

# 基于CBAM-ResNet50模型的水果图像识别

常添春\*, 王联国#

甘肃农业大学, 信息科学技术学院, 甘肃 兰州

收稿日期: 2023年12月25日; 录用日期: 2024年2月19日; 发布日期: 2024年2月29日

## 摘要

针对传统水果分类识别精度低, 人工成本高等问题, 提出一种基于CBAM-ResNet50模型的水果图像识别算法。首先, 采用迁移学习技术, 将ImageNet数据集上训练好的权重参数迁移到ResNet50网络水果图像分类模型中, 保留卷积层和池化层, 去掉分类器, 作为主干网络模型; 其次, 在主干网络模型后添加混合注意力机制模块, 根据不同的输入特征分配不同权重, 提取有效特征, 忽略无关信息。然后, 用全局平均池化(GlobalAvgPool2D)替换平均池化, 将高维数据转化为低维数据, 提高计算效率并简化模型训练过程。最后, 添加dropout正则化, 随机失活权重参数比例, 以确保网络对噪声和异常值的鲁棒性, 构建Batch Normalization层对输入数据进行归一化, 帮助网络更好地学习数据信息的特征分布, 进而提高网络模型性能。把收集到的水果图像按照随机取样的方法划分为80%训练集和20%测试集两部分, 采用旋转、平移和裁剪等技术扩充水果图像数据集的多样性和变化性, 本文提出CBAM-ResNet50网络模型与MobileNet-v3、VGG16、AlexNet、Xception、ResNet50网络模型的识别效果进行对比, 试验结果表明, 该模型能够有效识别出几种常见的水果图像, 相较于初始网络, 识别准确率增加了6个百分点, 测试准确率达99%。为了进一步验证模型性能, 分析了基于迁移学习下的数据集扩充与未扩充, 添加混合注意力机制对网络模型的影响, 由此得出, 该研究方法在水果分类识别中具有很好的实践意义。

## 关键词

ResNet50网络, 混合注意力机制, 迁移学习, 数据增广, 水果图像识别

# Fruit Image Recognition Based on CBAM-ResNet50 Model

Tianchun Chang\*, Lianguo Wang#

School of Information Science and Technology, Gansu Agricultural University, Lanzhou Gansu

Received: Dec. 25<sup>th</sup>, 2023; accepted: Feb. 19<sup>th</sup>, 2024; published: Feb. 29<sup>th</sup>, 2024

\*第一作者。

#通讯作者。

## Abstract

In response to the traditional fruit classification recognition accuracy and high labor cost, a fruit image recognition algorithm based on the CBAM-ResNet50 model is proposed. First of all, use migration learning technology to migrate the trained weight parameters trained on the ImageNet data set to the ResNet50 network fruit image classification model, retain the convolutional layer and pool layer, and remove the classifier as the main network model; second, in the main network of the main network, After the model, add a hybrid attention mechanism module, allocate different weights according to different input features, extract effective features, and ignore irrelevant information. Then, the average pooling is replaced with global average pooling (GLOBALAVGP-OOL2D), high -dimensional data is converted into low-dimensional data, increasing the calculation efficiency and simplifying the model training process. Finally, add Dropout regularization, randomly loss weight parameters ratio to ensure that the network has a robustness of noise and abnormal values, and builds the Batch Normalization layer to naturalize the input data to help the network better learn the characteristic distribution of data information. , By improve the performance of network models. The collected fruit images are divided into two parts: 80%training set and 20%test set in accordance with the method of random sampling. The diversity and variation of fruit image data sets are expanded by technology such as rotation, translation and cutting. The network model is compared with the recognition effects of MobileNet-V3, VGG16, Alexnet, Xception, and ResNet50 network model. The test results show that the model can effectively identify several common fruit images. Compared with the initial network, the accuracy rate of recognition has increased. 6 percentage points, the test accuracy rate is as high as 99%. In order to further verify the performance of the model, the expansion and unprepared data of data sets based on migration learning were analyzed, and the impact of a hybrid attention mechanism on the network model. From this, this research method has good practical significance in fruit classification recognition.

## Keywords

ResNet50 Network, Hybrid Attention Mechanism, Transfer Learning, Data Augmentation, Fruit Image Recognition

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着生活水平不断提高,人们更加注重饮食健康,水果给人们带来丰富的营养,但识别水果品种效率低下,人工成本提高,大量水果种类识别成为亟待解决的问题,水果品质的好坏直接影响水果是否能够正确分类,因此,研究水果品质与分级分类[1]具有重要的现实意义。

传统的水果识别方式大多是利用图像获取颜色、纹理等特征完成识别。Chen 等人[2]采用多个网络模型融合和注意力机制方法,完成对数据图像从低维到高维的图像特征提取,减少特征提取干扰信息,最后通过 SoftMax 函数实现对苹果图像的分类,识别的准确率高达 98.14%。吕伟[3]提出金字塔匹配模型和 SIFT 相融合的水果分类方法,采用 SVM 分类器进行分类,最后分类准确率达到 98.9%。Dubey 等人[4]提出一种改进的和差直方图作为纹理特征,并使用支持向量机分类器进行水果分类,最后得出的分类

准确率高达 99%。Muhammad 等人[5]提出自动分类不同类型日期系统, 把日期的彩色图像分解成颜色分量, 然后将局部二进制或韦伯局部描述符直方图形式的局部纹理描述应用于每个组件, 最后用支持向量机输出分类结果, 该系统对日期分类的准确率超过 98%。卢巍[6]采用 K-means 聚类算法对异常大树杨梅 RGB 图像进行聚类, 并选取未成熟和成熟杨梅颜色数据均值作为两个质心, 通过多次迭代, 能够较好提取到数据边界的模糊信息, 从而更好的区分果实和背景数据信息。

深度学习作为机器学习领域[7]的一个重要分支, 在计算机视觉图像识别领域有着广泛的应用。Zhang 等人[8]提出了新型的人工智能系统, 使用旋转角度矢量网络的 2D 分数傅里叶熵, 提取水果图像中的数据特征, 之后, 将五层堆叠的稀疏自动编码器作为分类器, 以对 18 种类别水果进行分类, 分类识别准确率达到 95.08%。Majid 等人[9]通过串行方法的最大平均值合并预先处理好的两种类型特征, 然后使用基于谐波阈值遗传算法优化所得到的融合向量, 最后使用多个分类器对所选功能进行分类, 对七种类型水果分类准确率达到 99%。王前程[10]提出深度信念神经网络的水果图像识别方法, 通过 Census 变换提取图像特征, 并在深度信念神经网络中训练, 得到良好的识别效果, 减少模型训练时间。Wang 等人[11]提出一种改进的 CNN 应用于水果类别分类, 先创建了一个 8 层的卷积神经网络, 并利用参数整流线性单元代替了普通的整流线性单元, 在每个全连接层之前放置了 dropout 层, 最终网络达到 95.67% 的准确率。Jahanbakhshi [12]提出一种改进的深度卷积神经网络对胡萝卜样本图像进行分类, 通过混合池的批标准化方法, 最终达到分级的准确率为 99.43%。

为了解决传统水果分类识别算法成本高, 效率低下等问题[13], 本文采用迁移学习技术、混合注意力机制、全局平均池化、BN、dropout 等方法对 ResNet50 网络模型进行改进, 提出基于 CBAM-ResNet50 网络模型的水果图像识别算法, 并通过仿真实验验证了模型的有效性。

## 2. 材料与方

卷积神经网络是在传统的多层感知机(MLP)的基础上发展起来的, 是经典的机器学习算法, 通过权值共享、局部感受野来模拟人脑结构, 利用卷积、池化等操作提取输入图像的特征, 当前比较流行的网络结构模型有 AlexNet [14]、GoogleNet [15]、VGG [16]等, 网络结构对训练模型表达能力有较大的影响, 结构太浅的网络使得模型在训练过程中不能充分的学习到输入图像的特征, 网络结构太深会造成模型的训练时间过长, 网络的性能退化。CNN 主要由输入层、卷积层、池化层和全连接层三部分构成。卷积层和池化层是特征提取的核心模块, 卷积核体现的是参数共享机制。池化层利用图像局部相关性的原理对图像保留有用信息, 同时减少处理信息的数量, 通常在卷积层之后会得到维度很大的特征, 将特征切成几个区域, 取其最大值或平均值, 得到新的特征。

卷积神经网络经过多次卷积和池化堆叠操作, 提取到原始图像中的一般抽象化表示, 再通过全连接层对高维特征输入数据进行降维展平操作, 最后激活函数将结果输入到分类器中进行分类, 运算表达式如式

$$\mathbf{X}_j^L = f \left( \sum_{i \in M_j} \mathbf{X}_i^{L-1} * \mathbf{K}_{ij}^L + \mathbf{b}_j^L \right) \quad (1)$$

其中:  $L$  表示网络模型的层数,  $\mathbf{K}$  表示过滤器,  $\mathbf{b}$  表示为输出特征图的偏置项。

### 2.1. 实验环境配置与参数设置

计算机是 64 位 windows10 操作系统, 256 GB 内存, 4 GB 显存, 使用 pycharm 集成开发环境编写代码, 连接远程服务器, 租用 GPU, 在 TensorFlow 框架下完成的模型的训练和测试, 采用 Anaconda 4.5.11, CUDA 4.10.3; Python 3.9.7。

## 2.2. 数据收集及处理

实验数据集选用部分 kaggle 水果图像数据集和从网上爬取[17]一些水果图片, 以苹果、香蕉、葡萄、橙子、梨这 5 常见的水果类别为研究对象, 实验采集原始水果图像 1590, 将这 5 类常见的水果图片放在不同的文件夹, 一个类别代表一个文件。标签 0 为苹果, 标签 1 为香蕉, 标签 2 为葡萄, 标签 3 为橙子, 标签 4 为梨。将水果数据集按照随机取样的方法划分为 80% 训练集和 20% 测试集两部分, 训练集图片数量为 1274 张, 测试集图片数量 316 张。不同类别水果图像数量如表 1 所示。

**Table 1.** Number of images of different types of fruits

**表 1.** 不同种类水果图片数量

数据集	苹果	香蕉	葡萄	橙子	梨
训练集	308	256	173	267	270
测试集	76	64	43	66	67

网络模型在训练过程中, 样本空间不够或样本数量不足会出现影响模型泛化能力, 准确率不高等问题, 采用平移、旋转、水平翻转、错切变换等数据扩充方法对数据样本进行扩充, 提高了图像的准确性和多样性, 扩大样本空间容量。同时, 为了满足模型的输入要求, 需要将图像进行归一化操作[18], 通过数据规范化操作将数据样本图像尺寸整理成 224\*224\*3 像素, 然后将整理好的水果图像数据集送到网络模型中进行训练。

## 2.3. 迁移学习技术

卷积神经网络在图像分类[19]和识别方面发挥着重要作用, 从头开始训练深度 CNN 模型通常需要大量的数据和高计算能力, 此外, 如果使用少量样本数据集来训练深度卷积神经网络模型, 则会导致过拟合现象, 因此, 迁移学习技术[20]是解决这个问题的方法。

定义 1 领域(Domains)

$$D = \{X, P(x)\} \quad (2)$$

式中:  $X$  表示特征空间;  $P(x)$  表示为  $x = \{x_1, x_2, \dots, x_n\} \in X$ ,  $x_i \in \mathbf{R}^d, i = 1, 2, \dots, n$  的边缘分布,  $x_i$  表示第  $i$  幅图的特征,  $\mathbf{R}^d$  表示领域的特征集合。

定义 2 任务(Task)

$$T = \{Y, f(\cdot)\} \quad (3)$$

式中:  $Y$  表示类标签空间;  $f(\cdot)$  表示目标预测函数, 预测实例类标签。

把一个问题中已经训练好的模型通过一些简单的调整策略使其适用于一个新的问题。将任务训练的第二次重复使用作为模型的起点任务, 来解决目标领域中新数据集的学习任务, 减少模型的开发时间, 提高模型的性能。迁移主要是基于 CNN 模型和架构, 并将其作为特征提取器与逻辑回归分类相结合。首先, 对参数随机初始化设置, 通过迭代网络模型和优化参数, 使模型达到收敛某一值时输出, 然后将预训练好的网络模型迁移到目标数据集上, 再进行微调和网络优化, 使得模型能够更好的适应目标任务。将训练好的模型适用于目标任务, 从 ImageNet 数据集中学习到的特征与逻辑回归相结合作为通用分类器, 预测新域中的类标签。借助迁移学习, 可以节省时间和计算资源, 并通过有限的训练数据集成功的学习新任务。

## 2.4. ResNet50 网络

随着神经网络的不断发展, 过多的堆叠网络模型的深度会出现网络退化等问题, 从而使得模型的训练变得困难, he 等人[21]提出了深度残差网络 ResNet, 主要通过恒等映射将输入跨层传递, 残差块结构示意图如图 1, 在反向传播时, 为了避免网络出现梯度弥散问题, 该网络将输入中的一部分数据直接送入到输出, 而不经过程卷积网络, 并保留了一部分原始的信息, 从而大大加深了网络的深度, 减少参数信息的丢失。

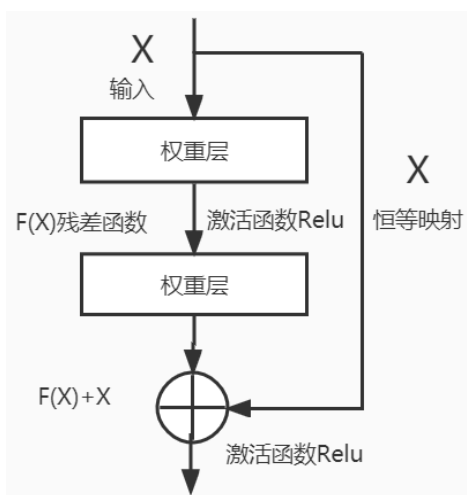


Figure 1. Residual structure of ResNet network

图 1. ResNet 网络残差结构

残差元有恒等映射和快捷连接, 恒等映射加深网络模型的深度, 而快捷连接使得残差变得可能, 在普通模块连接的过程中, 将其后面再加上几个恒等映射层(Identity mapping, 即  $y = x$ , 输出等于输入), 使得在增加网络深度的同时有效避免了训练误差的上升, 残差结构表示为

$$y_l = h(x_l) + F(x_l, W_l) \quad (4)$$

$$x_{l+1} = f(y_l) \quad (5)$$

$x_l$  和  $x_{l+1}$  分别表示残差单元第  $l$  和第  $l+1$  个输入,  $F(x_l, W_l)$  是要学习的残差映射,  $W$  是当前层需要学习的参数,  $f(y_l)$  是 relu 激活函数,  $h(x_l)$  表示恒等映射, 即输出等于输入, 由式(4) (5)进行递归运算, 得到残差单元  $l$  的特征表示, 计算公式如(6)所示

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i, W_i) \quad (6)$$

运用求导法则, 得出反向传播过程的梯度, 计算公式如式(7)所示:

$$\frac{\partial \text{loss}}{\partial x_l} = \frac{\partial \text{loss}}{\partial x_L} \cdot \frac{\partial x_L}{\partial x_l} = \frac{\partial \text{loss}}{\partial x_L} \cdot \left( 1 + \frac{\partial \text{loss}}{\partial x_L} \sum_{i=l}^{L-1} F(x_i, W_i) \right) \quad (7)$$

其中  $\frac{\partial \text{loss}}{\partial x_L}$  表示损失函数的梯度,  $L$  层的梯度可以直接传递到任何一个比它浅的  $L$  层,  $\frac{\partial \text{loss}}{\partial x_L} \sum_{i=l}^{L-1} F(x_i, W_i)$

表示残差梯度并且等于 1, 该梯度不会直接传递, 而是通过带有权重的层来传播梯度, 从而有效避免了学习过程中信息丢失的问题。



## 2.5. 注意力机制

注意力机制[22] [23] [24] [25]在计算机视觉、语音识别等领域中有着广泛的应用, 该技术的核心重点是对信息数据资源的高效分配, 能够让网络模型更多的关注应该关注的特征, 而不是什么都关注。注意力机制常见的实现方式[26]:

一是通道注意力机制: 是聚焦图像中特征内部之间关系的有用信息, 重点是获得输入进来的特征层和每一个通道中的特征信息;

二是空间注意力机制: 在通道上对输入进来的每一个特征点取最大值和平均值。之后将这两个结果进行一个堆叠, 利用一次通道数为1的7\*7卷积和激活函数sigmoid, 将获得了输入特征点的权值乘上原输入特征层, 进而能够更好的学习空间上的特征信息。

为了能够同时获取通道和空间上地有效信息, 采用混合注意力模块(CBAM)如图 2:

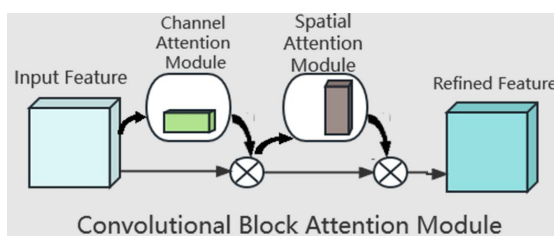


Figure 2. Mixed attention module  
图 2. 混合注意力模块

该模块是一种在图像分类中常用的深度学习模型, 能够将不同的注意力机制分配给不同的卷积核或池化层, 进而更好的捕捉到图像的特征信息。通过串联通道和空间注意力的方法, 避免了通道注意力机制不能很好地关注空间上的特征信息和空间注意力机制弱化了通道上的信息, 将两个不同的注意力机制结合起来, 能够获取更加全面的图像信息, 强化有效特征, 避免了不重要信息的提取, 进而提高了网络模型的性能。

## 3. 结果与分析

### 3.1. 网络模型构建

由于水果图像分类数据集和 ImageNet 数据集相差较大, 直接用 ResNet50 网络模型识别水果图像数据集, 会导致模型识别出来的效果不佳, 所以本文提出了 CBMA-ResNet50 模型对水果图像进行识别, 在初始 ResNet50 网络结构的卷积模块 2~4 直接后添加 CABM 注意力机制, 其后将平均池化(average pool)替换成全局平均池化(GlobalAvgPool2D), 并添加一个 Dropout 层和 BN 层, 来进行网络结构的优化, 网络结构如图 3 所示:

主要改进内容入下:

1) 迁移学习: 采用迁移学习技术[27], 将 ImageNet 数据集上训练好的权重参数迁移到 ResNet50 网络模型中, 使用这些训练好的权重信息, 能够加快对新任务数据集进行训练, 提高网络模型的训练速度和学习效率, 同时一定程度上避免了网络出现过拟合现象。

2) 添加 CBAM 注意力机制模块: 在卷积模块 2~4 之后添加混合注意力机制 CBMA, 通过空间和通道两个维度自动的学习输入到的权重信息, 使用学习权重分配强调数据区域中的重要特征, 进而提高网络模型的识别效果。同时 CBMA 模块内部只有少量的池化层和特征融合操作[28], 没有卷积结构, 这样的网络结构降低了模型的复杂度, 减少计算量, 进而提高网络模型的性能。

3) 设置全局平均池化(GlobalAvgPool2D), 添加至主干网络之后, 调整网络的输出尺寸, 按照中心化的方式将高维数据转化为低维数据, 然后对特征图求平均值, 这样, 每个数据点只能被计算一次, 提高了计算的效率, 进而提高了模型的泛化能力。

4) 添加 dropout 正则化, 将参数设置为 0.5, 使得网络模型在训练时将神经元随机丢弃一半, 梯度在正向传播时屏蔽对下游神经元的贡献如图 4。梯度在反向传播时, 不会更新神经元的权重, 也不会依赖其他神经元的输入特征, 不仅提升了模型的泛化能力, 也能够减少中间特征的数量, 从而减少冗余。

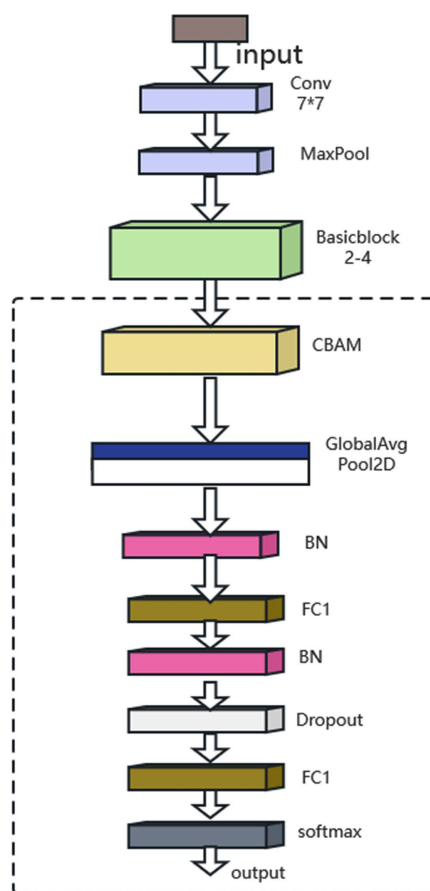


Figure 3. Improved ResNet50 network model  
图 3. 改进后 ResNet50 网络模型

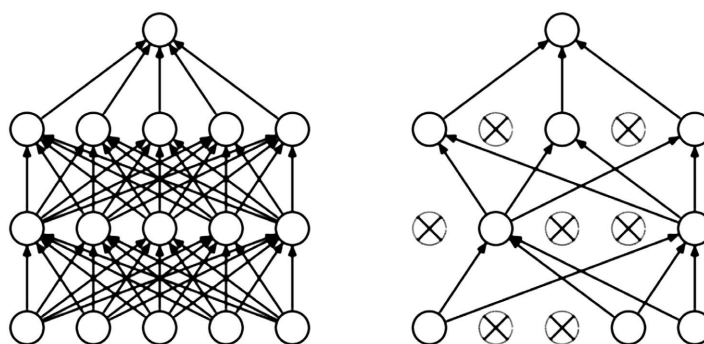


Figure 4. Neural networks without and with Dropout forward propagation  
图 4. 未使用和使用 Dropout 正向传播的神经网络

5) 构建 Batch Normalization 层, 大部分放在激活函数的后面[29], 主要是对隐藏层的输入服从均值为 0, 方差为 1 标准的正态分布, 使得领域的分布不再处于边界, 防止梯度的消失, 加快网络模型的收敛速度。

### 3.2. 水果图像识别性能对比试验

保持其他条件不变的情况下, 用本文模型与其他 5 种经典的神经网络模型进行比较, 根据水果图像数据集样本的情况和分类特点, 本实验选择准确率和损失值作为评价模型性能的指标, 每个模型均迭代 40 轮次, 每次迭代 5 就选择保存最好的一轮权重信息, 结合不同模型在测试集上的表现, 选择最好的网络模型, 结果如图 5 所示。

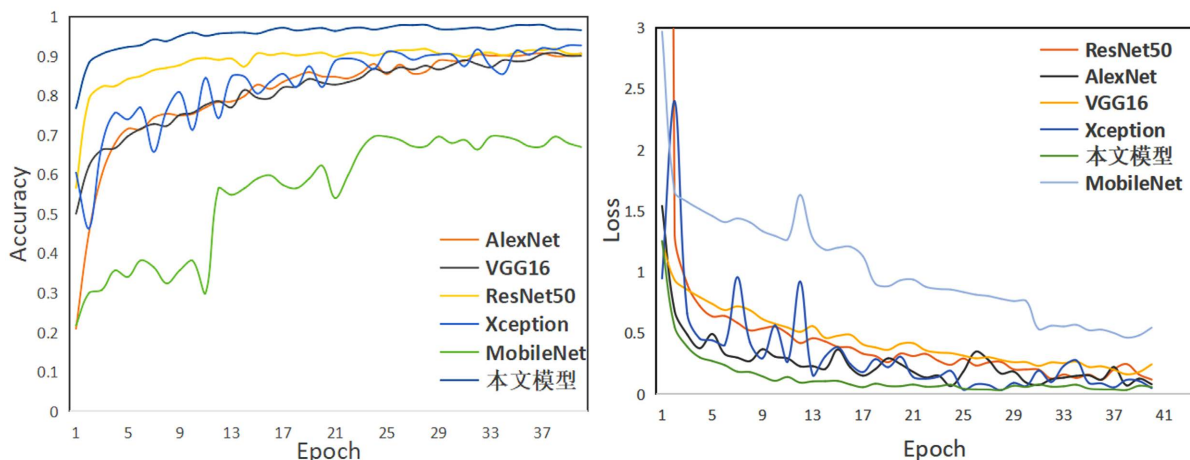


Figure 5. Accuracy and loss value curves of different network models  
图 5. 不同网络模型准确率和损失值曲线图

从图 5 可以看出, 本文提出的基于 CBMA-ResNet50 网络模型在损失值最低点和准确率达到最高点, 损失值为 0.032, 准确率为 99.3, 对水果图像分类的效果上优于其他模型。损失曲线的收敛速度也是最快的, 在迭代到第 5 次世代左右模型的收敛趋向于平稳, 这说明与其他模型相比较, 本文模型能够在较短的时间内能够达到最好的性能, MobileNet 模型在该数据集上模型性能表现最差, 准确率精度最高是 0.70, 识别效果不好; 且收敛速度较慢, 曲线不平滑, 网络波动大, 迭代到 20 世代模型才开始收敛; Xception 模型训练曲线全程波动较大, 模型不稳定; 其他网络模型在数据集上表现较为稳定, 均是在 25 轮次开始趋于收敛, 且曲线平滑。本文提出的 CBAM—ResNet50 模型训练集的准确率高达 99.6% 如表 2 所示, 与 MobileNet-v3、VGG16、AlexNet、Xception、ResNet50 模型相比较分别提高了 29、7、14、9、4 个百分点, 测试集准确率高达 99.3%, 与其他模型相比分别提高了 13、13、7、8、6 个百分点, 本文模型有效平衡了模型识别精度与收敛速度, 进而提高了网络模型的性能。

Table 2. Comparison effect diagram of network models  
表 2. 网络模型对比效果图

名称	训练精度	测试精度	训练损失	测试损失
MobileNet	70	60	0.2	0.26
VGG16	92	86	0.21	0.77
AlexNet	85	92	0.38	0.27

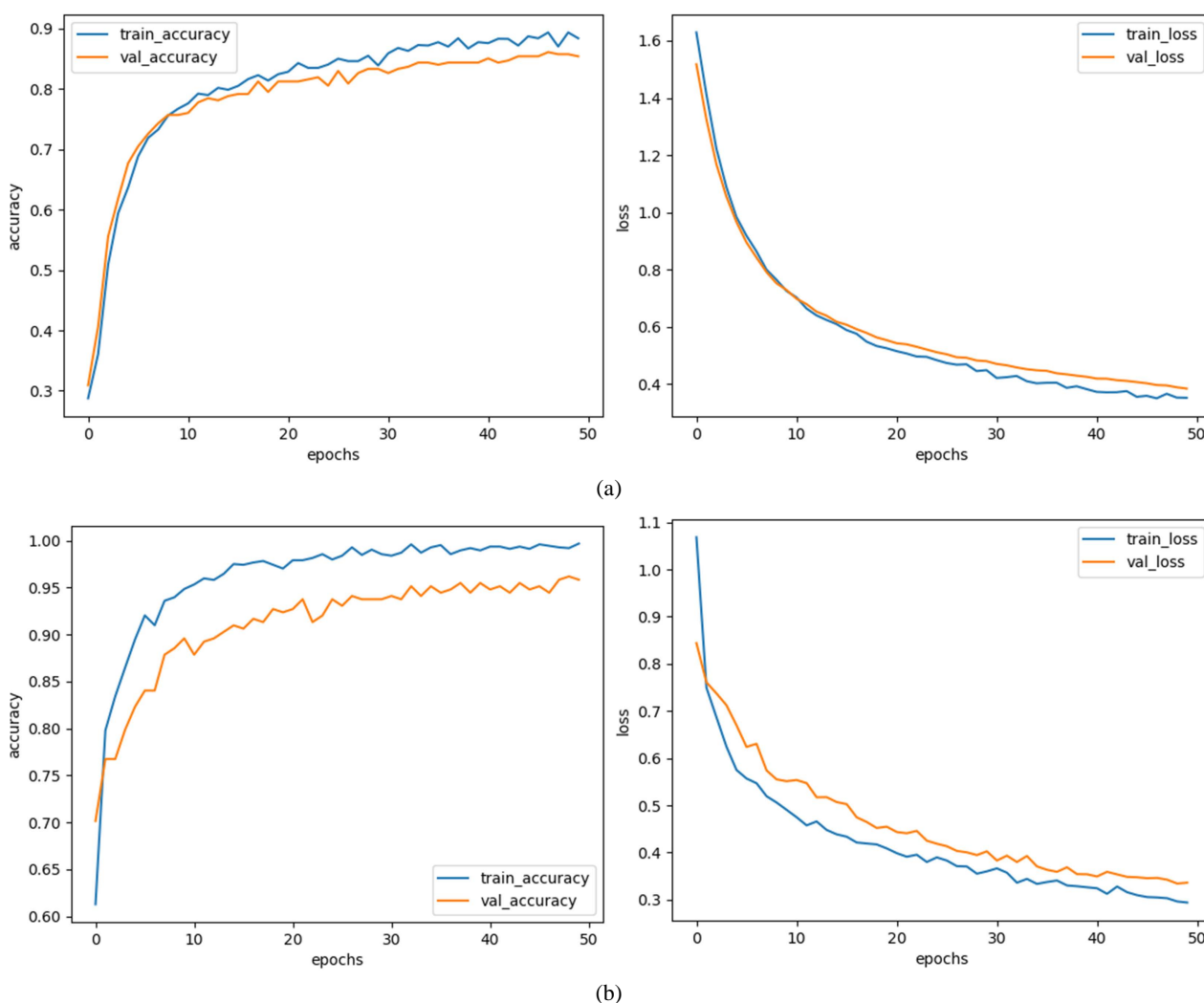


续表

Xception	90	91	0.09	0.03
ResNet50	95	93.7	0.07	0.09
本文模型	99.6	99.3	0.016	0.032

### 3.3. 迁移学习的数据增广对模型性能影响

为了验证数据增广对模型性能的影响, 在其他参数条件不变的情况下, 基于迁移学习分别对数据扩充和数据未扩充网络进行训练, 准确率和损失曲线如图 6 所示。



**Figure 6.** Data augmentation and non augmentation curves based on transfer learning; (a) Unexpanded training curve and loss curve of data; (b) Data augmentation training curve and loss curve

**图 6.** 基于迁移学习的数据扩充和未扩充曲线; (a)数据未扩充训练曲线和损失曲线; (b) 数据扩充训练曲线和损失曲线

没有数据增广的训练曲线和损失曲线之间存在一定的间隙, 数据质量得不到保障, 影响模型的识别度, 获取数据的特征不稳定。数据扩充之后得到的曲线在刚开始出现剧烈震荡, 随着迭代次数的增加波动的幅度慢慢趋于平稳, 减少了数据的丢失, 确保了数据的可靠性, 并且验证集的准确率达到 0.95 高于

未扩充后的验证集准确率 0.91, 提高了 4 个百分点。由此表明, 数据增广能够补全额外信息, 提高数据分析的速度和效率, 有效提升网络的性能。在迁移学习方式, 可以看出经过数据增广的训练集准确率并未有显著提升, CBMA-ResNet50 网络的迁移学习已经在 ImageNet 大型数据集上学习到了较好的特征提取能力, 进而导致数据增广的作用不那么明显, 模型在训练过程中波动较小; 验证集准确率刚开始达到了 0.9 以上, 网络模型迅速收敛, 在迭代 10 次后曲线趋于平稳, 且测试集准确率到达最高 0.95。上述结果表明, 经过迁移学习数据增广的测试集准确率高于没有扩充的测试集准确率, 加快了网络模型的收敛速度。数据增广优化了网络模型的效果, 一定程度上扩充了数据样本的多样性和变化性, 避免网络模型出现过拟合现象。

### 3.4. 注意力机制对实验结果的影响

为了验证注意力机制对模型性能的影响, 在其他参数条件不变的情况下, 通过加载预训练好的网络模型的参数, 使用迁移学习方法在残差结构之后添加混合注意力机制对水果图像数据集进行识别[30], 由图 7 可以看出

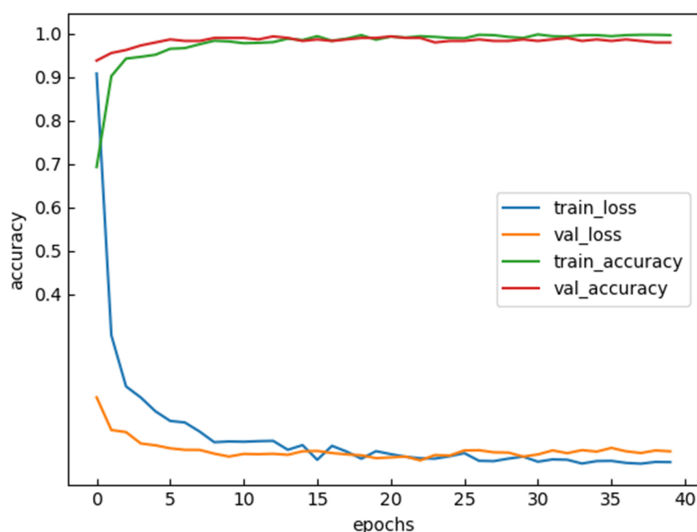


Figure 7. Accuracy of training and validation sets with attention mechanism  
图 7. 带有注意力机制的训练集和验证集的准确率

添加的混合注意力机制能够更快定位到水果图像有用的特征, 可以自适应地调整不同部分之间注意力权重信息, 获得有效的权重信息, 使得模型能够更好的理解输入的数据, 增强了网络模型的表达能力, 降低了噪声对实验结果的影响, 弱化了数据样本无效的特征, 提高了识别水果图像的效果, 准确率高达 0.99, 并且训练损失低于验证集损失, 曲线趋于平滑, 在水果图像数据集上表现良好, 进一步增强了网络模型的性能。

## 4. 讨论与结论

本文提出 CBAM—ResNet50 网络模型, 采用数据增强技术, 通过训练好的权重信息进行迁移学习, 在残差结构模块后添加混合注意力机制, 并优化模型参数对改进的网络进行训练。进一步分析了图像数据增广、注意力机制超参数等方法对模型的影响, 得出以下结论:

通过对大型数据集训练, 获取训练好的权重信息, 提取水果图像的特征, 水果图像数据集的增加, 可以了解数据的分布特征, 使得模型更好的适应了水果图像的数据分布, 提高了模型的鲁棒性, 数据扩

充不仅丰富了数据样本的多样性和变化性, 而且能够有效提高网络的性能。

2) 迁移学习将训练好的权重信息迁移到新的网络模型中, 使得模型能够快速学习, 提高新任务的学习速度, 降低模型的复杂度, 增强了模型的鲁棒性。

3) 混合注意力机制使得网络关注重要特征, 抑制无关信息, 弱化了图片噪声对网络的影响, 提高了提取信息的能力, 同时, CBMA 模块中没有大量的卷积操作, 不需要太多计算量, 从而降低网络模型复杂度, 提升模型的性能。

4) 网络结构中添加全局平均池化减少了模型的参数量, 降低了模型的复杂度; Dropout 正则化一定程度上避免了过拟合现象; 批量归一化(BN)加快了模型的收敛速度, 提高了模型的识别效率。

## 基金项目

国家自然科学基金项目(61751313); 甘肃省重点研发计划项目(21YF5GA088); 甘肃省教育信息化建设专项任务项目(2011-02)。

## 参考文献

- [1] 唐世裔, 钟高峰, 钟少伟. 特色水果高产优质栽培新技术[M]. 长沙: 湖南科学技术出版社, 2017: 255.
- [2] Chen, J., Han, J., Liu, C., *et al.* (2022) A Deep-Learning Method for the Classification of Apple Varieties via Leaf Images from Different Growth Periods in Natural Environment. *Symmetry*, **14**, Article No. 1671. <https://doi.org/10.3390/sym14081671>
- [3] 吕伟. 基于稀疏表示和卷积神经网络的水果图像分类与实现[D]: [硕士学位论文]. 广州: 华南农业大学, 2016.
- [4] Dubey, S.R. and Jalal, A.S. (2012) Robust Approach for Fruit and Vegetable Classification. *Procedia Engineering*, **38**, 3449-3453. <https://doi.org/10.1016/j.proeng.2012.06.398>
- [5] Muhammad, G. (2015) Date Fruits Classification Using Texture Descriptors and Shape-Size Features. *Engineering Applications of Artificial Intelligence*, **37**, 361-367. <https://doi.org/10.1016/j.engappai.2014.10.001>
- [6] 卢巍. 基于 K-means 聚类的计算机视觉下杨梅水果图像识别[J]. 现代计算机, 2022, 28(8): 78-81.
- [7] (美) 弗朗索瓦·肖莱. Python 深度学习 人工智能 Deep Learning with Python 中文版[M]. 张亮, 译. 北京: 人民邮电出版社, 2018.
- [8] Zhang, Y.D., Satapathy, S.C. and Wang, S.H. (2022) Fruit Category Classification by Fractional Fourier Entropy with Rotation Angle Vector Grid and Stacked Sparse Autoencoder. *Expert Systems*, **39**, e12701. <https://doi.org/10.1111/exsy.12701>
- [9] Majid, A., Khan, M.A., Alhaisoni, M., *et al.* (2022) An Integrated Deep Learning Framework for Fruits Diseases Classification. *Computers, Materials & Continua*, **71**, 1387-1402. <https://doi.org/10.32604/cmc.2022.017701>
- [10] 王前程. 基于深度学习的水果图像识别算法研究[D]: [硕士学位论文]. 保定: 河北大学, 2016
- [11] Wang, S.-H. and Chen, Y. (2020) Fruit Category Classification via an Eight-Layer Convolutional Neural Network with Parametric Rectified Linear Unit and Dropout Technique. *Multimedia Tools and Applications*, **79**, 15117-15133. <https://doi.org/10.1007/s11042-018-6661-6>
- [12] Jahanbakhshi, A., Momeny, M., Mahmoudi, M., *et al.* (2021) Waste Management Using an Automatic Sorting System for Carrot Fruit Based on Image Processing Technique and Improved Deep Neural Networks. *Energy Reports*, **7**, 5248-5256. <https://doi.org/10.1016/j.egy.2021.08.028>
- [13] 谭笑枫, 李广帅. 深度学习发展综述[C]//第十五届国家安全地球物理专题研讨会论文集. 北京: 清华大学出版社, 2019: 244-252.
- [14] 朱佳欣. 基于机器学习的图像检索算法研究 111[D]: [硕士学位论文]. 北京: 北京邮电大学, 2018.
- [15] Szegedy, C., Liu, W., Jia, Y., *et al.* (2015) Going Deeper with Convolutions. *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>
- [16] Weiss, K., Khoshgoftaar, T.M. and Wang, D.D. (2016) A Survey of Transfer Learning. *Journal of Big Data*, **3**, 1-40. <https://doi.org/10.1186/s40537-016-0043-6>
- [17] 李乔宇, 尚明华, 王富军, 等. 基于 Scrapy 的农业网络数据爬取[J]. 山东农业科学, 2018, 50(1): 142-147.

- [18] 张航, 程清, 武英洁, 等. 一种基于卷积神经网络的小麦病害识别方法[J]. 山东农业科学, 2018, 50(3): 137-141.
- [19] 黎英, 宋佩华. 迁移学习在医学图像分类中的研究进展[J]. 中国图象图形学报, 2022, 27(3): 672-686.
- [20] 邵浩著. 迁移学习理论与实践[M]. 上海: 上海交通大学出版社, 2013: 121.
- [21] He, K., Zhang, X., Ren, S., *et al.* (2016) Deep Residual Learning for Image Recognition. *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 770-778.  
<https://doi.org/10.1109/CVPR.2016.90>
- [22] 任欢, 王旭光. 注意力机制综述[J]. 计算机应用, 2021, 41(S1): 1-6.
- [23] 祁宣豪, 智敏. 图像处理中注意力机制综述[J]. 计算机科学与探索, 2024, 18(2): 345-362.
- [24] 皇甫晓瑛, 钱惠敏, 黄敏. 结合注意力机制的深度学习综述[J]. 计算机与现代化, 2023(2): 40-49+57.
- [25] 李清格, 杨小冈, 卢瑞涛, 等. 计算机视觉中的 Transformer 发展综述[J]. 小型微型计算机系统, 2023, 44(4): 850-861.
- [26] 杨环宇, 王军, 吴祥, 等. 一种坐标通道注意力深度学习网络的军用飞机识别方法[J]. 兵工学报, 2023: 1-15.
- [27] 邹建. 基于卷积注意力特征迁移学习的滚动轴承故障诊断[J]. 计算机测量与控制, 2024, 32(1): 23-29.
- [28] 林海波, 卢元栋, 丁荣诚, 等. 基于图像处理与改进 SVM 的苹果多特征融合分级方法[J]. 山东农业科学, 2022, 54(6): 141-149.
- [29] 赵小阳, 李仲年, 王文玉, 等. ADIC: 一种面向可解释图像识别的自适应解纠缠 CNN 分类器[J]. 计算机研究与发展, 2023, 60(8): 1754-1767.
- [30] 李大湘, 张玥. 基于多尺度 CNN 特征的国画图像分类算法[J]. 西安邮电大学学报, 2021, 26(1): 104-110.