

基于机器视觉的水果种类识别

武聪*, 黄健, 何才壮, 徐前进

西京学院计算机学院, 陕西 西安

收稿日期: 2024年12月22日; 录用日期: 2025年1月19日; 发布日期: 2025年1月28日

摘要

近年来, 随着计算机技术的发展, 机器视觉已经在各个领域得到了广泛的应用。其中, 利用机器视觉技术来对水果种类进行精确识别是一直以来的研究热点。水果销售作为餐饮零售业的重要分支, 正逐渐引入无接触服务的概念。通过非接触式水果销售机器人, 消费者可以享受到安全、卫生、快速的购买体验, 同时商家也因此获得更高效的销售方式和更为便捷的管理手段, 其目的是实现自动化水果种类识别, 节省农业劳动力资源, 提高生产效率, 促进智慧农业的发展。因此本课题利用K210视觉模块对水果进行图像采集, 将采集的图像进行预处理、立体校正, 接着在Mixhub开源数据平台进行模型训练, K210视觉模块不仅可以降低成本的使用, 还可以精确识别水果的种类。

关键词

机器视觉, 水果识别, K210视觉模块

Fruit Type Recognition Based on Machine Vision

Cong Wu*, Jian Huang, Caizhuang He, Qianjin Xu

School of Computer Science, Xijing University, Xi'an Shaanxi

Received: Dec. 22nd, 2024; accepted: Jan. 19th, 2025; published: Jan. 28th, 2025

Abstract

In recent years, with the development of computer technology, machine vision has been widely applied in various fields. Among them, the precise identification of fruit types using machine vision technology has been a research hotspot for a long time. Fruit sales, as an important branch of the catering and retail industry, are gradually introducing the concept of contactless service. Through

*通讯作者。

contactless fruit vending robots, consumers can enjoy a safe, hygienic, and fast purchasing experience, while businesses obtain more efficient sales methods and more convenient management tools. The aim is to achieve automated fruit type recognition, save agricultural labor resources, increase production efficiency, and promote the development of smart agriculture. Therefore, this project utilizes the K210 vision module to capture images of fruits, preprocess and rectify the captured images in three dimensions, and then trains models on the Mixhub open-source data platform. The K210 vision module not only reduces costs but also accurately identifies fruit types.

Keywords

Machine Vision, Fruit Recognition, K210 Vision Module

Copyright © 2025 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

自古以来,我国农业的发展备受重视,其中水果种植更是占据重要地位,种植面积广阔且种类繁多,使得我国水果产量在全球名列前茅,成为继蔬菜和粮食之后的第三大农业种植产业[1]。然而,随着水果交易量的急剧增长,水果采摘后的商业化处理效率却显得捉襟见肘,高交易量与低分类效率之间的矛盾日益凸显。当前,水果商户在交易过程中仍普遍依赖人工识别水果种类。在批发市场和水果超市中,消费者往往需要排队等待工作人员人工识别所购水果的种类,才能获得价格数据、计算交易金额并贴上标签[2]。这种交易方式不仅耗时较长、效率低下,还造成了人力和财力的浪费,同时也给消费者带来了不便[3]。针对这一问题,将水果图像的自动识别技术应用于水果超市或批发市场的分类中,无疑是一个明智的选择[4]。这不仅可以解放工作人员繁重的水果分类任务,节省大量劳动力资源,提高效率,还能推动我国智慧农业的发展和无人超市的普及[5]。

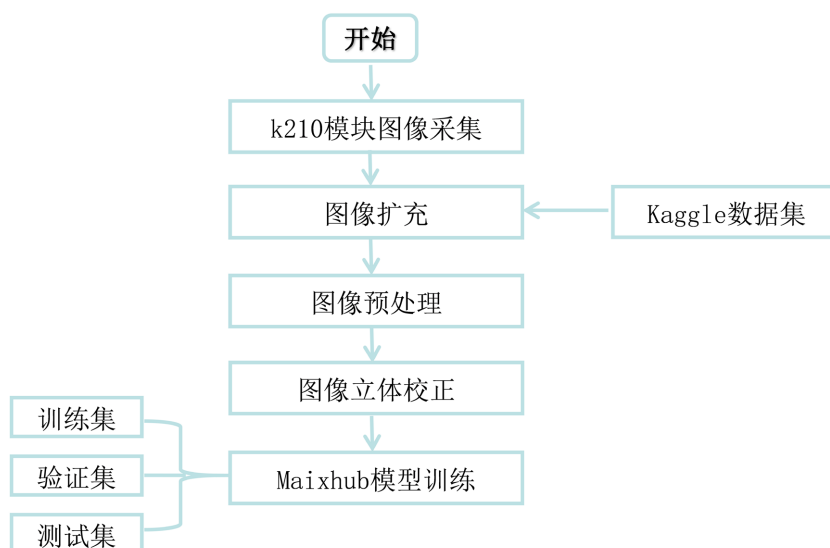


Figure 1. Overall design process

图 1. 整体设计流程

近年来, 计算机视觉技术和数字图像处理技术取得了飞速发展, 同时, 我国也出台了一系列有利于农业发展的政策, 加速了农业数字化、信息化、自动化的进程[6]。随着传统农业与信息技术的深度融合, 我国农业生产效率显著提升, 农业焕发出新的生机, 为经济发展注入了新动力[7]。作为我国主要农作物之一的水果, 其自动识别技术的研究对于提高水果采摘、交易等处理环节的效率具有重要意义[8]。目前, 这一领域已经吸引了国内外众多学者的关注和研究[9]。水果分类技术的研究主要包括水果特征集的获取和利用训练特征集进行模式识别两个环节, 涉及图像处理、机器识别等多个领域的知识。在水果的二维图像分析方面, 研究人员已经取得了诸多成果[10]。然而, 在机器识别精确度方面, 仍存在一些问题。

综上所述, 本文针对以上问题, 利用 K210 视觉模块对水果进行图像采集, 将采集的图像进行预处理、立体校正, 接着在 Mixhub 开源数据平台进行模型训练, K210 视觉模块不仅可以降低成本的利用, 还可以精确识别水果的种类。整体流程如图 1 所示。

2. 系统预处理设计

2.1. K210 视觉模块概述

K210 芯片集成了先进的卷积神经网络硬件加速器(KPU), 专为深度学习推理任务设计, 能高效处理图像数据。其内置 64 位双核 CPU, 主频达 400 MHz, 搭配 8MB 片内 SRAM, 满足图像缓存与中间数据暂存。K210 视觉模块水果识别功能详解 K210 视觉模块在水果识别领域表现卓越, 为智能农业、生鲜零售等行业注入革新力量, 解锁高效精准的水果分拣与管理方案。实物如图 2 所示。

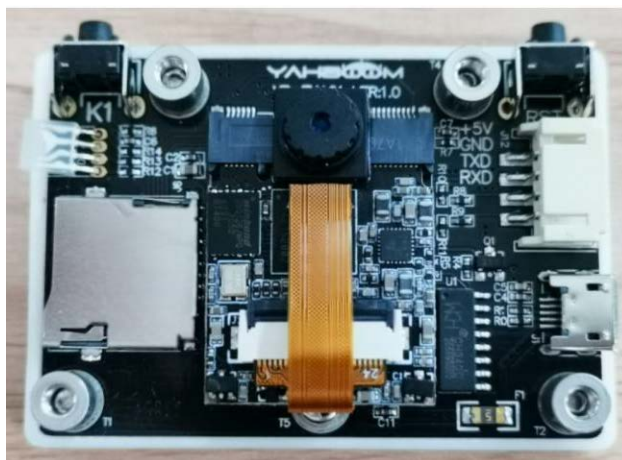


Figure 2. K210 vision module
图 2. K210 视觉模块

2.2. 数据集收集

在本课题的研究进程中, 水果检测网络训练所需的数据集主要源于自制。于水果类别拣选方面, 精准聚焦了四大品类, 分别为红苹果、葡萄、梨、香蕉, 为构建充足的数据集, 在自制数据与网络图片采集双轨并行之下, 累计收获各类水果图片各 300 张, 总量达 2000 张之多。自制数据使用 K210 视觉模块自带的 TFT 卡将拍摄的照片存在里面, 灵活切换拍摄角度, 从俯拍、仰拍、侧拍等多元视角全方位捕捉水果的千姿百态, 只为选取水果在不同情境下的形态特征, 极大扩充数据多样性。收集的水果图片尺寸参差不齐, 为契合训练规范, 运用专业图像处理手段, 一律将所有采集到的水果数据规整至 224×224 像素规格, 达成数据标准化, 随后即刻步入关键的数据标注环节, 数据集创建如图 3 所示。



Figure 3. Dataset creation

图 3. 数据集创建

2.3. 训练平台概述

Maixhub 训练平台简介: Maixhub 是一个功能强大的训练平台, 主要用于机器学习和深度学习相关模型的训练, 特别是在嵌入式视觉领域发挥着重要作用。对于模型训练, Maixhub 拥有丰富的模型库和训练算法。用户可以根据自己的需求选择合适的模型架构, 如卷积神经网络(CNN)等。同时, 平台能够根据用户的数据自动调整训练参数, 以达到较好的训练效果。在应用场景上, Maixhub 训练出的模型可以很好地应用于 K210 视觉模块。在水果识别、物体检测等多个领域实现智能识别, 为开发智能视觉应用提供了极大的便利, 如图 4 所示。



Figure 4. Maixhub dataset training platform

图 4. Maixhub 数据集训练平台

3. 水果识别软件系统设计

3.1. 算法原理

(一) 卷积神经网络(CNN)原理: K210 视觉模块采用 CNN 来提取水果图像的特征。CNN 的卷积层是核心部分, 其包含多个卷积核。以识别苹果为例, 当图像输入后, 卷积核在图像上滑动, 每个卷积核对局部像素进行加权求和运算。例如, 一个 3×3 的卷积核在图像上滑动时, 会与对应的 3×3 像素区域进行乘法和加法操作, 提取出如苹果的圆形轮廓、颜色分布等特征, 生成特征图。池化层则对特征图进行下采样操作。

(二) 目标检测 YOLOV3 算法: 在水果种类识别中, YOLOV3 算法将输入图像划分成多个网格单元。

每个网格单元负责预测水果的边界框和类别概率。例如，将图像划分为 13×13 个网格单元，对于不同光照条件下的苹果，该网格单元会预测草莓的边界框位置(包括中心坐标、宽度和高度)和它是苹果的概率。YOLOV3 把目标检测看作一个回归问题，直接从图像像素预测目标的类别和位置，这样能够快速地进行检测，适合在 K210 这种资源有限的模块中使用。

3.2. 数据训练参数

学习率是训练过程中的重要参数。对于 K210 视觉模块进行水果种类识别，学习率一般选择一个较小的值，如 0.001~0.01。如果学习率过大，模型可能无法收敛，甚至出现梯度爆炸的情况；如果学习率过小，训练过程会非常缓慢。为了增加数据集的多样性，提高模型的泛化能力，通常会进行数据增强。数据增强的参数包括旋转角度、平移量、缩放比例等。这些操作可以使模型更好地适应水果在不同姿态和位置下的识别，同时对训练次数，学习率，还有批量大小做了数据统计，下表是水果识别对数据进行采样，一些数据统计，如表 1 所示。

Table 1. Statistics on fruit recognition data

表 1. 水果识别数据统计

参数配置	训练次数	批量大小	学习率	标注框限制	数据均衡	允许负样本
苹果	200	32	0.001	10	关	关
香蕉	210	34	0.002	13	开	关
梨	220	31	0.003	24	关	开
葡萄	320	34	0.002	28	关	关

3.3. 图像采集

在面对零售场景时，例如精品水果店的陈列展示区域或是智能生鲜售卖机的内部，摄像头更多聚焦于采集水果的静态图像。这些静态图像能够清晰定格水果在最佳展示状态下的外观全貌，其色泽、形态等细节一览无余，从而极为契合零售场景下对于水果精准识别与产品推介的需求，助力消费者在选购过程中获得准确无误的水果信息引导。

3.4. 图像预处理

图像预处理阶段作为承上启下的关键环节，首先，是灰度化转换操作。借助成熟且精准的加权平均法，依据人眼对不同颜色敏感程度赋予红、绿、蓝三色通道相应权重，将彩色图像有条不紊地转换为灰度图像；或是运用亮度提取公式，巧妙提取图像亮度信息以达成灰度化目的。

3.5. 图像特征提取

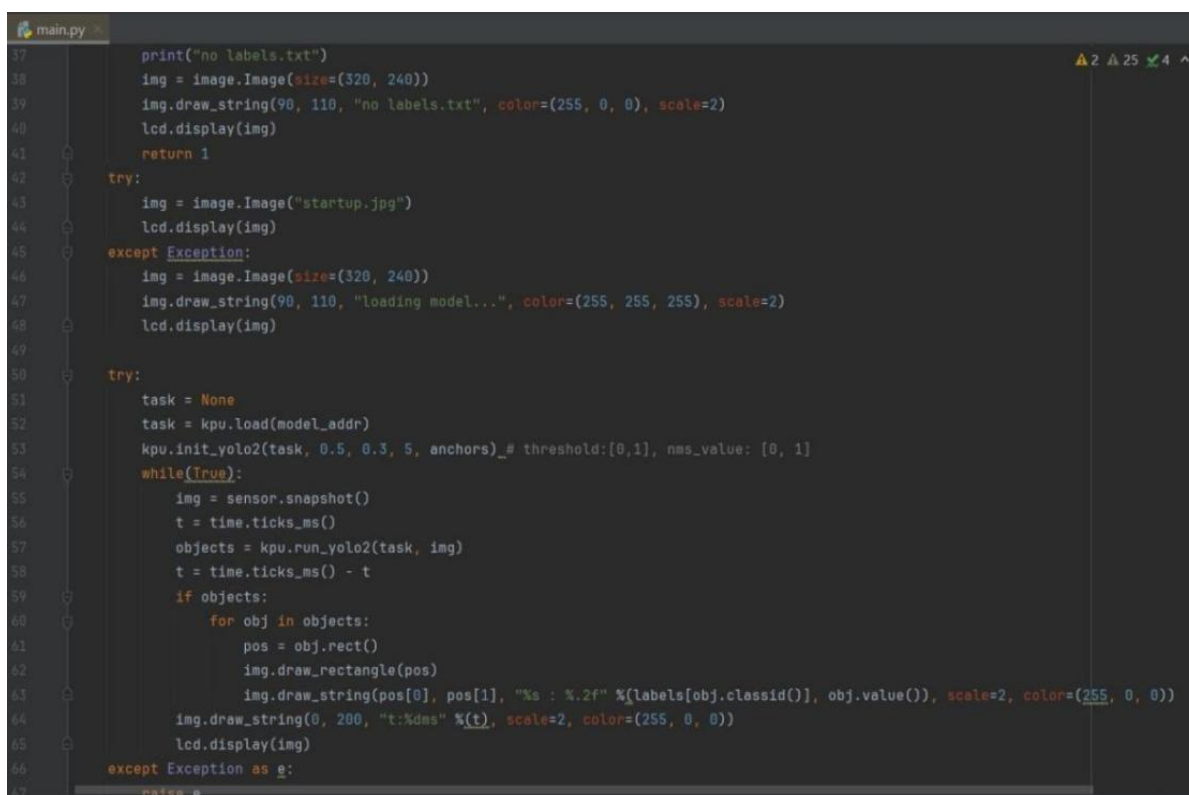
通过细致划分颜色区间，统计各区间内像素数量占比，精准量化水果表面不同颜色分布态势，红色苹果的鲜艳红调为水果初步分类提供鲜明色彩线索。转向深度学习领域，充分发挥 K210 视觉模块内置 KPU 强大算力优势，加载精心设计的轻量级卷积神经网络模型。

4. 实验与结果分析

4.1. 配置 K210 开发环境

下载 K210 对应的 SDK (软件开发工具包)，并解压到合适的本地目录。SDK 中包含了编译、烧录等

相关工具以及 K210 相关的底层代码资源等。在 Pycharm 中配置好相应的编译器路径,使其能找到 K210 的交叉编译工具链(在 SDK 里),以便后续能将代码编译成可在 K210 上运行的二进制文件。图像采集代码编写假设 K210 连接到电脑的 COM3 端口(实际根据你电脑上识别到的端口号为准),波特率等参数按 K210 默认或实际配置来设置,代码调试如图 5 所示。



```

37 print("no labels.txt")
38 img = image.Image(size=(320, 240))
39 img.draw_string(90, 110, "no labels.txt", color=(255, 0, 0), scale=2)
40 lcd.display(img)
41 return 1
42
43 try:
44     img = image.Image("startup.jpg")
45     lcd.display(img)
46 except Exception:
47     img = image.Image(size=(320, 240))
48     img.draw_string(90, 110, "loading model...", color=(255, 255, 255), scale=2)
49     lcd.display(img)
50
51 try:
52     task = None
53     task = kpu.load(model_addr)
54     kpu.init_yolo2(task, 0.5, 0.3, 5, anchors)_# threshold:[0,1], nms_value: [0, 1]
55     while(True):
56         img = sensor.snapshot()
57         t = time.ticks_ms()
58         objects = kpu.run_yolo2(task, img)
59         t = time.ticks_ms() - t
60         if objects:
61             for obj in objects:
62                 pos = obj.rect()
63                 img.draw_rectangle(pos)
64                 img.draw_string(pos[0], pos[1], "%s : %.2f" % (labels[obj.classid()], obj.value()), scale=2, color=(255, 0, 0))
65         img.draw_string(0, 200, "t:%dms" % (t), scale=2, color=(255, 0, 0))
66         lcd.display(img)
67     except Exception as e:
68         raise e

```

Figure 5. Image acquisition

图 5. 对图像进行采集

4.2. 实验模型分析

4.2.1. 模型性能评估

精确率衡量的是模型预测为正类(某水果种类)且预测正确的样本数占有所有被预测为正类样本数的比例。以识别苹果为例,模型预测有 200 个样本为苹果,其中真正是苹果的有 175 个,苹果类别的精确率是 87.5%。召回率反映的是实际为正类(某种水果)的样本中,被模型正确识别出来的比例。数据集中共有 200 个苹果样本,模型识别出了 160 个,召回率是即 80%。召回率越高,意味着模型越不容易遗漏目标水果,对数据集中各类水果的覆盖越全面。当召回率低时,可能有大量水果样本未被识别,影响实际应用效果。

4.2.2. 模型局限性

小目标识别困难,当水果尺寸较小,如葡萄,K210 结合 YOLOV3 算法表现欠佳。原因在于 YOLOV3 的网络结构在处理图像时会经过多次下采样,小水果原本就稀少的特征信息在这个过程中大量流失。光照适应性差水果在不同光照条件下外观差异巨大。强光下苹果的红色区域反光强烈,部分纹理被掩盖;弱光中香蕉的黄色暗沉,颜色特征模糊,与其他类似颜色水果难以区分。

4.2.3. 改进方向

运用模型剪枝技术, 依据神经元、卷积核对识别结果的重要性, 去除贡献小的部分。对于 K210, 剪掉模型中对水果识别影响微弱的结构, 既能缩小模型体积, 减轻内存压力, 又能在一定程度上维持甚至提升性能, 缓解模型复杂度与资源的矛盾。数据增强与预处理, 为克服光照问题, 丰富数据采集, 涵盖各种光照下的水果图像, 从不同时段、天气、室内外场景收集。训练前对图像归一化, 将像素值规范到统一区间, 如 $[0, 1]$, 弱化光照影响, 让模型聚焦水果本质特征而非表面光影变化, 提升光照条件下的精确率。采用多样化图像增强, 像随机裁剪、翻转、旋转、添加噪声等。以随机裁剪为例, 模拟水果在不同视角、遮挡下的形态, 增加样本多样性。对于 K210, 这些操作贴合实际水果场景, 扩充的样本有助于模型泛化, 提高应对复杂实际情况的能力。

4.2.4. 不同模型对比分析

假设存在基于传统机器学习方法(如 SVM)和其他深度学习模型(如 SSD)构建的水果识别模型, 与 K210 搭载的 YOLOV3 模型进行对比: 精确率与召回率 SVM 模型通常依赖手工设计特征, 对于水果复杂多变的外观, 特征提取能力有限, 可能导致精确率和召回率在多类别水果识别时相对较低。资源占用对比: SVM 模型结构简单, 只需存储支持向量等少量数据, 模型极小, 在 K210 内存中占用空间微不足道。YOLOV3 模型即使经过优化, 由于深度学习架构复杂, 仍需占用一定内存空间存储权重、参数等。SSD 模型同样需要一定存储空间, 与 YOLOV3 大致相当, 但在具体参数设置不同时会有差异。

4.3. 数据增强对比分析

数据增强是一种在机器学习和深度学习领域广泛应用的技术手段。以水果种类识别为例, 原始数据可能是从不同角度拍摄的水果照片, 但角度范围有限。通过数据增强的旋转操作, 可以模拟出更多角度下水果的外观, 让模型能够学习到水果在各种姿态下的特征。颜色变换则有助于模型适应不同光照条件下水果的颜色变化, 比如在强光、弱光或者不同颜色灯光下水果颜色的改变。添加噪声则能使模型对图像采集过程中可能出现的干扰更加鲁棒。数据增强前后对比如图 6、图 7 所示, 在验证集上第 100 次训练迭代有最佳准确率: 0.603, 经过数据增强后, 在验证集上第 100 次训练迭代有最佳准确率: 0.855。

在自然语言处理领域, 数据增强技术也有诸多应用, 图 8 是没有进行数据增强对实验, 图 9 是进行数据增强后损失精度的实验。

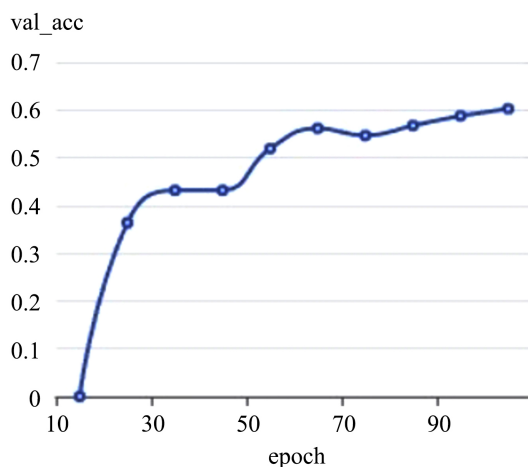


Figure 6. Best accuracy before data augmentation

图 6. 数据增强前最佳准确率

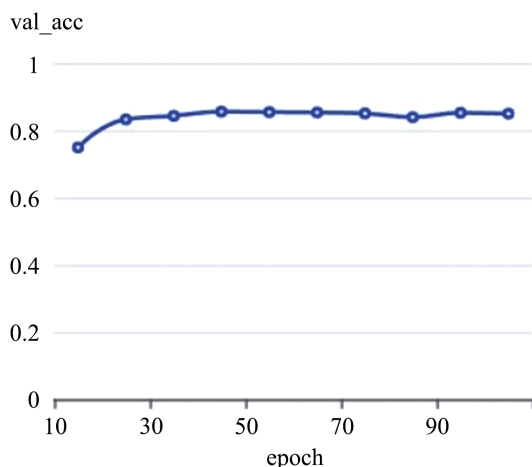


Figure 7. Best accuracy after data augmentation
图 7. 数据增强后最佳准确率

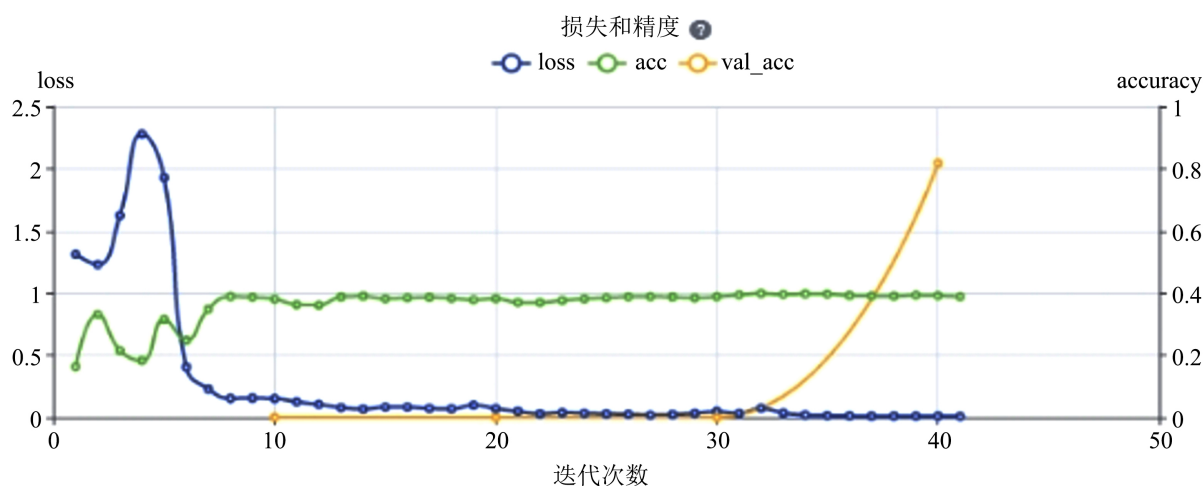


Figure 8. Experiment on loss and accuracy without data augmentation
图 8. 无数据增强损失精度实验

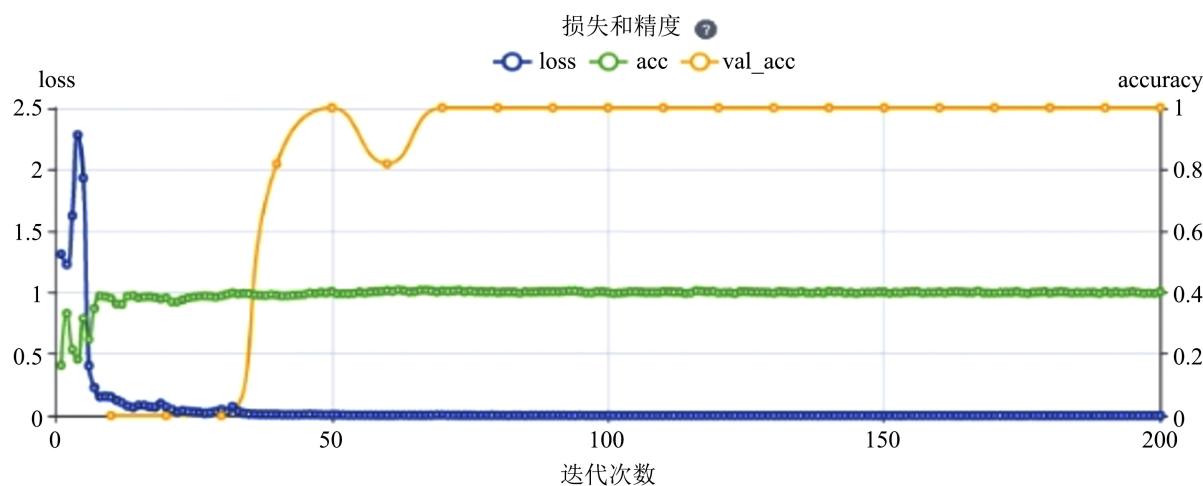


Figure 9. Experiment on loss and accuracy after data augmentation
图 9. 数据增强后损失精度实验

从图 8 可以看到, 损失函数从第 43 次开始迭代, 数据增强后损失函数从第 25 次开始迭代, 数据网络优化经过数据增强后从 20, 变成了 12, 数据越小网络优化越快, 数据精度从 30, 变成了 60, 精度提高了, 总体来说训练的准确度比没数据增强的高出不少, 而且迭代次数也比较快。

用 K210 视觉模块来检测识别目标, 选取香蕉, 苹果, 梨为例, 从图 10 可以看出香蕉识别率很高。图 11 看出受光照影响, 梨的识别精度会降低, 图 12 经过水果识别精度检测可以观察到, 通过数据增强和特定预处理去除恶劣环境对检测识别的影响后, 苹果检测精度提高, 证明 K210 模块能够更准确地检测自然条件下的目标, 为智慧农业的发展奠定了理论基础。



Figure 10. Effect diagram of banana detection
图 10. 香蕉检测效果图



Figure 11. Detection effect diagram of pears
图 11. 梨检测效果图

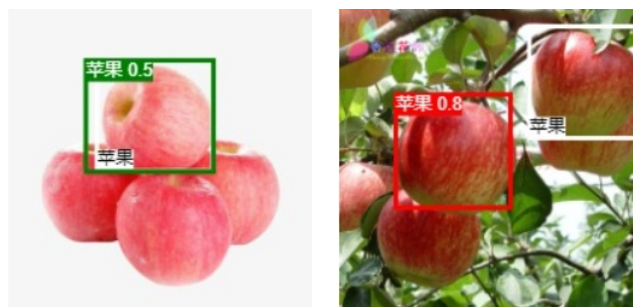


Figure 12. Effect diagram of apple detection
图 12. 苹果检测效果图

5. 结论与展望

本研究利用 K210 视觉模块成功搭建高效水果种类识别系统,经严谨实验验证其精度与速度可行性,为水果产业前端分拣、后端销售赋能,降本增效显著。未来,拓展水果品类至小众、异形种类,融合更多传感数据如近红外光谱补全特征;优化模型压缩技术,将复杂网络部署更小型设备;结合 5G 实现果园云端协同识别,远程监控水果生长、品质全流程,全方位深化机器视觉在水果产业革新应用。综上, K210 视觉模块于水果种类识别前景广阔,持续迭代将重塑水果产业链智能化格局,开启精准农业与智能商业新篇章,在全球水果贸易与消费升级浪潮中扮演关键技术推手角色,助力产业多元创新发展。

参考文献

- [1] Yu, Y., Liu, Y., Li, Y., Xu, C. and Li, Y. (2024) Object Detection Algorithm for Citrus Fruits Based on Improved YOLOv5 Model. *Agriculture*, **14**, Article 1798. <https://doi.org/10.3390/agriculture14101798>
- [2] 盛德庆. 基于深度学习的水果采摘机器人目标识别定位研究[D]: [硕士学位论文]. 吉林: 吉林化工学院, 2024.
- [3] 罗家梅. 改进 YOLOv5 的水果品质检测与分类方法研究[D]: [硕士学位论文]. 赣州: 赣南师范大学, 2023.
- [4] 洪馨. 基于改进型残差网络的水果品种分类方法研究[D]: [硕士学位论文]. 南昌: 南昌工程学院, 2023.
- [5] Al-Nawah, Mohammed. 基于神经网络的蔬菜水果识别和价格评估方法研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2023.
- [6] 郑浩. 基于深度学习的水果识别与分割技术研究[D]: [硕士学位论文]. 西安: 西安工业大学, 2023.
- [7] 姚猛. 基于 TOF 相机的水蜜桃检测与抓取方法研究[D]: [硕士学位论文]. 无锡: 江南大学, 2023.
- [8] Fu, Y., Li, W., Li, G., Dong, Y., Wang, S., Zhang, Q., *et al.* (2024) Multi-Stage Tomato Fruit Recognition Method Based on Improved YOLOv8. *Frontiers in Plant Science*, **15**, Article 1447263. <https://doi.org/10.3389/fpls.2024.1447263>
- [9] Shi, Y., Qing, S., Zhao, L., Wang, F., Yuwen, X. and Qu, M. (2024) YOLO-Peach: A High-Performance Lightweight YOLOv8s-Based Model for Accurate Recognition and Enumeration of Peach Seedling Fruits. *Agronomy*, **14**, Article 1628. <https://doi.org/10.3390/agronomy14081628>
- [10] Xiao, X., Wang, Y., Zhou, B. and Jiang, Y. (2024) Flexible Hand Claw Picking Method for Citrus-Picking Robot Based on Target Fruit Recognition. *Agriculture*, **14**, Article 1227. <https://doi.org/10.3390/agriculture14081227>