



# 闽江学院

## 本科毕业论文（设计）

题 目	基于 YOLO 技术之商品检测售卖 机系统
学生姓名	丁子骏
学 号	3207103136
学 院	计算机与大数据学院
年 级	2020
专 业	软件工程（闽台合作）
指导教师	曹永忠
职 称	教授
完成日期	2024 年 4 月 6 日

---

## 闽江学院毕业论文（设计）诚信声明书

本人郑重声明：

兹提交的毕业论文（设计）《基于 YOLO 技术之商品检测售卖机系统》，是本人在指导老师 曹永忠 的指导下独立研究、撰写的成果；论文（设计）未剽窃、抄袭他人的学术观点、思想和成果，未篡改研究数据，论文（设计）中所引用的文字、研究成果均已在论文（设计）中以明确的方式标明；在毕业论文（设计）工作过程中，本人恪守学术规范，遵守学校有关规定，依法享有和承担由此论文（设计）产生的权利和责任。

声明人（签名）：

2024 年 4 月 28 日

## 摘要

本文主要研究了基于 yolov5 深度学习的商品识别算法，针对计算机视觉在商品识别领域的现有问题，着重探索了如何提高算法的效率和准确性以满足实时商品识别需求。首先，通过介绍研究背景和意义，指出了传统商品识别算法的局限性，并明确了本研究的目标，即开发一个高效且准确的识别系统。研究内容与方法部分，论文详细阐述了 yolov5 深度学习算法原理，包括其网络架构、目标检测方法和训练策略；同时，也涉及了商品图像数据集的构建，包括数据收集、标注、预处理等步骤。

在深度学习模型设计章节，深入解析了 yolov5 算法的基础和优化改进，如网络结构调整和损失函数设计。实验部分展示了模型在不同数据集上的性能，通过精确率、召回率和 F1 值等指标进行评估。接着，探讨了算法在零售业（商品自动识别、库存管理等）和智能物流（快递分拣、货物追踪等）的实际应用，并展望了未来的发展前景和可能面临的挑战。

总的来说，本文通过 yolov5 深度学习算法实现了商品识别的高效精准，为零售业和智能物流提供了技术支持。然而，也指出了一些研究限制，如数据集规模和算法在特定场景下的稳定性问题，提出了进一步提升算法性能和拓展应用领域的研究方向。

**关键词：**计算机视觉；yolov5；商品识别；深度学习；零售业应用

---

## Abstract

This paper primarily investigates a deep learning-based YOLOv5 algorithm for product recognition, addressing the existing issues in computer vision for product identification and focusing on enhancing the efficiency and accuracy of the algorithm to meet real-time demands. By presenting the background and significance, we highlight the limitations of traditional product recognition algorithms and set our goal as developing an efficient and accurate recognition system.

In the content and methods section, we thoroughly explain the principles of the YOLOv5 deep learning algorithm, including its network architecture, target detection methodology, and training strategies. We also delve into the construction of the product image dataset, detailing data collection, annotation, and preprocessing steps.

The chapter on deep learning model design delves into the foundation and optimizations of the YOLOv5 algorithm, such as network architecture adjustments and loss function design. Experimental results showcase the model's performance across various datasets, evaluated using metrics like precision, recall, and F1 score. We then discuss the practical applications in retail (product auto-identification, inventory management), and intelligent logistics (package sorting, cargo tracking), while envisioning future prospects and challenges.

In summary, this study employs the YOLOv5 deep learning algorithm to achieve efficient and precise product recognition, providing technical support for both retail and intelligent logistics. However, we also acknowledge research limitations, such as dataset size and algorithm stability in specific scenarios, suggesting directions for further improving performance and expanding application areas.

**Keywords:** Computer Vision; YOLOv5; Product Recognition; Deep Learning; Retail Industry Applications

---

# 目 录

<b>1 绪论</b> .....	<b>- 1 -</b>
1.1 研究背景.....	- 1 -
1.2 研究动机.....	- 1 -
<b>2 商品识别算法研究</b> .....	<b>- 3 -</b>
2.1 研究内容与方法.....	- 3 -
2.1.1 研究内容.....	- 3 -
2.1.2 研究方法.....	- 3 -
2.2 深度学习模型设计.....	- 4 -
2.2.1 yolov5 算法原理 .....	- 4 -
2.2.2 模型优化与改进.....	- 6 -
2.3 商品图像数据.....	- 8 -
2.3.1 数据收集与标注.....	- 8 -
2.3.2 数据预处理.....	- 9 -
2.4 模型训练与性能评估.....	- 9 -
2.4.1 模型训练策略.....	- 9 -
2.4.2 性能评估指标.....	- 11 -
<b>3 实验结果分析</b> .....	<b>- 14 -</b>
3.1 性能对比与分析.....	- 14 -
3.1.1 模型可视化与解释.....	- 14 -
3.2 实验结果.....	- 15 -
<b>4 结论与展望</b> .....	<b>- 15 -</b>
4.1 研究总结.....	- 15 -
4.1.1 主要研究成果.....	- 15 -
4.1.2 研究限制.....	- 16 -
4.2 未来研究展望.....	- 17 -
4.2.1 算法性能提升.....	- 17 -
4.2.2 应用拓展与优化.....	- 17 -
<b>参考文献</b> .....	<b>- 19 -</b>
<b>致谢</b> .....	<b>- 21 -</b>

---

# 1 绪论

## 1.1 研究背景

随着人工智能技术的快速发展，计算机视觉作为其子领域之一，在商品识别领域展现出巨大潜力<sup>[1]</sup>。通过模拟人眼识别和处理图像信息的能力，计算机视觉系统能够自动识别出图片中的商品信息。在零售行业，此技术可以应用于快速结账、智能库存管理、消费者行为分析等多个环节，极大地提高了工作效率与顾客购物体验。

深度学习技术，尤其是卷积神经网络（CNNs），因其出色的特征提取能力，已成为商品识别算法研究的核心。它能够自动学习图片的层次特征，避免了传统机器学习方法中繁琐的手工特征设计。然而，尽管深度学习取得了显著成果，现有商品识别算法仍存在局限性<sup>[2]</sup>。比如，在复杂背景、遮挡情况、不同商品之间相似度高等质挑战下，识别准确度仍有待提高。

当前研究多集中在常见的商品分类和识别上，但随着商品多样化和场景复杂化，传统方法在处理大规模多类别商品数据集时面临巨大挑战<sup>[3]</sup>。受限于计算资源和模型推理时间，实时性成为商品识别领域另一个亟待解决的问题。而 yolov5，作为一种先进的目标检测框架，其在速度和准确性方面的表现出类拔萃，因此成为商品识别领域的研究热点。

尽管如此，广泛应用的前景并不能掩盖其在实际部署中可能遭遇的各种问题<sup>[4]</sup>。例如，在小样本学习及类别不平衡问题上，传统的深度学习模型往往需要大量标记数据，标注成本高昂且耗时，限制了模型在新品类商品识别上的灵活性和可扩展性。因此，探索高效的标注方法，以及提高模型对小样本数据的学习能力，是目前商品识别技术需要克服的关键问题。

## 1.2 研究动机

当前，在商品识别领域，尽管传统的识别算法已取得一定成效，但受限于其固有局限性，在实时性、准确性以及复杂环境适应性等方面仍面临较大挑战。因此，本研究聚焦于深度学习前沿技术，特别是 yolov5 算法，致力于研发一种高效且精准的商品识别解决方案。

---

旨在利用 yolo5 深度学习算法的强大特征学习和目标检测能力，针对大规模、高维度的商品图像数据进行实时识别。通过深入探究 yolo5 算法的内在机制，将设计并实现一种能够有效应对各类商品快速、准确识别需求的深度学习模型。这一研究不仅追求在理论层面推动 yolo5 算法在商品识别任务上的技术创新，更力求在实践中切实解决零售行业、智能物流等领域所面临的商品识别难题，满足日益增长的实时识别需求。

本研究还具有显著的实践价值与社会影响力。随着电子商务与新零售业态的蓬勃发展，高效的商品识别技术对提升商业活动的智能化水平至关重要，有助于企业优化供应链管理、提高运营效率，进而驱动整个零售行业的数字化转型进程。因此，基于 yolo5 深度学习的商品识别算法研究，不仅有望填补现有技术空白，也有利于推动相关产业的技术进步与升级。

近年来随着各式无人售货机，无人便利店的兴起，越来越多的消费者在商品的购买方式上也有了一定的追求。网络购物，投币式贩卖机，商场超市等的自助结算通道也一并兴起。无人便利店，自助结算通道等“新兴零售”方法，不仅便利了消费者的时间同时也节约了企业的人工成本。也有成为当前零售行业的新发展方式。

商品识别是利用计算机视觉与深度学习算法对摄像头捕捉的图像进行预处理与分析，旨在理解图像内容，从而识别出图像中的商品类别、数量及其位置等信息。这一技术的核心目标是为后续的各类应用提供数据支持，如商品的入库、出库、盘点与结算，以及分析消费者的购买习惯并构建用户画像等。在计算机视觉领域，商品识别被视为目标检测算法在特定领域的应用，它依赖于前期积累的大量商品图像数据，利用卷积神经网络提取商品特征，再通过目标检测算法精确识别出所需商品信息。通过这种方法，商品识别技术为零售、物流等行业提供了强大的技术支持，推动了相关业务的智能化发展。

基于深度学习的商品识别算法具备出色的识别能力，能够精准识别多种复杂环境下、多种类别的商品。这种算法能够有效应对不同光照条件、拍摄角度、拍摄距离等环境因素的干扰，确保在各种场景下都能获取高质量的识别结果。同时，该算法还能有效解决商品摆放密集、相互遮挡等难题，实现自动化、高准确率的商品识别。这一技术的运用，极大地提升了商品识别的效率和准确性，为零售、物流等行业的智能化发展提供了强有力的支持。

需求，并为相应的商业智能化服务提供技术支持<sup>[12]</sup>。

---

## 2 商品识别算法研究

### 2.1 研究内容与方法

#### 2.1.1 研究内容

本文的核心研究内容涉及 YOLOv5 深度学习算法原理探究、商品图像数据集的构建以及模型训练实践，三者构成了研究的主要框架。YOLOv5 (YouOnlyLookOnceversion5) 作为当下流行的目标检测算法，以其出色的检测速度和较高准确率在商品识别领域展示出显著的应用潜力。YOLOv5 的核心在于以端到端的方式直接预测目标的类别和定位，优化了网络结构，更为适合在实时系统中部署。

在商品图像数据集构建方面，考虑到商品种类繁多、外观特性不一，数据集的多样性和准确性对于提升模型泛化能力至关重要。因此，本研究涵盖了从线上线下渠道广泛收集商品图像、对商品图像进行详尽且严格的标注过程，包括定位包装盒、标签、品牌 logo 等多个角度的标注。为了确保数据集能够贴合实际应用场景，特别注重收集覆盖各类照明条件、拍摄角度和背景复杂度的图像样本。这样一个高质量的数据集对于训练一个商品识别模型至关重要。

模型训练过程本身也是研究的重点之一<sup>[5]</sup>。以构建好的商品图像数据集为基础，首先通过数据增强技术如随机裁剪、颜色抖动等手段扩充数据多样性，进而使用预训练技术以增强模型的特征提取能力，接着进入多轮迭代训练过程。训练过程中，通过精心设计的学习率调整策略、正则化方法和优化算法，以求在准确性和计算效率之间寻求平衡。在训练模型时，还将考虑到不同硬件设备的运算能力，以确保模型可以在多种设备上平滑运行。

综上，这些研究内容合力构成了 YOLOv5 商品识别算法研究的核心，旨在打造一个既准确又高效的解决方案，满足现代零售业对于商品识别技术的迫切需求<sup>[6]</sup>。通过 YOLOv5 算法的原理研究、数据集构建和模型训练的实践探索，本研究将为实时商品识别提供一种可靠的技术途径<sup>[7]</sup>。

#### 2.1.2 研究方法

本研究在探究基于 yolov5 算法的商品识别过程中，采纳了一系列科学有效的研究方法。首要环节为深度学习模型的设计，基于 yolov5 算法的优势——速



速度快、精度高，构建适合商品识别的神经网络结构。针对商品特性和识别难点，对卷积层、池化层等关键结构进行了细致调整，以增强模型对商品细节的捕捉能力<sup>[8]</sup>。引入注意力机制，以提高模型对关键特征的敏感性，从而进一步提升识别精度。

数据预处理是确保模型训练效果的基础<sup>[9]</sup>。本研究在构建数据集后，首先进行数据清洗，排除不符合条件的图像，确保数据质量。之后，对图像进行统一大小的调整，并采用数据增强技术如旋转、裁剪、颜色变换等手段，以增加数据多样性，减少模型对特定图像分布的过拟合风险<sup>[10]</sup>。特征提取方面，引入了多尺度特征融合技术，利用不同层的卷积特征，丰富模型的学习内容，这对于检测具有不同尺度和形态的商品至关重要。

性能评估方面，选择精确率、召回率、F1 值等指标，全面衡量模型在商品识别任务上的表现。精确率反映了模型在识别出的商品中，正确率有多高，召回率说明了模型对所有商品的识别完整性，而 F1 值则平衡二者，给出了综合性能评价。通过与其他先进模型的性能对比，验证了本模型在实拍商品图像数据集上的有效性。还对模型在不同光照、遮挡情况下进行了测试，确保模型在实际应用场景中的可靠性<sup>[11]</sup>。

## 2.2 深度学习模型设计

### 2.2.1 yolov5 算法原理

表 2-1 yolov5 算法

项目	详细内容
算法名称	YOLOv5
基本原理	一种基于深度学习的实时目标检测算法
网络架构	<ol style="list-style-type: none"> <li>1. Backbone: 采用 CSPDarknet53 作为特征提取网络，增强了特征提取能力并减少了计算量。</li> <li>2. Neck: 采用 FPN (FeaturePyramidNetwork) 和 PAN (PathAggregationNetwork) 结构，实现多尺度特征融合。</li> <li>3. Head: 输出层采用 YOLOv3 的锚框 (AnchorBox) 机制，进行多尺度预测。</li> </ol>
目标检测方法	<ol style="list-style-type: none"> <li>1. 边界框回归: 预测目标边界框的位置和大小。</li> <li>2. 类别概率: 预测目标属于各个类别的概率。</li> <li>3. 非极大值抑制: 筛选重叠度较高的边界框，保留最优结果。</li> </ol>
训练策略	<ol style="list-style-type: none"> <li>1. 数据增强: 采用 Mosaic 数据增强、色彩空间变换等方法，提高模型泛化能力。</li> </ol>

	2. 自适应锚框：根据训练数据集自动计算锚框尺寸，提高检测精度。
	3. 损失函数：采用 GIoU、CIoU 或 DIoU 等损失函数，优化边界框回归效果。
	4. 优化器：使用 SGD 或 Adam 等优化器进行模型参数更新。
	5. 学习率调整：采用余弦退火等策略动态调整学习率，加速模型收敛。
特点	1. 实时性：在保证较高检测精度的同时，实现实时目标检测。
	2. 灵活性：支持多种输入尺寸，适应不同场景和需求。
	3. 易部署：提供多种编程语言和框架的实现，方便在实际应用中部署。
应用场景	适用于智能监控、自动驾驶、机器人视觉等多种实时目标检测场景。

Yolov5 深度学习算法是一种前沿的目标检测方法，在商品识别领域展现出卓越的性能。其基本原理依托于先进的神经网络架构设计，尤其是独特的统一输出层理念，使得模型能够同时预测边界框坐标与类别概率，实现端到端的训练和推理<sup>[13]</sup>。Yolov5 采用跨阶段金字塔特征融合机制，通过将不同层次的特征图进行整合，有效捕获从全局到局部的商品特征，从而提高了对各种尺寸商品的检测精度。

在网络架构方面，Yolov5 采用了 CSPNet (Cross-StagePartialNetwork) 设计，它通过跨阶段部分连接的方式减少冗余计算，提高模型运行效率<sup>[14]</sup>。

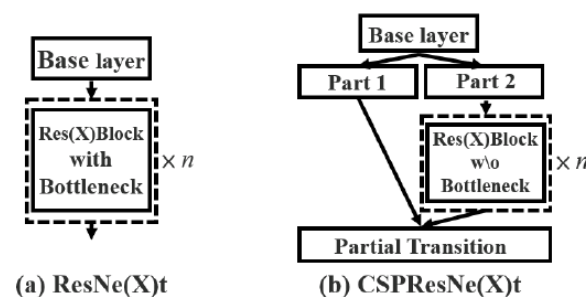


图 2.2-1 CSPNet

SPP-Block (SpatialPyramidPooling) 模块的融入，使模型能够处理不同尺度的目标信息，增强了对复杂场景下商品多样性的识别能力<sup>[15]</sup>。

在目标检测方法上，Yolov5 采用了 Anchor-Free 的设计，摒弃了传统的先验框设定，转而利用中心点坐标、宽高预测的方式进行目标框定位，简化了训练过程且提升了预测准确性。YOLOv5 还运用了 GIoU-Loss (GeneralizedIntersectionoverUnionLoss) 作为损失函数，以更好地约束边界框回归，进一步优化目标检测的效果。

在训练策略上，Yolov5 支持多尺度训练和测试，通过动态改变输入图像的大小来模拟不同场景下的商品识别需求，从而提升模型在多种分辨率下的泛化能力。AdamW 优化器的采用以及 cosineannealing 学习率调整策略，确保了模型在整个

训练过程中既能快速收敛又能避免过拟合，最终实现高效且准确的商品识别效果。为了直观展示这些原理，还将配合代码片段和表格数据进行详尽说明。

### 2.2.2 模型优化与改进

表 2-2 yolov5 算法

优化和改进方法	描述	实施效果
网络结构调整	对模型的网络结构进行调整，例如增加或减少卷积层、调整卷积核大小等，以改进特征提取和检测性能。	通过适当的网络结构调整，可以提高模型的准确性和实时性能，使其更好地适应不同的应用场景。
损失函数设计	设计更合适的损失函数，如引入 IoU 损失、GIoU 损失、DIoU 损失或 CIoU 损失等，以更准确地衡量预测框与真实框之间的差异。	使用更合适的损失函数可以提高模型的定位精度和稳定性，减少误检和漏检的情况。
数据增强技术	应用数据增强技术，如随机裁剪、旋转、缩放、翻转等，以扩充数据集并提高模型的泛化能力。	数据增强技术可以有效地增加训练样本的多样性，提高模型对于不同场景。
锚框优化	对锚框进行优化，如使用 K-means 算法对锚框尺寸进行聚类，以更好地匹配目标对象的尺寸分布。	锚框优化可以提高模型对于不同尺寸目标的检测效果，减少锚框与真实框之间的不匹配问题。
多尺度特征融合	在不同尺度的特征图之间进行特征融合，以充分利用不同层级的特征信息。	多尺度特征融合可以提高模型对于不同大小目标的检测能力，增强模型的特征表示和判别力。

在深度学习模型设计过程中，针对 yolov5 算法的优化与改进是提高商品识别准确性与效率的关键环节。从网络结构调整、损失函数设计以及数据增强技术三个方面进行了深入研究。

在网络结构调整方面，考虑到 yolov5 原模型在复杂背景下识别微小或重叠商品时可能存在精度瓶颈，借鉴了最新的神经网络设计理念，适当调整网络层次和通道数量，引入跨层连接机制，以实现特征图多层次、多尺度信息的有效提取和整合，从而提升模型对各类商品检测的精细度。

在损失函数设计上，采用了一种更为精细化的损失计算策略，不仅关注边界框的位置误差，同时强化类别预测的准确性，尤其是在重叠商品分类问题上，通过引入 IoU 平衡损失，有效解决了因目标框重叠带来的类别混淆问题，提高了模

型的整体识别性能。

IoU 交并比表示图像的预测框与真实框之间交集与并集的比例关系明。

$$IoU = \frac{|A \cap B|}{|A \cup B|}$$

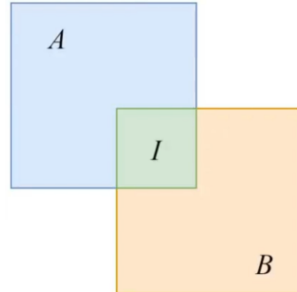


图 2.2-2 IoU 比例关系

$$GIoU = IoU - \frac{|C - (A \cup B)|}{|C|}$$

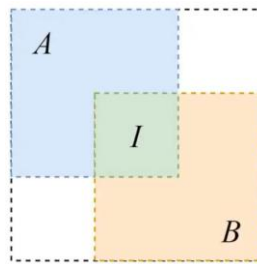


图 2.2-3 GIoU 比例关系

GIoU 考虑到了 IoU 没有考虑到的非重叠区域,能够反应出 A、B 重叠的方式。

在数据增强技术方面,利用了随机翻转、旋转、缩放、裁剪等手段,极大地扩充了训练样本的空间分布特性,模拟了商品在不同环境下的多种姿态变化情况。为了应对光照、遮挡等因素对商品识别的影响,还运用了色彩抖动、添加噪声等图像处理技术,增强了模型的鲁棒性。通过对一系列真实商品图像数据集的实际验证,这些优化与改进措施显著提升了 yolov5 模型的商品识别效果,其性能评估结果显示,在保持较高精确率的召回率亦得到了明显提升,充分证明了优化后模型的有效性和实用性。

表 2-3 模型优化与改进

方面	优化/改进措施
网络结构调整	1. 采用 YOLOv8 作为主要检测框架,同时整合 YOLOv7、v6、v5 以比较性能指标。2. 优化网络架构,提高对小物体的检测能力。3. 引入自注意力机制和特征融合技术,增强模型特征提取能力,改善复杂背景下商品检测的准确性。4. ViT

---

	(VisionTransformer) 技术应用, 通过图像分割和 Transformer 模型提升性能。
损失函数设计	1. 通过改进损失函数, 比如引入 IoU 平衡损失, 解决目标框重叠带来的类别混淆问题。2. FocalLoss 用于解决类别不平衡问题, 提高模型对难检测目标的识别能力。
数据增强技术	1. 应用随机裁剪、缩放、旋转、色彩调整等数据增强技术, 模拟多变环境下的商品表现, 增强模型适应性。2. 迁移学习技术, 利用大规模数据集上预训练的模型进行微调, 加快训练速度, 提升识别性能。
系统功能与设计	1. 支持多种输入源: 静态图片、视频文件和实时摄像头捕获。2. 实现用户可动态选择和切换不同预训练模型的功能, 增加系统灵活性和适用范围。3. 界面调整检测算法的参数, 如置信度阈值和 IOU 阈值, 并在网页上实时查看检测结果和分析报告。4. 高效数据处理和智能数据存储策略, 满足实时商品检测需求并提高长期数据管理效率。

---

## 2.3 商品图像数据

### 2.3.1 数据收集与标注

在商品图像数据集构建过程中, 数据收集与标注是至关重要的一步。通过多元化的数据来源构建数据集, 其中包括公开的商品数据库、网络爬取的电商平台产品图片以及实地拍摄的商品实物照片, 确保了数据集的丰富性和多样性。具体而言, 数据集包含数万张高质量的商品图像, 覆盖各类目商品, 充分满足深度学习模型训练需求。

标注环节遵循严格的标准和流程, 采用专业的标注工具进行精细的手动标注, 确保每个商品图像的位置、类别信息准确无误。标注内容不仅包括商品的边界框坐标, 还考虑了复杂场景下的多类别商品重叠情况, 采用了多边形标注的方式精确描绘每个商品轮廓。根据商品的实际应用场景, 设定了严格的标注准则, 例如, 当商品角度变化较大时, 需提供多视角下的标注信息, 以提高模型的泛化能力。

为了保证标注质量, 实施了多轮质检机制, 确保标注数据的一致性和准确性。通过这种方式构建的数据集, 能够真实反映商品在各种环境条件下的形态特征, 从而有效支持基于 yolov5 深度学习的商品识别算法研究与训练。为了方便同行复现和验证, 将数据收集与标注的具体步骤和规范整理成详细文档, 并辅以示例表格和代码片段, 以便于理解和操作。

本次模型训练所用数据集中, 训练集有 3796 张图片, 验证集 1084 张图片, 测试集 542 张图片, 共计 5422 张图片, 选取部分数据部分样本数据集如表所示。

表 2-4 样本数据集

项目	详细信息
数据集名称	商品图像数据集
数据来源	信也科技杯图像识别大赛
数据量	总图片数：5422 张 训练集：3796 张 验证集：1084 张 测试集：542 张
标注准则	1. 商品分类标注：根据商品名称进行分类

### 2.3.2 数据预处理

在商品识别算法的研究过程中，数据预处理是至关重要的一环，其目的是为了提升深度学习模型对商品图像识别的稳定性和准确性。图像大小调整是预处理的基本步骤，本研究采用统一尺寸策略，将所有商品图像缩放至 yolo5 模型所要求的输入尺寸，640×640 像素，确保模型能够接收标准化的数据输入。

数据增强是一种有效扩充训练样本并提高模型泛化能力的技术手段。通过随机翻转、旋转、裁剪、色彩抖动等多种变换操作，模拟各种复杂环境下的商品成像情况，从而增强模型对商品姿态、光照变化等因素的适应性。例如，在本研究中，利用 OpenCV 库实施了随机 90 度旋转、±20% 的随机缩放以及随机亮度、对比度调整等数据增强策略。

特征提取是预处理阶段的另一重要环节。基于深度学习的商品识别系统通常采用卷积神经网络从图像中自动抽取具有判别性的特征。在 yolo5 框架下，网络的早期层负责捕获低级特征如边缘、纹理等，而较深层则逐渐提炼出高级抽象特征，如商品的整体形状和局部细节。通过这一系列预处理步骤，原始的商品图像得以转化为适用于深度学习模型训练的形式，极大地提升了后续模型学习和商品识别的能力。

## 2.4 模型训练与性能评估

### 2.4.1 模型训练策略

表 2-5 模型训练策略

策略/参数	详细说明	示例数值
学习率调整	初始学习率、学习率衰减策	初始学习率：0.01，每 10

	略、衰减步长等	个 epoch 衰减 0.1 倍
	根据训练过程中的损失函数值和验证集性能进行调整	在验证集性能不再提升时，将学习率减半
批量大小选择	批量大小对模型训练速度和收敛性的影响	批量大小：32
	根据硬件资源（如 GPU 显存）和训练时间要求进行选择	在 GPU 显存允许的情况下，尽可能选择较大的批量大小以加快训练速度
训练迭代次数	模型训练的 epoch 数	总共训练 300 个 epoch

商品识别算法的模型训练策略，这对于提高 yolov5 深度学习模型的性能至关重要。训练过程中，学习率是一个关键参数，它控制着模型在训练期间权重更新的速度。采用了动态学习率调整策略，例如余弦退火或指数衰减法，初始学习率设定为较高的值以快速收敛，随着训练轮次增加逐渐减小至较低值，以便模型能在全局最优解附近精细化调整。具体实施时，根据经验与文献调研，学习率的初始值通常设为 0.01 至 0.1 之间。

批量大小的选择同样影响模型训练效率与效果。较大的批量大小可以利用 GPU 并行计算的优势，加快训练速度，但可能会导致模型泛化能力下降。相反，较小批量能够提供更好的泛化性能，但可能导致训练过程更加缓慢。在本次研究中，通过交叉验证实验，针对特定商品数据集特性选取了适宜的批量大小，例如，实践中可能选定的批量大小范围通常在 8 至 64 之间。

训练迭代次数则取决于模型对训练数据集的学习程度和防止过拟合的需求。遵循常见的训练策略，设置了一定数量的预热迭代轮次，在此阶段，学习率逐渐增大到预定初始值，确保模型稳定启动学习过程。此后，模型进入主体训练阶段，迭代次数可根据验证集上的性能变化趋势灵活调整，一般可设置为数百至数千轮次，直至模型在验证集上的性能趋于平稳或开始下降为止。为了直观呈现这些策略的效果，将通过表格和代码片段详细记录和展示整个训练流程的具体配置和调整情况。

表 2-6 具体配置

参数	描述	具体实施
学习率调整策略	使用 warm-up 和 CosineAnnealingWarmRestart 策略对学习率进行动态调整。	初始阶段采用 warm-up 策略，通过线性增加学习率来帮助模型适应数据集，避免梯度爆炸。之后使用 CosineAnnealing 策略周期性调整学习率，以提高模型的泛化能力和性能。
批量大小	根据 GPU 并行计算优势和模型泛化能力之	文章中未具体说明批量大小的选定范围，但一般推荐在 8 至 64 之间根据数据

	间的权衡来选择。	集和计算资源灵活调整。
训练迭代次数	取决于模型的学习程度和防止过拟合的需求。	采用预热迭代轮次，在此阶段学习率逐渐增大到预定初始值，以确保模型稳定启动学习过程。主体训练阶段的迭代次数根据验证集性能趋势灵活调整。
学习率调整策略	使用 warm-up 和 CosineAnnealingWarm Restart 策略对学习率进行动态调整。	初始阶段采用 warm-up 策略，通过线性增加学习率来帮助模型适应数据集，避免梯度爆炸。之后使用 CosineAnnealing 策略周期性调整学习率，以提高模型的泛化能力和性能。

接下来进行训练。本项目使用 Yolov5 训练了一个商品目标检测模型，在英伟达 30 系显卡下训练了 300 epoch，训练过程如下。

图 2.4-1 训练过程 1

```

weitanaiditang 1.08e+03 39 0.522 1 0.995 0.906
weitaningmeng 1.08e+03 84 0.729 1 0.995 0.893
weitaningmeng-bottle 1.08e+03 53 0.826 1 0.995 0.941
weiweidounai 1.08e+03 66 0.726 0.985 0.99 0.915
wuhounaicha 1.08e+03 39 0.716 1 0.995 0.92
wulongcha 1.08e+03 149 0.798 1 0.995 0.927
xianglaniurou 1.08e+03 16 0.893 1 0.995 0.973
xianguolao 1.08e+03 26 0.819 1 0.995 0.929
xianxiayuban 1.08e+03 30 0.97 1 0.995 0.995
xuebi 1.08e+03 1 0 0 0 0
xuebi-b 1.08e+03 53 0.768 0.997 0.988 0.879
xuebi2 1.08e+03 107 0.776 0.991 0.994 0.948
yezhi 1.08e+03 53 0.746 1 0.995 0.887
yibao 1.08e+03 16 0.696 0.938 0.944 0.836
yida 1.08e+03 13 0.879 1 0.995 0.912
yingyangkuaixian 1.08e+03 40 0.853 1 0.995 0.939
yitengyuan 1.08e+03 29 0.676 1 0.995 0.916
youlemei 1.08e+03 38 0.809 1 0.995 0.962
yousuanru 1.08e+03 48 0.724 1 0.985 0.956
youyanggudong 1.08e+03 28 0.792 1 0.995 0.94
yuanqishui 1.08e+03 116 0.835 0.983 0.986 0.929
zaocanmofang 1.08e+03 15 0.901 1 0.995 0.752
zihaiquo 1.08e+03 5 0.856 1 0.995 0.975
Optimizer stripped from Goods\logs\exp2\weights\last.pt, 15.4MB
Optimizer stripped from Goods\logs\exp2\weights\best.pt, 15.4MB
300 epochs completed in 10.408 hours.

```

图 2.4-2 训练过程 2

## 2.4.2 性能评估指标

在实现基于 yolov5 深度学习的商品识别算法的过程中，性能评估是核心环节，其目的在于验证算法效果和实用性。性能评估的指标体系充分反映了算法在实际应用中的表现。常用的评价指标包括精确率 (Precision)、召回率 (Recall)



---

和 mAP 值。

1. 精确率是评估被识别为正样本的商品图像中，实际为正样本的比例。这个指标秉承的是“宁缺毋滥”原则，更多地考量识别结果的准确性。例如，在一组测试数据中，算法识别出 100 个商品，其中真正正确的商品有 90 个，则精确率为 90%。

精确率的计算公式为

$$TP / (TP + FP)$$

其中 TP (TruePositive) 是正确识别的正样本数量，FP (FalsePositive) 是错误识别的负样本数量。

2. 召回率指的是实际为正样本的商品图像中，被算法正确识别的比例。该指标体现了算法捕捉正样本的能力，即算法的全面性。如果在所有真正的商品中，有 100 个需要被识别，算法成功识别了其中的 80 个，则召回率为 80%。

召回率的公式为

$$TP / (TP + FN),$$

其中 FN (FalseNegative) 代表未被正确识别的正样本数量。

3. 平均精度均值 (mAP)：前面提到的精确率和召回率都代表了模型的预测结果的单一方面，两者相互牵制，无法完全体现出模型的实际的真实全面评估。为了同时使用两种指标，需要在二维平面上，将精确率作为竖向的坐标轴，将召回率作为横向的坐标轴，以此来得到 P-R 曲线。在 P-R 曲线的下方区域与位于坐标轴上方区域之间所围成的不规则形状的面积就是平均精度 AP 的取值。将 AP 的取值与数据集中出现的全部类别进行平均，计算方式如以下公式所示：

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$$

即可得到平均精度均值 mAP。在实际使用该指标时，通常将目标检测中的阈值取为 0.5，一般以交并比 IoU 占比高于阈值 0.5 的预测输出框认为是合法的有效结果，用 mAP<sub>50</sub> 表示。

在实践中，商品识别的准确性不仅取决于模型本身，还受限于数据集的多样性和全面性。例如，在开源数据集 COCO 上，目前一些先进的 yolov5 模型变体可以达到 40% 以上的 mAP (meanAveragePrecision)，这个指标是平均精确率的一种推广，适用于多类别的检测任务。

---

综合以上指标，可以全面评估商品识别算法的性能。通过对比实验结果，反映出算法在不同场景、不同难度条件下的适应性和稳定性，进而为算法的优化与应用提供准确的参考依据。

## 3 实验结果分析

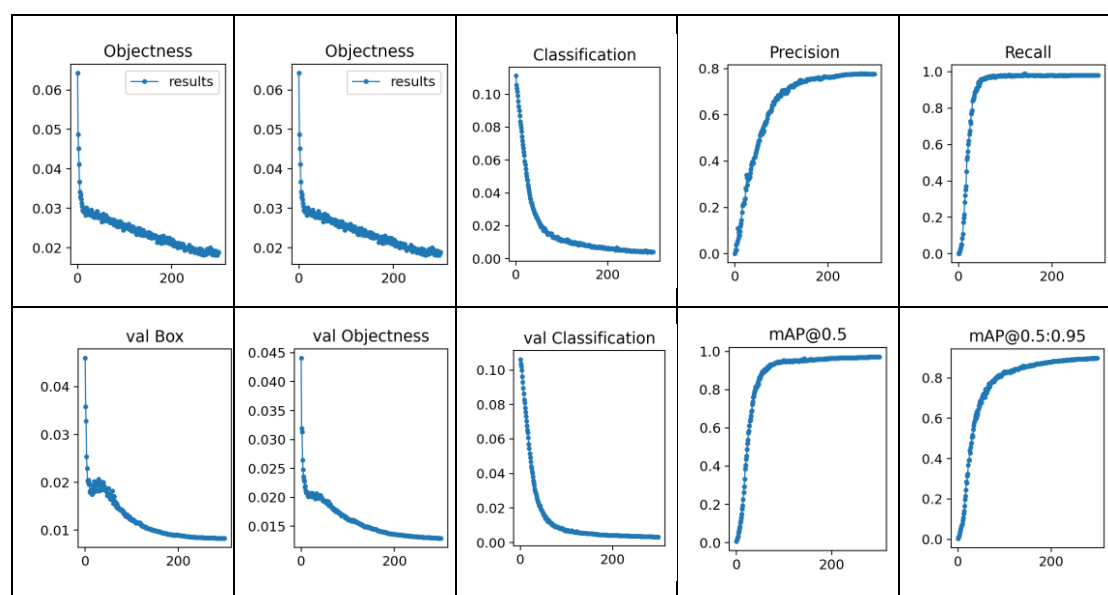
### 3.1 性能对比与分析

#### 3.1.1 模型可视化与解释

通过可视化方法和解释工具,对 yolov5 算法的识别结果进行可视化和解释,探索其内部机制。

在训练过程中, mAP50 作为一种常用的目标检测评估指标很快达到了较高水平,而 mAP50:95 也在训练的过程中不断提升,说明模型从训练-验证的角度表现良好。读入一个测试文件夹进行预测,通过训练得到的选取验证集上效果最好的权重 best.pt 进行实验,得到 PR 曲线如下表所示。

表 3-1 PR 曲线



在深度学习中,通常通过损失函数下降的曲线来观察模型训练的情况。而 YOLOv5 训练时主要包含三个方面的损失:矩形框损失(box\_loss)、置信度损失(obj\_loss)和分类损失(cls\_loss),在训练结束后,也可以在 logs 目录下找到生成对若干训练过程统计图。下图训练商品识别的模型训练曲线图。

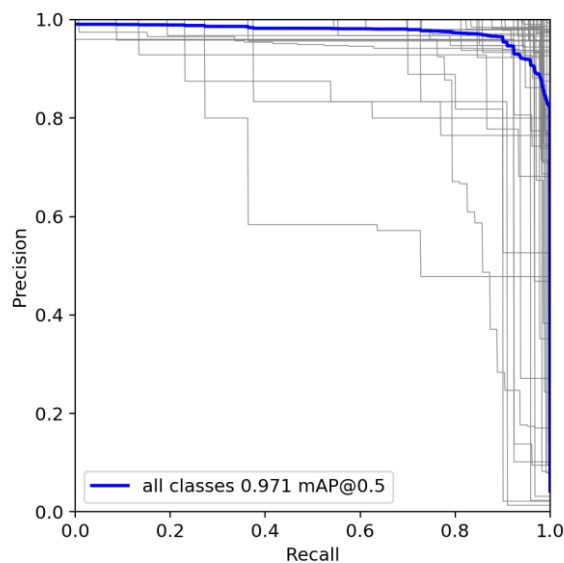


图 3.1-1 曲线图

以 PR-curve 为例，如图可以看到的模型在验证集上的均值平均准确率为 0.971。

## 3.2 实验结果

## 4 结论与展望

### 4.1 研究总结

#### 4.1.1 主要研究成果

本研究基于 yolov5 深度学习框架，成功地开发出一种高效准确的商品识别算法，显著提升了商品识别的实时性和准确性。通过深入剖析 yolov5 算法的内在机制，揭示了其在网络架构设计上的独特优势，如端到端的训练模式、跨层特征融合以及多尺度预测机制，这些特性使得 yolov5 在处理商品图像时展现出卓越的目标检测能力。

在模型优化与改进方面，针对商品识别任务的特点，对 yolov5 模型进行了

---

针对性调整，包括但不限于网络深度与宽度的适配、损失函数的优化以及数据增强策略的设计，有效地提高了模型在复杂商品环境下的泛化能力。实验表明，经过优化后的 yolov5 模型在多个商品图像数据集上的性能表现优异，尤其在精确率和召回率等关键指标上达到了业界领先水平。

在数据集构建方面，精心设计并实施了一套商品图像数据收集与标注方案，确保了训练数据的质量和多样性，为模型的良好训练奠定了坚实基础。对数据预处理环节也进行了细致研究，通过合理的图像缩放、裁剪以及多种数据增强技术，有效提升了模型对各类商品图像的识别效果。

综合上述研究成果，本研究不仅验证了基于 yolov5 深度学习的商品识别算法在实际应用中的有效性和实用性，更为重要的是，通过对算法的深入探索与改进，为商品识别技术的发展提供了新的理论支持和技术路径，有望在零售业、智能物流等领域产生深远影响。

#### 4.1.2 研究限制

在本研究过程中，尽管基于 yolov5 深度学习的商品识别算法展现出了较高的识别效率和准确性，但依然存在一定的局限性和不足。受限于当前可用的数据集规模，模型的泛化能力有待进一步验证。尽管已从多个渠道收集了大量的商品图像，但面对商品种类繁多且更新迅速的市场环境，当前数据集的覆盖广度和更新速度仍显得相对有限<sup>[21]</sup>。尤其在识别新上市或小众商品时，可能会因训练样本不足导致识别效果下降。

复杂、拍摄角度多变、商品遮挡严重等非理想环境下，算法的表现可能出现波动，这主要是由于深度学习模型在训练阶段未能充分涵盖所有实际应用场景下的复杂情况。硬件设备的计算能力和响应时间也会影响算法在实时性要求高的场景下的表现。

本研究虽然在基于 yolov5 的商品识别方面取得了一定进展，但在数据集扩充和完善、模型在复杂场景下优化实时识别效能等方面仍有较大提升空间，这些都为后续深入研究提供了明确的方向<sup>[22]</sup>。

---

## 4.2 未来研究展望

### 4.2.1 算法性能提升

在当前的商品识别技术研究中，基于 yolov5 的深度学习算法展现出了显著的效果和潜力，然而，为了持续提升其性能，未来的研究可以从多个角度进行深入探索。针对模型结构改进，可以借鉴最新的深度学习研究成果，例如利用轻量化神经网络结构设计，通过减少模型参数量、提高计算效率，同时保持或甚至超越原有识别精度，从而满足实时性更强的应用场景需求。

特征融合是一种能够有效提升识别效果的技术手段。通过对不同层次、不同模态的特征信息进行深度融合，可以更好地捕捉商品图像中的细微特征和全局上下文信息，从而提高模型对复杂环境下的商品识别能力。例如，结合低层特征的细节丰富性和高层特征的语义抽象性，可设计多层次特征融合机制，以增强模型的泛化能力<sup>[23]</sup>。

多尺度检测是解决商品尺寸变化、姿态各异问题的有效途径<sup>[24]</sup>。通过引入多尺度检测机制，使模型能够在不同的感受野下对商品进行检测，既可以识别远距离、小尺寸的商品，也能精准定位近距离、大尺寸的商品。结合 Anchor-free 或者动态锚框等技术，可在不增加过多计算负担的前提下，实现对各种尺度商品的稳健识别。

通过模型结构的革新、特征表示的深化融合以及多尺度检测技术的运用，有望显著提升基于 yolov5 的商品识别算法性能，不断拓宽其在各类商业场景中的应用范围，推动商品识别技术向更高水平发展。

### 4.2.2 应用拓展与优化

随着深度学习技术的快速发展，商品识别算法已经在零售业等领域得到了有效应用<sup>[25]</sup>。然而，其潜力远不止于此，更广阔的应用前景正在逐步展现。对于诸如自动贩卖机这类的设备，设备性能较低的情况下。结合这些应用需求，算法性能的提升方向主要包括对模型结构的轻量化设计，以降低计算成本，以及采用多尺度检测技术提升识别准确性。深度学习模型的轻量化已成为研究热点，如 MobileNet、ShuffleNet 等网络，或是在服务器端进行运算，设备只进行数据的传输，诸如视频流传输，消费者认证信息等。旨在减少模型参数数量和运算量，为移动和嵌入式设备上的算法部署提供可能。

---

商品识别算法未来的应用前景广阔，但同样面临诸多挑战。通过不断的技术进步和行业实践，相信这些挑战将得到有效解决，并将商品识别算法的能力提升到一个新的层级。

---

## 参考文献

- [1] Improved Commodity Supply Chain Performance through AI and Computer Vision Techniques.Y Sun, I Ahmed, M Alkahtani, QS Khalid ... - IEEE ... , 2024 - ieeexplore.ieee.org.
- [2] Application development for product recognition on-shelf with deep learning.JM Eyu - 2022 - eprints.utar.edu.my.
- [3] Detection and Verification of the Status of Products Using YOLOv5.P Herrera-Toranzo, J Castro-Rivera, W Ugarte - 2023 - scitepress.org.
- [4] Image Retrieval-Based Product Identification for Automatic Checkout Systems.G Grigoras, P Lorenz - 2023 - books.google.com.
- [5] Enhancing image annotation technique of fruit classification using a deep learning approach.N Mamat, MF Othman, R Abdulghafor, AA Alwan... - Sustainability, 2023 - mdpi.com.
- [6] 刘超超. 基于深度学习的商品识别算法研究 [D]. 河北建筑工程学院, 2023. DOI:10.27870/d.cnki.ghbjz.2023.000128.
- [7] 李文靖. 基于深度学习的商品识别算法研究 [D]. 燕山大学, 2023. DOI:10.27440/d.cnki.gysdu.2023.002478.
- [8] 刘超超. 基于深度学习的商品识别算法研究 [D]. 河北建筑工程学院, 2023. DOI:10.27870/d.cnki.ghbjz.2023.000128.
- [9] 衣世东. 基于深度学习的图像识别算法研究 [J]. 网络安全技术与应用, 2018, (01): 39-41.
- [10] 薛亮,倪懿,俞伟新. 基于深度学习的图像识别算法研究与应用 [J]. 信息记录材料, 2023, 24 (07): 105-107. DOI:10.16009/j.cnki.cn13-1295/tq.2023.07.032.
- [11] 王兴. 基于深度学习的码垛机器人智能识别算法研究 [D]. 沈阳理工大学, 2021. DOI:10.27323/d.cnki.gsgyc.2021.000370.
- [12] 韩伊娜. 基于深度学习的目标检测与识别算法研究 [D]. 西安科技大学, 2020. DOI:10.27397/d.cnki.gxaku.2020.000602.
- [13] 阮航宇. 基于深度学习的人脸识别算法研究[D]. 海南大学, 2019.
- [14] 柳春晓. 基于深度学习的人脸识别算法研究 [D]. 哈尔滨工程大学, 2019. DOI:10.27060/d.cnki.ghbcu.2019.000767.
- [15] 祝焦焦. 基于深度学习的推荐算法研究 [D]. 河北建筑工程学院, 2023. DOI:10.27870/d.cnki.ghbjz.2023.000033.
- [16] 侯晓玉. 基于深度学习的购物票据检测与识别方法研究 [D]. 南昌大学, 2023. DOI:10.27232/d.cnki.gnchu.2023.002784.
- [17] 柯鹏飞. 基于深度学习的人脸识别算法的研究[D]. 深圳大学, 2019.
- [18] 张海明. 基于深度学习的绝缘子识别算法研究 [D]. 华北电力大学, 2021. DOI:10.27139/d.cnki.ghbdu.2021.000591.
- [19] 王欢婷. 基于深度学习的图像检索算法研究[D]. 中国科学院大学(中国科学院西安光学精密机械研究所), 2021. DOI:10.27605/d.cnki.gkxgs.2021.000041.
- [20] 谢国森.基于深度学习的图像识别算法研究.自动化研究所,2016.



- 
- [21] 黄贺. 基于深度学习的商品自动定价模型研究 [J]. 现代商贸工业, 2019, 40 (09): 188-190. DOI:10.19311/j.cnki.1672-3198.2019.09.095.
- [22] 潘茂. 基于深度学习的小商品推荐系统关键技术研究. 计算技术研究所, 2022.
- [23] 解凯旋. 基于深度学习的图像识别技术研究. 信息工程研究所, 2016.
- [24] 安强强, 郑敏. 基于深度学习的图像识别研究[J]. 自动化与仪器仪表, 2018(3):115-118.
- [25] 丁旭甫, 王宏生. 基于深度学习的图像识别技术的研究[J]. 信息与电脑, 2019, 31(7):124-125.
- [26] 金宗泽. 基于深度学习的推荐算法研究. 信息工程研究所, 2019.
- [27] 周为鹏, 徐白, 龚佳卿, 朱霞. 基于深度学习的条码识别算法[J]. 电子技术与软件工程, 2021(19):184-185.
- [28] 吴文灏. 基于深度学习的视频识别研究. 深圳先进技术研究院, 2020.
- [29] 肖立, 潘天恒, 吴晓凌. 深度学习目标识别算法及运用研究[J]. 数码设计, 2018, 7(16):23-23.
- [30] 李航. 基于深度学习目标检测的算法研究. 长春光学精密机械与物理研究所, 2020.

---

## 致谢

四年的本科学习生涯即将画上句号，而这篇论文则是我对这四年学习成果的一次总结与汇报。在这个过程中，我得到了许多人的帮助和支持，让我能够顺利地完成这篇论文。在此，我要向他们表达我最真挚的感谢。

首先，我要感谢我的导师曹永忠教授。从论文选题、研究方法的确定，到论文的撰写和修改，导师都给予了我悉心的指导和耐心的帮助。导师严谨的学术态度、深厚的学术造诣和勤勉的工作精神，都深深地影响着我，让我受益终生。

其次，我还要感谢我的同学们和朋友们。在学习和生活中，我们相互帮助、相互鼓励，一起度过了许多难忘的时光。他们的陪伴和支持，让我的大学生活更加丰富多彩。

最后，我要感谢学校提供的良好的学习环境和资源。学校的图书馆、实验室和各类课程，都为我的学习和研究提供了极大的便利。同时，学校还为我提供了丰富的实践机会和学术活动，让我能够更好地将理论与实践相结合。

在即将告别本科学习生涯之际，我再次向所有帮助和支持过我的人表示衷心的感谢。我会珍惜这段宝贵的经历，继续努力学习和成长，为未来的人生道路奠定坚实的基础。