

# 基于卷积神经网络的果蔬识别系统设计

郑欢欢 赵高瑞朗 骆开发

榆林学院 信息工程学院, 中国·陕西 榆林 719000

**摘要:** 构建一个基于卷积神经网络的果蔬智能识别系统可以大大推进农业的智能化建设。该系统利用 TensorFlow 为深度学习框架构建了 Mobile Net 深度学习模型, 主要用于图片识别功能。在视频检测方面, 采用 YOLOv5 模型进行目标检测和分类, 并结合 PyQt5 图形化界面实现用户友好的交互体验。为了在具体应用中达到精确度和召回率之间的最佳平衡, 系统在 YOLOv5 模型调整 IoU 阈值, 并修改了损失函数算法以进一步改善模型性能。此外, 系统采用了简洁的条件判断和赋值结构, 避免了代码中的冗余和复杂性, 从而在保证准确性的同时降低计算复杂度, 减少 CPU 和 GPU 的占用率。实验结果表明, 改进后的系统在识别准确率和效率方面的表现优异, 有效提高了果蔬识别的精度和速度, 促进了农业智能化的发展。

**关键词:** 果蔬识别; YOLOv5; MobileNet; IoU 阈值; 损失函数处理

## Design of Fruit and Vegetable Recognition System Based on Convolutional Neural Network

Huanhuan Zheng Ruilang Zhaogao Kaifa Luo

School of Information Engineering, Yulin University, Yulin, Shaanxi, 719000, China

**Abstract:** Building an intelligent recognition system for fruits and vegetables based on convolutional neural networks can greatly promote the intelligent construction of agriculture. The system utilizes TensorFlow as a deep learning framework to construct a MobileNet deep learning model, mainly used for image recognition functions. In terms of video detection, the YOLOv5 model is used for object detection and classification, and combined with the PyQt5 graphical interface to achieve a user-friendly interactive experience. In order to achieve the best balance between accuracy and recall in specific applications, the system adjusted the IoU threshold based on F1 score in the YOLOv5 model and modified the loss function algorithm to further improve model performance. In addition, the system adopts a concise conditional judgment and assignment structure, avoiding redundancy and complexity in the code, thereby reducing computational complexity and CPU and GPU usage while ensuring accuracy. The experimental results show that the improved system performs excellently in recognition accuracy and efficiency, effectively improving the accuracy and speed of fruit and vegetable recognition, and promoting the development of agricultural intelligence.

**Keywords:** fruit and vegetable recognition; YOLOv5; MobileNet; IoU threshold; loss function processing

## 0 前言

随着人工智能的飞速发展, 果蔬产业也随之发展。在这一背景下, 果蔬产业作为农业和食品领域的重要组成部分, 对于满足日益增长的食品需求和提升农业生产效率具有至关重要的作用。当前, 果蔬产业在生产、加工和销售过程中需要快速而准确的果蔬分类和识别技术, 以提高生产效率、优化资源利用, 同时确保产品质量和安全。在这一背景下, 本研究旨在开发一种智能化的果蔬识别系统, 通过融合先进的计算机视觉和机器学习技术, 以更快速、准确地进行果蔬分类和识别。果蔬智能识别系统的研发对农业和食品生产领域具有重要的实际意义, 该系统将有助于解决传统人工分类方式存在的劳动力成本高和效率低的问题, 通过引入先进的图像识别技术, 提高果蔬分类的自动化水平。此外, 果蔬智能识别系统在果蔬分拣中由于人工和机械分拣往往成

本高, 灵活性低, 智能化程度差。而通过果蔬智能识别系统能够快速准确地识别出不同种类、大小、成熟度的果蔬, 极大地提高了分拣效率、减少了人为错误, 并且能够支持更精细化的分类和管理。

## 1 果蔬识别系统设计

### 1.1 系统结构设计

通过图 1 的果蔬识别系统总体结构图可以看出, 所设计系统的图像采集模块利用摄像机以及视频和图片采集果蔬图像<sup>[1]</sup>。标记处理现场图像后, 利用通信网络将其传送至图像预处理模块。图像预处理模块对图像进行清洗和标准化处理, 再将预处理后果蔬图像传送至行为识别模块。果蔬识别模块利用卷积神经网络识别果蔬种类, 输出果蔬识别结果。可视化界面展示果蔬识别结果, 为农业智能化提供基础。

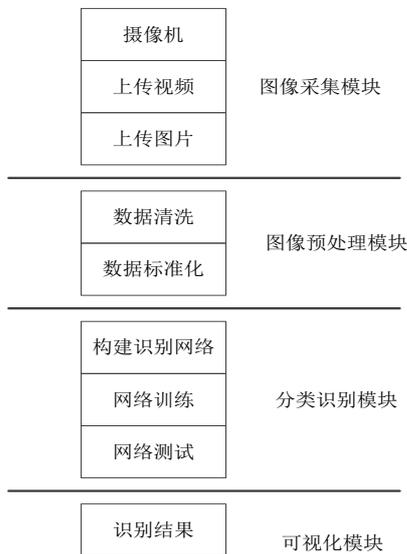


图 1 果蔬识别系统结构设计

### 1.2 基于 Mobile Net 果蔬识别

MobileNet 是一种专为移动和嵌入式设备设计的轻量级卷积神经网络模型，其核心设计原理在于通过深度可分离卷积 (Depthwise Separable Convolutions) 来减少模型的计算量和参数数量，从而提高模型的运行效率。深度可分离卷积将传统的标准卷积分解为两个独立的步骤：深度卷积 (Depthwise Convolution) 和逐点卷积 (Pointwise Convolution)。深度卷积对每个输入通道分别进行卷积操作，而逐点卷积则使用 1x1 的卷积核来组合深度卷积的输出。这种分解方式显著降低了计算复杂度，使得 MobileNet 能够在保持较高识别准确率的同时，实现较低的延迟和能耗。

在果蔬智能识别系统中，MobileNet 模型被用作图片识别功能的核心组件。通过 TensorFlow 深度学习框架，MobileNet 能够高效地提取果蔬图片的特征，并进行准确的分类识别。由于 MobileNet 的轻量级特性，它非常适合在资源受限的农业环境中部署，如智能农业机器人、便携式果蔬检测设备，为农业生产提供了强有力的技术支持。

### 1.3 基于 YOLOv5 视频检测和分类

YOLOv5 是基于卷积神经网络 (Convolutional Neural Network, CNN) 的目标检测模型。卷积神经网络具有强大的特征提取能力。YOLOv5 架构包括多个卷积层、池化层、批量归一化层和激活函数等组件，用于提取图像特征并进行目标检测。模型主要部分有 Backbone (主干网络)、Neck (颈部网络) 和 Head (头部网络)，都是基于卷积神经网络设计的。通过卷积神经网络的层层堆叠和信息传递，YOLOv5 能够高效地从输入图像中提取特征，并将这些特征用于目标检测任务。

## 2 系统评估与改进

### 2.1 损失函数

YOLOv5 模型采用 BoundinGIoU 损失函数计算损失，

gBox 损失则采用了 GIoU (Generalized Intersection over Union) 解决了 IoU (Intersection over Union) 检测框和真实框没有重叠时 Loss 等于 0 的问题。GIoU 在 IoU 的基础上考虑了最小包围矩形的面积，但是当检测框和真实框出现包含的时候 GIoU 退化成 IoU，当两个框相交时，在水平和垂直方向上收敛变慢。DIoU 要比 GIoU 更加符合目标框回归的机制，增加了中心点距离的考虑，使回归机制变得更加稳定，不会像 IoU 和 GIoU 一样在训练过程中出现发散的问题。CIoU (Complete-Intersection over Union) 相比 DIoU 充分考虑了预测框的长宽比和目标框之间的长宽比的一致性，从而提升预测框接近真实框的概率，所以将 YOLOv5 中计算 Bounding box 的损失函数由 GIoU 改进为 CIoU<sup>[2]</sup>，在使用预测框标注图像时可以使预测框更加接近真实框的位置。根据这些理论原理，对损失函数算法进行改进。

### 2.2 损失函数处理

在研究 YOLOv5 模型损失函数处理中可发现其调用了 ComputeLoss 类来初始化损失函数。而其中对调用了 bbox\_iou 方法，在训练过程中，bbox\_iou 用于计算预测框和真实框之间的 IoU，从而计算损失函数。损失函数会反馈给模型进行权重更新，以提升模型的预测精度，推理阶段，bbox\_iou 用于非极大值抑制 (NMS)，以移除多余的重叠框，最终确定检测结果。针对此对 bbox\_iou 进行改进。根据果蔬识别系统的应用需求分析，可以减少函数调用和条件判断，直接使用张量运算计算 IoU 和 CIoU。并且根据需要设计 CIoU 损失函数计算方法。在一定程度上提高性能。计算 CIoU (Complete IoU) 损失函数公式如下：

对于 iou 的计算采用相交面积与并集面积的比值计算，其中相交面积公式为：

$$width = \max(0, x_2 - x_1 + 1) \tag{1}$$

$$height = \max(0, y_2 - y_1 + 1) \tag{2}$$

$$inter = width \times height \tag{3}$$

其中 x1, x2, y1, y2 分别是相交区域四个边角的坐标，因为边界框的坐标可能是闭区间表示，也就是包括边界点本身，为了确保计算准确，对坐标上的差值加 1。

$$iou = inter / (b_1 + b_2 - inter + \epsilon) \tag{4}$$

式中，inter 表示相交区域面积，b1 表示第一个边界框面积，b2 表示第二个边界框面积，ε 一个很小的正数，用于防止除以零的情况，确保计算的稳定性。

CIoU 在 iou 的基础上添加了中心点距离计算以及长宽比的考虑。其中，包围框面积计算如下：

$$w_{enclose} = \max(0, x^{(2)} - x^{(1)} + 1) \tag{5}$$

$$h_{enclose} = \max(0, y^{(2)} - y^{(1)} + 1) \tag{6}$$

$$Area_{enclose} = w_{enclose} \times h_{enclose} \tag{7}$$

式中,  $x^{(1)}$ 、 $x^{(2)}$ 、 $y^{(1)}$ 、 $y^{(2)}$  分别是包围矩形的四个直角坐标,  $w_{enclose}$ 、 $h_{enclose}$  分别是包围矩形的宽度和高度为了包含边界上的像素, 在计算高度和宽度时我们加 1。

接下来需要使用中心点距离以及包围框面积来计算中心点距离项。其中, 中心点距离计算公式如下:

$$cx = \frac{b1\_x1 + b1\_x2}{2} \quad (8)$$

$$cy = \frac{b1\_y1 + b1\_y2}{2} \quad (9)$$

$$cxp = \frac{b2\_x1 + b2\_x2}{2} \quad (10)$$

$$cyp = \frac{b2\_y1 + b2\_y2}{2} \quad (11)$$

式中,  $cx$  和  $cy$  分别表示  $box1$  的中心点的  $x$  坐标和  $y$  坐标,  $cxp$  和  $cyp$  分别表示  $box2$  的中心点的  $x$  坐标和  $y$  坐标。  $b1\_x1, b1\_y1, b1\_x2, b1\_y2$  分别是  $box1$  的左上角和右下角坐标,  $b2\_x1, b2\_y1, b2\_x2, b2\_y2$  分别是  $box2$  的左上角和右下角坐标。

有了中心点距离可以推导出中心点距离项公式如下:

$$\rho^2 = \frac{(cx - cxp)^2 + (cy - cyp)^2}{Area_{enclose}} \quad (12)$$

式中,  $\rho^2$  衡量了两个边界框中心点之间的欧氏距离的平方, 通过归一化到最小闭合矩形的面积来进行调整, 以便在计算  $CIoU$  时使用。

由于边界框的形状差异, 简单的  $IoU$  只考虑了两个边界框的重叠面积, 而没有考虑边界框形状的相似性。可以引入宽高比例项,  $CIoU$  可以更好地捕捉到边界框之间的几何差异, 从而提升模型的精度, 公式如下:

$$v = \frac{4}{\pi^2} (\tan^{-1}(\frac{w_{enclose}}{h_{enclose}}) - \tan^{-1}(\frac{b2\_x2 - b2\_x1}{b2\_y2 - b2\_y1}))^2 \quad (13)$$

式中,  $w_{enclose}$  和  $h_{enclose}$  分别是两个框的最小和最大宽度和高度。其中, 两个  $\tan^{-1}$  所计算的分别是两个框的宽高比。

通过引入权重因子  $\alpha$  来平衡  $IoU$  和宽高比项  $v$  的影响, 可以使得  $CIoU$  更加准确地反映两个边界框的实际差异。  $\alpha$  值如下所示:

$$\alpha = \frac{v}{1 - iou + v + \epsilon} \quad (14)$$

式中, 确保了当  $IoU$  较高时,  $\alpha$  较大,  $v$  的影响较大; 当  $IoU$  较低时,  $\alpha$  较小,  $v$  的影响较小。

最终将中心点距离项  $\rho^2$ , 宽高比项  $v$ 、权重因子  $\alpha$  综合起来来计算出最终的  $CIoU$  值。如下所示:

$$CIoU = IoU - (\rho^2 + \alpha \cdot v) \quad (15)$$

式中, 使得  $CIoU$  能够同时考虑边界框的重叠面积、中心点距离和宽高比差异, 从而提供更全面和准确的边界框相

似性度量。

### 3 实验与优化

#### 3.1 实验参数设置

为了测试相同条件下不同模型的性能, 在模型训练过程中均使用相同的数据集和参数设置, 并对训练结果进行对比, 部分训练参数设置如表 1 所示。

表 1 参数设置

参数名	参数值
Momentum	0.937
Batch_size	4
Iou_thres	0.2
Weight_decay	0.0005
Mosaic	1.0
Copy_paste	0.0
Mixup	0.0
Epoch	100

#### 3.2 评价指标

论文基于图 1 构建了基于神经网络的果蔬识别系统, 并进行神经网络训练, 实验步骤如下: 第一, 数据准备, 即通过准备和下载数据集, 并对图像进行预处理, 包括尺寸调整、归一化等; 第二, 卷积神经网络设计, 即基于 YOLOv5s 模型; 第三, 模型训练, 通过修改 YOLOv5 中对于  $IoU$  阈值调整损失函数优化来对模型进行评估。其中模型评估标准使用平均精度均值  $mAP$  (mean Average Precision)、准确率  $P$  (Precision)、召回率  $R$  (Recall)、 $F1$  值 ( $F1$ -score) 等评价指标对模型进行评价<sup>[3]</sup>, 计算公式如上述所示。

#### 3.3 实验结果与分析

基于 YOLOv5 改进了损失函数算法, 在训练时将改进后的损失函数计算模型在手动标注的果蔬图像数据集上进行 100 轮训练, 训练获得  $mAP_{0.5}$  达到 92%,  $mAP_{0.5:0.95}$  达到 57%, 准确率 Precision 达到 81%, 召回率 Recall 达到 93%, 如图 2 所示。

研究对 YOLOv5 模型损失函数进行了改进。提高了 YOLOv5s 模型的性能, 如图 3 所示。

在模型训练完成以后对果蔬识别系统进行构建, 论文构建的图像采集界面, 包括上传本地图片, 视频识别, 摄像头检测按钮, 在上传完成图片或视频检测时数据会发送到服务器端进行识别并且返回结果, 结果会在相框中显示出来。在进行测试时, 上传数据后的识别结果如图 4 所示。可以看出, 该系统可以准确判断识别出数据, 验证了该方法的可行性和有效性。

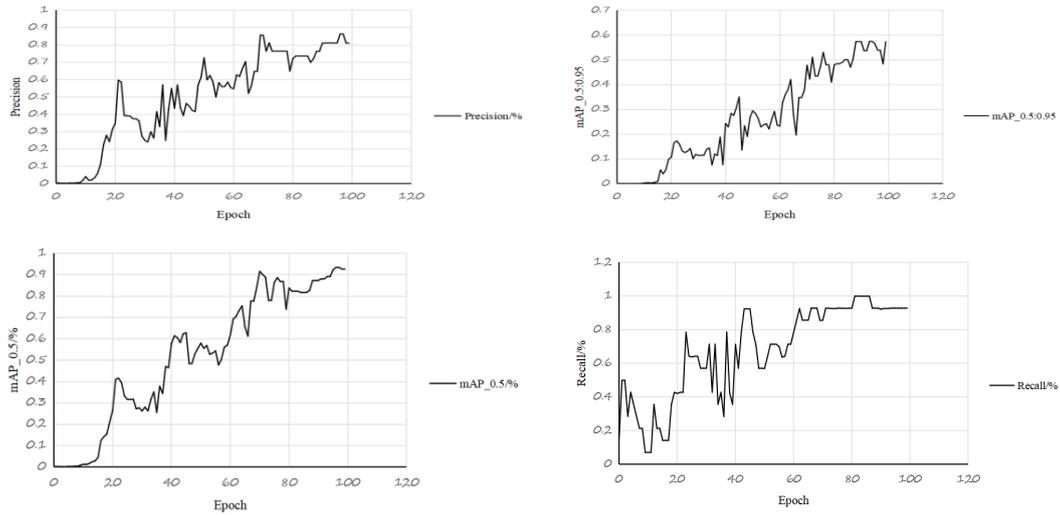


图 2 系统训练性能

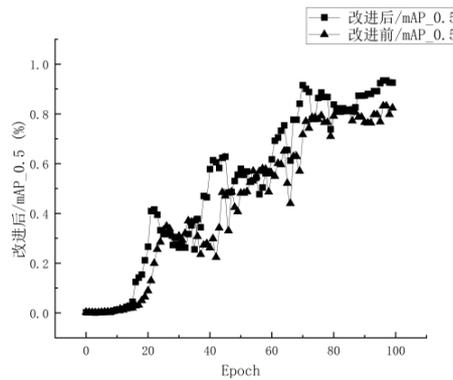


图 3 改进后性能

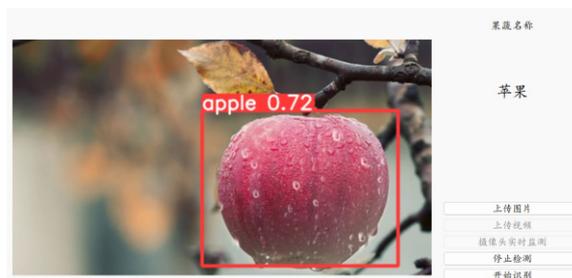


图 4 系统检测效果

### 4 结语

本研究提出了一种优化的 CIoU 计算方法。与原方法相比，我们的改进版本专注于简化代码结构并提高计算效率，去除了对 GIoU 和 DIoU 的支持，仅保留 CIoU 计算。具体改进包括在计算中心点距离和长宽比项时引入更高效的计算公式。实验结果表明，改进后的 CIoU 方法在保持精度的同时，显著提高了计算速度，为实时目标检测应用提供了更加高效的解决方案。

### 参考文献：

[1] 刘琳.基于卷积神经网络的学生课堂行为识别系统设计[J].现代

电子技术,2024,47(6):142-146.

[2] 马文宝,田芳明,谭峰.基于改进YOLOv5模型的农作物病斑图像自动标注[J].黑龙江八一农垦大学学报,2024,36(1):108-116+130.  
 [3] 覃琼花,张春燕,徐百宁.基于深度学习的微信小程序图像识别系统[J].信息与电脑(理论版),2024,36(1):73-75.

作者简介: 郑欢欢(1986-),女,中国陕西榆林人,硕士,副教授,从事物联网技术和图像处理研究。

基金项目: 榆林市科技计划项目(项目编号: 2022CXY-12-6)。