Safety*, 2019, 39(4):1-8.
- [30] 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(2):146-150.
- [31] Lee S H , Chan C S , Mayo S J , et al. How deep learning extracts and learns leaf features for plant classification[J]. *Pattern Recognition*, 2017, 71:1-13.
- [32] Gao X W , Hui R , Tian Z . Classification of CT brain images based on deep learning networks[J]. *Computer Methods & Programs in Biomedicine*, 2017, 138:49-56.
- [33] P. Wang, W. Li, S. Liu, et al. Large-scale Isolated Gesture Recognition using Convolutional Neural Networks[C]//In 2016 23rd International Conference on Pattern Recognition (ICPR), 2016: 7-12.
- [34] 曾平平,李林升.基于卷积神经网络的水果图像分类识别研究[J].*机械设计与研究*,2019,35(1):23-26, 34.
- [35] Raikar M M , Meena S M , Kuchanur C , et al. Classification and Grading of Okra-ladies finger using Deep Learning[J]. *Procedia Computer Science*, 2020, 171:2380-2389.
- [36] Xue, G., Liu, S., & Ma, Y. A hybrid deep learning-based fruit classification using

- attention model and convolution autoencoder [J]. *Complex & Intelligent Systems*, 2020, : 1-11.
- [37] Zhenbo Li 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.
- [38] Fu Yuesheng et al. Circular Fruit and Vegetable Classification Based on Optimized GoogLeNet[J]. *IEEE ACCESS*, 2021, 9 : 113599-113611.
- [39] 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.
- [40] Gill, H.S., Khalaf, O.I., Alotaibi, Y., Alghamdi, S. and Alassery, F. Multi-Model CNN-RNN-LSTM Based Fruit Recognition and Classification[J]. *Intelligent Automation & Soft Computing*, 2022, 33(1) : 637-650.
- [41] Jayakumari, R., Nidamanuri, R. R., & Ramiya, A. M. 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.
- [42] Li X , Rai L . Apple Leaf Disease Identification and Classification using ResNet Models[C]// 2020 IEEE 3rd International Conference on Electronic Information and Communication Technology (ICEICT) . IEEE,2020 . 738–742.
- [43] Son C H. Leaf Spot Attention Network for Apple Leaf Disease Identification[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. IEEE, 2020: 229–237.
- [44] Wang J, Liu Q, Xie H, Yang Z, Zhou H. Boosted EfficientNet: detection of lymph node metastases in breast cancer using convolutional neural networks.[J] *Cancers*. 2021;13(4): 661-661.
- [45] 王振亚,赵继红,王艳鹏,葛广英,孙群.基于改进 EfficientNet 网络的药片检测分类的研究[J].*现代计算机*,2021,27(28):27-32.
- [46] Guan Shengqi et al. Strip Steel Defect Classification Using the Improved GAN and EfficientNet[J]. *Applied Artificial Intelligence*, 2021, 35(15) : 1887-1904.
- [47] 光金正,梁鉴如,刘义生.基于改进 EfficientNet 的植物图像分类算法[J].*传感器*

- 与微系统,2022,41(8):136-139.
- [48] 王芳,刘小虎,罗艺闯.基于显著性检测和 EfficientNet 的垃圾分类算法研究[J]. 计算机与数字工程,2022,50(6): 1163-1166.
- [49] S M Jaisakthi et al. Classification of skin cancer from dermoscopic images using deep neural network architectures.[J]. Multimedia tools and applications, 2022, : 11-16.
- [50] 姜天宇,赵晓林,赵搏欣,李伟龙,吴梦瑶.基于 EfficientNet 的木薯叶病变自动分类模型[J]. 计算机应用,2022,42(S1): 64-70.
- [51] 帖军,隆娟娟,郑禄,牛悦,宋衍霖.基于 SK-EfficientNet 的番茄叶片病害识别模型[J]. 广西师范大学学报(自然科学版),2022,40(4): 104-114.
- [52] 苏赋,吕沁,罗仁泽.基于深度学习的图像分类研究综述[J]. 电信科学,2019,35(11): 58-74.
- [53] 郑欣悦. 基于深度学习的少样本图像分类方法[D]. 中国科学院大学(中国科学院国家空间科学中心),2019.
- [54] 赵艳芹,童朝娣,张恒.基于 LeNet-5 卷积神经网络的车牌字符识别[J]. 黑龙江科技大学学报,2019,29(3): 382-386.
- [55] Ioffe S , Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[J]. JMLR.org, 2015: 448–456.
- [56] 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.
- [57] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016: 770-778.
- [58] Huang G, Liu Z, Laurens V, et al. Densely Connected Convolutional Networks[J]. IEEE Computer Society, 2016: 4700–4708.
- [59] Jie, Shen, Samuel, et al. Squeeze-and-Excitation Networks.[J]. IEEE transactions on pattern analysis and machine intelligence, 2019.42: 7132–7141.
- [60] Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design[C]// Proceedings-34th IEEE Conference on Computer Vision and Pattern

- Recognition (CVPR 2021). IEEE Computer Society, 2021: 13713-13722.
- [61] Robbins Herbert and Monro Sutton. A Stochastic Approximation Method[J]. The Annals of Mathematical Statistics, 1951, 22(3) : 400-407
- [62] Kingma D, Ba J. Adam: A method for stochastic optimization[C]//3rd International Conference on Learning Representations (ICLR 2015). San Diego, USA: International Conference on Learning Representations, 2015: 1-15.
- [63] 黄玉萍,梁炜萱,肖祖环.基于 TensorFlow 和 PyTorch 的深度学习框架对比分析 [J].现代信息技术,2020,4(4): 80-82,87.
- [64] 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.
- [65] 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.
- [66] Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks[C]//International Conference on Machine Learning. PMLR, 2019: 6105-6114.
- [67] Han K. et al. GhostNet: More features from cheap operations[J]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020, : 1577-1586.
- [68] 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
- [69] Liu L, Jiang H, He P, et al. On the Variance of the Adaptive Learning Rate and Beyond[J]. 2019.

## 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] 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.
- [2] 宋文琴, 尚庆生, 巩晴. 旅游评论短文本的改进 ERNIE-RCNN 模型分类[J]. 宜宾学院学报, 2021, 21(12):53-56.



## 致谢

花开花落万物道，聚散离别终有时。从 2020 年的金秋到 2023 年的盛夏，从懵懵懂懂到成熟稳重，三年的时光仿佛弹指一挥间，在这段旅程中所发生的所有事情，也许不会终生铭记，但定会一生感激。

饮水思其源，学成念吾师。首先要向我的导师尚庆生老师致以最衷心的感谢，无论是传道授业还是生活琐事，都感谢您的体谅与包容；无论是论文选题、修改还是最后定稿，都感谢您的悉心指导，无数次交流探讨，无数次文件的传阅，积淀成感激又感动的记忆，使我在学习的过程中不断进行自我反思，不断前进。感谢管理科学与工程专业的所有老师们，是你们的无私授课与倾囊相授，让我在这三年的学习中，受益匪浅，衷心祝愿各位老师工作顺利，生活美满。

万爱千恩百苦，疼我孰知父母？在 20 多年的求学时光里，有着数不清的迷茫，感谢我的父母二十多年来无怨无悔地付出，无微不至的照顾；感谢你们永远做我最坚实的靠面，支持我的想法；感谢你们让我无忧无虑长大又备受宠爱，我最美好的品格都归功于你们，愿你们碎碎念念可以岁岁年年。

愿岁并谢，与友长兮。感谢我的室友和研究室的每一位小伙伴，生活中的交流、困难时的鼓励、高光时的祝福、疫情时期的陪伴，点点滴滴的事情构成了一段美好的回忆，未来依旧会遇到许多人，但这段回忆我会珍藏于心。我们相遇在彼此的黄金时代，用最饱满的热情投身于自己选择的道路中，无论前方是鲜花遍地还是荆棘丛生，都希望我们不骄不躁、不气不馁，愿我们前路漫漫亦灿灿。

有幸遇见，恰好合拍。感谢在我想尝试新领域时全力支持的你，感谢在我紧张不安时告诉我手心画人的你，感谢在我不懂交流时耐心引导的你，感谢在悄然流逝的时光里都有你的存在，希望我们在相互陪伴中一起努力，一起成长。愿可共白首，余生常相伴。

少年不惧岁月长，彼方尚有荣光在。感谢在前行道路中一直坚持的自己，虽有气馁，但从未放弃。坚信所有经历都是成长，希望在未来前行的道路中，忠于自己，不忘初心，永葆热爱。

至此落笔，我的学生时代就要落下帷幕。韶光易逝，终有别离，愿我们都能成为最好的自己，顶峰相见！