
分类号: TP391

单位代码: 10363

密 级: 公开

学 号: 2220320197

题 目 基于深度学习的工业钢材表面缺陷检测算法研究

题 目 Research on industrial steel surface defect detection
algorithms based on deep learning

学生姓名: 李晓宇

指导教师: 瞿成明 副教授

专 业: 电子信息

研究方向: 人工智能

论文答辩日期: 2025/05/12

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：李 晓 宇

日期：2025 年 5 月 15 日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 __ 年解密后适用本版权书。

本学位论文属于

不保密 ☐。

学位论文作者签名：李 晓 宇 指导教师签名：瞿成明

日期：2025 年 5 月 15 日

日期：2025 年 5 月 15 日

基于深度学习的工业钢材表面缺陷检测算法研究

摘要

在“中国制造 2025”战略的影响下，钢铁作为制造业的基础材料之一，在加工过程中，工业检测钢材表面缺陷的方法，与基于深度学习的目标检测技术深度融合，取代了人工肉眼检测 and 传统物理方法检测。借助深度学习的目标检测技术，不仅能够实现自动化检测，还可以同时识别多种工业钢材表面缺陷，显著提升了检测精度，展现出了更强的鲁棒性。然而，基于深度学习的目标检测技术在检测一些尺度变化较大的缺陷和与图片背景相似的缺陷时，存在平均检测精度低的问题。本文基于 YOLOv8n 算法，针对平均检测精度低的问题对算法进行改进，考虑将算法移植到计算资源有限的工业设备时，需要兼顾算法的平均检测精度和实时性问题，对算法进行了轻量化，并设计了工业钢材表面缺陷检测系统。本文所作的研究如下：

(1) 在算法模型进行训练时，损失函数可以减少预测结果与真实标签之间的差异，通过对比实验，分析了 GIoU、DIoU 和 CIoU 三种损失函数的损失曲线，选出使用 CIoU 损失函数提升模型的检测性能。针对缺陷的特征提取受到图片背景的干扰问题，提出使用注意力机制提高缺陷的特征提取，抑制不重要特征信息。在算法模型不同位置设计了三种注意力机制添加方案，通过对比实验，分析了 CBAM、ECA、SK、EMA、SE 五种注意力机制在模型检测中对缺陷特征的提取情况。其中，SE 注意力机制对缺陷特征提取效果最佳，模型的检测结果在平均检测精度中最高。针对工业钢材表面缺

陷种类复杂、尺度变化大的问题，引入可变形卷积，并与 C2f 模块结合构建出 C2f_DCNv2 模块，提高模型对各种缺陷特征的学习能力。

(2) 将各种最佳改进结构融入 YOLOv8n 算法模型中，确定最终的 YOLOv8n-Improved 算法模型；使用 NEU-DET 数据集并搭建工业钢材表面缺陷检测实验环境，对 YOLOv8n-Improved 算法模型进行消融实验。实验结果表明，该模型在 mAP 上达到了 79.2%，比 YOLOv8n 模型的 mAP 高出 1.9%。分析检测后的缺陷图片，YOLOv8n-Improved 算法改善了 YOLOv8n 算法对一些缺陷的漏检、误检情况，而且在各类缺陷的平均检测精度上都有所提高。

(3) 将改进的算法模型移植到工业中计算资源较小的嵌入式设备时，需要降低算法的计算量，同时保持算法的平均检测精度尽可能不降低。因此，选择对 YOLOv8n-Improved 模型进行轻量化。通过实验，对比分析了分别使用 GhostNet 卷积、深度可分离卷积和 DSConv 卷积对模型进行轻量化的效果，选择平衡性较好的 DSConv 卷积作为轻量化的改进方案，并结合 PySide6 设计了工业钢材表面缺陷检测系统。该系统可以对工业钢材表面各类缺陷进行检测，并把检测缺陷的种类、个数以及平均检测精度显示到系统检测界面中。

关键词：缺陷检测，深度学习，注意力机制，YOLOv8n，卷积网络

RESEARCH ON INDUSTRIAL STEEL SURFACE DEFECT DETECTION ALGORITHMS BASED ON DEEP LEARNING

ABSTRACT

Under the influence of the “Made in China 2025” strategy, steel, as one of the basic materials for manufacturing, is processed in the process of industrial detection of steel surface defects, and deep learning-based target detection technology is deeply integrated, replacing manual visual inspection and traditional physical methods. With the help of deep learning target detection technology, not only can automated detection be realized, but also can simultaneously identify a variety of industrial steel surface defects, which significantly improves the detection accuracy and shows stronger robustness. However, deep learning-based target detection technology has the problem of low average detection accuracy when detecting some defects with large scale changes and defects similar to the background of the picture. In this paper, based on the YOLOv8n algorithm, the algorithm is improved for the problem of low average detection accuracy, and considering the need to take into account the average detection accuracy of the algorithm and real-time problem when the algorithm is transplanted to the industrial equipment with limited computational resources, the algorithm is lightened, and the defect detection system for the surface defects of industrial steel is designed. The research done in this paper is as follows:

(1) When the algorithm model is trained, the loss function can reduce the difference between the prediction results and the real labels, and the

loss curves of the three loss functions of GIoU, DIoU and CIoU are analyzed through comparative experiments, and the use of the CIoU loss function is elected to improve the detection performance of the model. Aiming at the problem that the feature extraction of defects is interfered by the background of the picture, it is proposed to use the attention mechanism to improve the feature extraction of defects and suppress the unimportant feature information. Three attention mechanism addition schemes are designed in different positions of the algorithm model, and the extraction of defect features by the five attention mechanisms CBAM, ECA, SK, EMA, and SE in the model detection is analyzed through comparison experiments. Among them, the SE attention mechanism has the best effect on defect feature extraction, and the detection results of the model are the highest in the average detection accuracy. Aiming at the problem of complex types of defects on the surface of industrial steel with large scale changes, deformable convolution is introduced and combined with the C2f module to construct the C2f_DCNv2 module, which improves the model's ability to learn various defect features.

(2) Incorporate various best improved structures into the YOLOv8n algorithm model to determine the final YOLOv8n-Improved algorithm model; use the NEU-DET dataset and set up an experimental environment for industrial steel surface defects detection to conduct ablation experiments on the YOLOv8n-Improved algorithm model. The experimental results show that the model achieves 79.2% in mAP, which is 1.9% higher than the mAP of the YOLOv8n model. Analyzing the pictures of the detected defects, the YOLOv8n-Improved algorithm improves the leakage and misdetection of some defects by the YOLOv8n algorithm, and it also improves the average detection accuracy in all types of defects.

(3) When porting the improved algorithmic model to embedded

devices in industry with small computational resources, it is necessary to reduce the computational effort of the algorithm while keeping the average detection accuracy of the algorithm as undiminished as possible. Therefore, the YOLOv8n-Improved model was chosen to be lightweighted. Through experiments, the effects of using GhostNet convolution, depth-separable convolution and DSConv convolution respectively to lighten the model were compared and analyzed, and the better-balanced DSConv convolution was chosen as an improved solution for lightening, and a defect detection system for industrial steel surfaces was designed in combination with PySide6. The system can detect all kinds of defects on the surface of industrial steel, and display the type, number and average detection accuracy of the detected defects into the system detection interface.

KEY WORDS: Defect detection, Deep learning, Attention mechanism, YOLOv8n, Convolutional networks

目 录

摘 要	I
ABSTRACT	III
第 1 章 绪论	1
1.1 研究背景及意义	1
1.2 工业钢材表面缺陷检测国内外研究现状	3
1.2.1 基于传统物理的方法	3
1.2.2 基于传统目标检测的方法	4
1.2.3 基于深度学习的目标检测方法	5
1.3 研究现状分析	10
1.4 主要研究内容与章节安排	11
第 2 章 深度学习目标检测理论	14
2.1 卷积神经网络	14
2.1.1 卷积层	15
2.1.2 池化层	15
2.1.3 全连接层	16
2.1.4 激活函数	17
2.2 两阶段目标检测算法	18
2.3 单阶段目标检测算法	19
2.4 评价指标	24
2.5 本章小结	25
第 3 章 YOLOv8n 网络模型改进	26
3.1 选取损失函数	26
3.1.1 GIoU 损失函数	26
3.1.2 DIoU 损失函数	27
3.1.3 CIoU 损失函数	27
3.1.4 实验结果对比与分析	27

3.2 引入注意力机制模块	28
3.2.1 CBAM 注意力机制	29
3.2.2 ECA 注意力机制	30
3.2.3 SK 注意力机制	30
3.2.4 EMA 注意力机制	31
3.2.5 SE 注意力机制	32
3.2.6 实验结果对比与分析	33
3.3 可变形卷积改进	36
3.3.1 可变形卷积 v1	37
3.3.2 可变形卷积 v2	38
3.3.3 可变形卷积改进模型结构	39
3.4 本章小结	40
第 4 章 基于改进 YOLOv8n 的工业钢材表面缺陷检测算法	42
4.1 改进 YOLOv8n 的算法模型	42
4.2 实验过程及分析	43
4.2.1 工业钢材表面缺陷数据集	43
4.2.2 实验环境	44
4.2.3 消融实验	46
4.2.4 实验结果对比与分析	46
4.2.5 主流算法对比与分析	52
4.3 本章小结	53
第 5 章 工业钢材表面缺陷检测算法的轻量化及系统设计	54
5.1 模型轻量化	54
5.1.1 GhostNet 卷积	55
5.1.2 深度可分离卷积	55
5.1.3 DSCConv 卷积	57
5.1.4 实验结果对比与分析	58
5.2 工业钢材表面缺陷检测系统设计	59
5.2.1 系统方案设计	59
5.2.2 系统开发环境	59

5.2.3 PySide6 简介	60
5.3 系统主界面设计	61
5.4 系统功能测试	63
5.5 本章小结	64
第 6 章 总结与展望	65
6.1 总结	65
6.2 展望	66
参考文献	67
攻读学位期间主要研究成果	73
致谢	74

第 1 章 绪论

1.1 研究背景及意义

钢铁作为制造业的基础材料之一，具有广泛的应用领域和重要的经济作用。钢铁产品被广泛应用于家电、军工、交通、机械设备等，如图 1-1 所示，钢铁就是制造飞机和机器人的材料之一。而在经济领域，中国已经成为钢铁的生产大国，其钢铁产品在国际贸易中占据了重要地位。2024 年，中国累计出口钢材约为 10115.2 万吨，同比增长 22.6%，据估算，2024 年中国钢材出口量将达到 1.1 亿吨。钢铁产品的出口贸易不仅为国家赚取了大量外汇，还促进了国内钢铁行业的发展和繁荣。钢铁产品大规模的使用，使人们更加关注其性能，中国作为钢铁生产以及出口大国，要对生产钢铁的质量要严格把控，出产高质量和高品质的钢铁已经成为了各钢材制造企业的迫切追求。在“中国制造 2025”战略^[1]的影响下，我国钢铁行业逐步与新一代信息技术融合，朝着智能化方向发展^[2]。



(a)飞机



(b)机器人

图 1-1 钢铁材料的应用

中国钢铁行业正在经历规模迅速扩张、产量飞速提升的一段时期，该时期着力于解决国民经济发展和用钢问题。但是，钢铁在生产和加工过程中，容易受到环境中不良因素的影响，导致钢材表面出现多种类型的缺陷。划痕、开裂、斑块等典型的钢材表面缺陷会导致钢材的性能和使用年限降低，甚至直接造成安全事故的发生^[3]。企业更加重视对钢材表面缺陷的检测，甚至联合学者一起对相关的技术进行研究。起初，工业钢材表面缺陷检测主要依赖人工目视检测法，但是该方法在检测的过程中不仅会消耗大量的人力财力，而且检测时间耗时过长，容易使检测人员产生疲劳，从而导致检测效率下降，严重时甚至会出现误判，影响检测结果的准确性。随后是频闪检测法，该方法通过利用频闪光

源快速照射到钢材表面，观察者通过巡视频闪下的钢材表面来发现缺陷，但是此种方法本质上仍然依赖人眼进行判断，这意味着它同样受限于观察者的主观因素。随后又出现一些传统的物理检测方法，进行特定产品的缺陷检测，例如微波检测^[4]、无损检测^[5]、线结构光扫描检测^[6]等，但是，工业中产品的形式多种多样，因此这些方法不适用于多类型产品的检测。

随着“中国制造 2025”战略三步走中，第二步智能化升级阶段即将完成，制造业与新一代信息技术将深度融合，制造业向数字化、网络化、智能化发展速度加快。智能制造成为“中国制造 2025”的制高点、突破口和主攻方向，其带来的制造新理念和新技术给制造行业带来极大的影响。近年来，计算机视觉领域中基于深度学习的目标检测算法有着较好的发展，在图像识别与分类方面，通过利用深度学习技术，计算机不仅可以准确识别图像中的物体，还可以处理复杂背景和多物体的图像。在智能决策方面，计算机视觉系统可以辅助汽车自动驾驶以及为汽车的行驶规划路线^[7]。在与传统的计算机视觉技术比较时，这些深度学习的目标检测算法模型有着卓越的性能，展现出了深度学习技术的潜力。目标检测在于定位和分类物体，并识别物体类别。当我们在检测工业钢铁表面缺陷时，利用基于深度学习的目标检测算法，使用训练好的网络模型，缺陷所在区域可以在毫秒级别时间内被检测出来。相比于传统检测方法，深度学习缺陷检测不再局限于某种固定程序以及固定材料，而是通过神经网络的自主学习，克服了传统方法只能检测单一缺陷的局限性^[8]。深度学习中的卷积神经网络因其本身具有检测稳定性高、泛化能力强的特性，运用到工业钢材表面缺陷检测时，不仅可以为企业节约人力成本，减少开支，还可以提高对缺陷检测的精度和速度，从而减少生产钢铁时出现废品的概率，促进工业钢材表面缺陷自动化检测的发展。总之，在“中国制造 2025”战略中的第二步收官之际，深度学习与工业领域必将深度结合，深度学习对于提高工业产品的质量与生产的效率，将起到至关重要的技术支撑作用。在智能制造行业中，为积极配合“中国制造 2025”战略的实施，开展基于深度学习的工业钢材表面缺陷检测算法研究，将基于深度学习的目标检测技术同工业钢材表面缺陷检测相结合具有重要意义。

1.2 工业钢材表面缺陷检测国内外研究现状

目前工业钢材表面缺陷检测技术经历了从人工目测检测技术，到涡流、漏磁检测技术，然后到计算机视觉检测技术三个阶段的发展历程。而在计算机视觉领域中，目标检测是其中的关键环节之一，它用来对检测目标的定位以及类别的预测。目标检测算法大致分为两个阶段：传统的目标检测算法和基于深度学习的目标检测算法。而随着基于深度学习的目标检测算法的发展，基于深度学习的目标检测算法逐渐融合到工业钢材表面缺陷的检测中，并且算法的检测性能在不断改善。

1.2.1 基于传统物理的方法

人工目视检测法和频闪检测法是钢材表面缺陷检测最初使用的方法，这两种方法的核心都是工作者利用眼睛对钢材表面进行观察来发现缺陷，因此这两种方法不仅需要工作人员在识别钢材表面缺陷上具有丰富的经验，还需要具有较强的长时间工作能力。由于过度依赖人，而且对工作者本身的要求还比较高，显而长时间的进行钢材表面缺陷的检测会使检测效率变低，甚至还有可能会出现严重的误检、漏检情况。

随后出现了一些物理检测方法对工业钢材表面缺陷进行检测，例如涡流检测法^[9]、红外检测法^[10]、漏磁检测法^[11]、超声波检测法^[12]等，其中涡流检测法是利用电磁感应原理对缺陷进行检测，当遇到缺陷时会出现涡流变化；红外检测法主要针对温度敏感的缺陷，当检测钢材表面时根据温度的不同分布来确定缺陷；漏磁检测法是基于铁磁性材料的高磁导率特性和磁偶极子理论，当钢材表面出现缺陷时，这些缺陷会引起材料磁导率的变化进而确定缺陷；超声波检测法通过利用超声波在不同介质界面会发生反射来确定钢材表面是否有缺陷，根据反射波的时间和幅度等特征来确定缺陷的位置、大小和形状等。虽然这些物理检测法为钢材表面缺陷的检测提供了更多的选择，然而这些物理检测法的一些局限性不足以应对现阶段工业场景中复杂多变的缺陷类型，而且本身在应用场景方面也有许多的局限性，更加不能满足工业场景中准确性与实时性的要求。

1.2.2 基于传统目标检测的方法

在深度学习中的卷积神经网络运用到工业钢材表面缺陷检测之前，传统的机械视觉目标检测一直是工业钢材表面缺陷检测的主流，它通过使用 CCD 相机获取图片，然后对图片进行预处理，包括对图像进行灰度化、去噪等操作，随后手工设计特征提取方式，常用的特征提取方式有尺度不变特征变换和方向梯度直方图等，最后利用机械学习分类器对图片进行训练和分类，常用的分类器有支持向量机（Support Vector Machine,SVM）^[13]和自适应提升（Adaboost）^[14]。

随着检测技术的不断发展，研究者在传统机械视觉工业钢材表面缺陷检测中也做出了许多贡献。2020 年，Zhang 等人^[15]在大部分人使用同一种方法检测工业钢材表面所有缺陷类型时，提出了一种基于模糊测度的新颖方法来进行带钢表面的缺陷检测。首先根据调查统计带钢图像的背景强度服从高斯分布这一条件，创建给定测试图像的直方图，从中估计背景模型参数，然后定义隶属度函数以估计每个灰度级属于缺陷的程度，通过该隶属度函数，测试图像的每个像素都将获得隶属度值。最后，结合像素连通性，模糊连接区域的最大值和总和用于定位缺陷。实验结果表明，该方法检测率高达 96.8%。

2021 年，Chaudhari 等人^[16]综合各种特征提取技术设计了一款基于钢材表面缺陷检测系统，该系统的特征提取技术有灰度共生矩阵、Gabor 小波和 HOG 梯度直方图，在最后的特征分类时采用随机森林分类器进行分类。实验结果表明，该系统的分类准确率达到 95.0%。

2024 年，肖玲等人^[17]针对采集钢材表面缺陷时，图片会受到光照等外部环境的影响，从而使得图片中的缺陷特征不明显以及噪声大等问题，提出使用图像处理进行钢材表面缺陷检测的方法。首先，采用 CLAHE 算法和双边滤波算法用来突出缺陷特征和去除噪声；然后，采用改进的 Bottom-hat 算法用来增强缺陷特征和图片背景的对比度，最后，采用基于 LBP 与 SVM 相结合的方法用于检测和分类。实验结果表明，此方法在钢材表面缺陷数据集中的检测准确率达 81.67%。

虽然传统的机械视觉目标检测采用机器进行检测，在一定程度上避免了人眼对钢材表面缺陷进行检测的局限性，但是在设计一个完整的特征提取方式时，需要进行大量的实验，即使一个相同的特征提取方式在进行不同的目标检测任

务时，产生的结果也是不同的。因此这种只能针对固定特征进行检测的方法不适用于现代各类工业环境。

1.2.3 基于深度学习的目标检测方法

基于深度学习的目标检测在计算机视觉检测技术的不断发展中，逐渐走向了历史的舞台，基于深度学习的目标检测算法使用的是卷积神经网络学习特征的方式，这种特征学习方式能自动发现检测及分类目标所需要的特征。早在 2006 年，Hinton 就提出了能力很强的深层神经网络架构学习特征，这可以学习到样本更加本质的模式，如果将网络在大型数据集上进行预训练，可以使得深度网络架构更容易找到全局最优点^[18]。这些极其重要的观点为深度学习的发展奠定了基础。在 2012 年，Hinton 课题组通过构建的卷积神经网络 AlexNet^[19]，在 ImageNet 图像识别比赛中获得了冠军，并且网络的性能比第二名（SVM 方法）更好。在比赛结束后，卷积神经网络开始受到众多学者的关注，基于深度学习的算法开始快速发展，因此这一年被学术界称为深度学习的元年。

现阶段，基于深度学习的目标检测主要分为两种，单阶段目标检测和两阶段目标检测。在两阶段目标检测中，首先进行目标候选区域的生成，然后对候选区域中的物体进行分类并不断精修候选区域的位置，使候选区域可以更好的包含物体^[20]。而单阶段目标检测直接在输入的图像上使用卷积层生成目标类别的概率和目标位置的坐标。根据不同的场景与检测目的，各种检测算法均会展现出各自的优点，据此可以选择较为适合的算法对目标进行检测，学者们也在不断推进着各种算法的研究。

（1）两阶段目标检测算法

2022 年，Wang 等人^[21]发现分类和回归对目标翻译的敏感性不同，提出了一种名为 Decoupled R-CNN 的深度检测器的解耦采样策略。因为回归器的检测头具有比分类器对翻译更敏感的内在特征，所以采用解耦采样的策略来解耦这两个任务的建议采样，从而产生两个灵敏度特定的检测头。同时对解耦 R-CNN 的单个回归器头采用级联结构来提高目标检测的性能。通过在数据集上进行检测，其平均检测精度提高了 2.6%，达到了 82.2%。

2023 年，Fang 等人^[22]针对交通场景中的多目标检测存在精度低的问题，对 Mask R-CNN 算法进行了改进。将骨干网络 ResNet 替换为具有组卷积的

ResNeXt 网络，以进一步提高模型的特征提取能力；然后在特征金字塔网络（Feature Pyramid Network, FPN）中加入自下而上的路径，增强特征融合；同时加入 ECA 注意力机制，增强网络对目标的特征提取能力；最后，使用 CIoU 损失函数，加快模型收敛速度。通过实验，改进的 Mask R-CNN 算法在公开的 CityScapes 自动驾驶数据集上，平均检测精度达到了 62.62%。

2024 年，Hu 等人^[23]针对在图片中细胞检测和分割时，生物细胞形状多、尺寸差异大、灰度不均匀和生物细胞间分布密集的问题，由此对 Mask R-CNN 算法进行改进，提出了改进算法 Mask R-DHCNN。该算法中提出了一种新的卷积膨胀异构卷积，并将 Mask R-CNN 中的卷积替换成膨胀异构卷积，以适应检测和分割任务中细胞形状和大小的不同。在生物细胞数据集上进行检测并与 Mask R-CNN 算法进行比较，改进的算法在检测精度和检测速度上都有着较好的提升。

2024 年，Liu 等人^[24]在对微小物体检测时，提出了带有去噪 FPN 的 Transformer R-CNN 算法。首先，在 FPN 的特征融合中引入了一个 DN-FPN 模块，利用对比学习来抑制 FPN 自上而下路径中每个级别特征中的噪声。其次，基于两阶段框架，使用最新的 Transformer R-CNN 检测器替换传统的 R-CNN 检测器，使算法更加专注于微小物体。实验结果表明，改进后的算法在 AI-TOD 数据集上进行检测时，比改进前平均检测精度高出了 17.4%。

（2）单阶段目标检测算法

2022 年，Xu 等人^[25]针对 YOLOv3 算法处理行人检测时会出现漏检、误检以及行人多姿态变化的问题，对 YOLOv3 算法进行改进。在算法中引入 CSPNet 结构对特征图进行分割和跨阶段连接，实现特征的重复使用，提高算法的检测精度，并且大幅度降低了网络的参数量。同时，引入 ECA 注意力机制来提高对特征图的学习。实验结果表明，改进的 YOLOv3 算法在 INRIA 数据集中平均检测精度达到了 93.89%。

2023 年，Zhang 等人^[26]针对人工搜索野生蘑菇效率低下问题，提出了一种改进的 YOLOv4 算法模型。首先在 YOLOv4 为框架的前提下，采用高效的 ShuffleNetv2 特征提取网络和轻量级自适应空间特征融合结构，减少网络的参数量和计算量。然后对于检测分支，采用深度可分离卷积（DWConv）和金字塔卷积（PyConv）代替普通卷积，实现轻量级改进。最后，在网络的输出端引

入 SA 注意力模块，以较小的计算成本补偿轻量级改进带来的精度损失。实验结果表明，在 1112 张野生蘑菇图像上进行检测时平均检测精度达到了 88.76%。

2024 年，Li 等人^[27]以苹果为研究对象，引入 YOLOv5 算法来解决传统果园管理中的苹果检测和数量估算问题。首先，利用相机获得大量果园图片；然后，将大量果园图片使用 YOLOv5 算法进行训练，使模型成功学习苹果的特征表示；最后，在新的果园图像中对学习特征后的算法进行验证。实验结果表明，在进行有效的特征提取和学习后，该模型准确检测和估计了水果数量，为果园管理提供了可靠的数据支持。

2024 年，Hu 等人^[28]发现在无人车技术的不断发展和应用过程中，越来越多的无人车必须在恶劣天气和各种复杂环境条件下工作，这给无人车感知系统的研究带来了更高的要求。尤其在大雾天或其他恶劣天气条件下，感知复杂的道路状态变得更加困难，于是提出了一种基于 YOLOv5 的多功能无人车视觉感知系统。首先，通过使用 Ghost Bottleneck 模块改善网络模型中计算效率和准确性之间的平衡关系；其次，利用 CBATM 模块增强模型的目标感知能力，提高在有雾场景下的检测精度；最后，结合 MSR 算法增强模型在雾场景下的鲁棒性，提高模型在复杂环境下感知目标的能力。基于 YOLOv5 的多功能无人车视觉感知系统，在集配送和巡检一体的多功能无人车应用中具有广阔的应用前景，并为实现智能感知和决策提供有力的支撑。

近年来，深度学习的目标检测算法逐渐应用于工业钢材表面缺陷检测。基于深度学习的工业钢材缺陷检测主要有两个子任务组成：缺陷定位和缺陷分类。两阶段目标检测方法通常分两步进行，首先获取目标候选区域的位置，然后基于候选区域中目标的特征判断缺陷的类别^[29]。而单阶段目标检测方法直接将缺陷定位和缺陷分类当作同一个任务进行回归，虽然在检测缺陷精度方面有所下降，但是在缺陷检测的速度上却得到了极大的提升^[30]。因为算法的普适性和可移植性，使其在工业场景检测中得到了广泛应用。并且随着深度学习目标检测算法的飞速发展，企业家联合学者开始提升算法的平均检测精度，降低算法的模型大小。

2019 年，Li 等人^[31]采用 Faster R-CNN 算法和 FPN 网络，提出一种 RPN 结构。该结构在多尺度的特征融合中可以更好的结合低层的特征，使钢材表面缺陷中的一些小目标缺陷得到更好的检测。实验结果表面，改进的算法平均检测

精度达到了 98.26%。

2020 年, Zhang 等人^[32]针对人工检测钢材表面缺陷误差率高的问题, 提出了两种基于图像压缩的检测方法, 即基于重构图像的检测方法和基于压缩表示的检测方法。在基于重构图像的检测方法中, 将 CAE 网络中的卷积层替换 YOLO-tiny 网络中的卷积层进行网络重建。在基于压缩表示的检测方法中, 以深度学习 YOLO-tiny 网络为检测骨干框架, 以卷积自编码器 CAE 为压缩预处理框架, 提高对缺陷的检测。实验结果表明, 提出的两种检测方法在平均检测精度上都要高于 YOLO-tiny 的平均检测精度。

2022 年, Yang 等人^[33]针对钢材表面缺陷检测的准确率较低会对生产和经济造成巨大损失的问题, 对 YOLOv5 网络模型进行改进。首先, 在缺陷图像的预处理方面使用同态滤波; 其次, 在 YOLOv5 网络中加入 CBAM (Convolutional Block Attention Module) 注意力模块, 获得更多与检测任务相关的重要信息来提高检测精度。实验结果表明, 提出的改进策略具有更好的缺陷检测能力和更高的准确率。

2023 年, Hang 等人^[34]针对传统的检测算法在进行钢材表面缺陷检测存在平均检测精度低的问题, 提出了一种改进的 YOLOv5 算法。首先, 在骨干网络中引入可变形卷积以获取更多的特征信息; 其次, 在骨干中加入 CBAM 注意力机制, 增强网络的特征提取能力; 然后, 使用 Focal EIoU 损失函数替代 CIoU 损失函数, 加速模型的收敛速度; 最后, 采用 K-means 算法对 Anchor box 进行聚类, 得到更适合本文的 Anchor box 参数。实验数据显示, 改进的算法平均检测精度达到了 78.8%, 比 YOLOv5 算法提高了 4.3%。

2023 年, Wang 等人^[35]针对检测不同尺度的钢材表面缺陷存在效率低的问题, 提出了一种改进的 YOLOv5 算法。为了有效探索钢材表面缺陷多尺度的信息, 在检测网络中开发了多尺度探索模块, 该模块由三部分组成: 瓶颈、头部和检测器。首先, 对输入图像进行瓶颈处理, 以探索多尺度特征; 然后, 对提出的特征进行聚合, 并由头部进一步处理; 最后, 将多尺度特征发送到检测器进行分类和定位。实验数据显示, 改进的算法在平均检测精度上达到了 72%。

2024 年, Han 等人^[36]针对钢材表面缺陷类型多、背景杂乱、对比度低和噪声干扰导致平均检测精度低的问题, 提出了一种改进的 YOLOv5 算法。首先, 利用 BiFPN 方法优化特征图的融合。其次, 引入 CARAFE 上采样来扩展网络

的感知场，更有效地利用周围信息。然后，引入 WIoU 损失函数，加速网络的收敛速度。最后，在网络预测阶段使用动态磁头，提高算法的缩放感知、空间感知和任务感知性能。通过实验，改进的算法在 NEU-DET 数据集上的平均检测精度达到了 81.1%，比原始模型提高了 6%。

2024 年，Sun 等人^[37]针对工业生产中钢材质量的要求越来越高，而传统的钢材表面缺陷检测方法检测性能差、检测速度慢的问题，提出了一种基于 YOLOv8 的名为 LMS-YOLO 的算法模型，以实现高效的钢材表面缺陷检测。首先，在骨干网中，设计轻量级多尺度混合卷积（LMSMC）模块与 C2f 模块融合得到 C2f_LMSMC 模块，从而提取不同尺度的特征进行融合，实现网络的轻量化。同时，在骨干网络中加入高效的全局注意力机制，增强跨维度信息交互和特征提取能力。在颈部网络中，使用 BiFPN 加权双向特征金字塔网络实现特征的跨尺度融合。最后，该模型使用三个独立的解耦头进行回归和分类，并采用 NWD 代替 CIoU 作为回归损失，提高对小目标缺陷的检测效果。实验结果表明，LMS-YOLO 在 NEU-DET 数据集上平均检测精度达到了 81.1%。

2021 年，Hao 等人^[38]针对需要大量计算能力的高精度深度卷积网络难以直接应用于工厂的问题，提出了一种基于 MobileNetv3 骨干模型的轻量级目标检测框架，同时为提高带钢表面缺陷的检测精度，对网络模型做了三方面的改动。分为用于多尺度检测的额外输出层、用于更丰富深度特征的空间金字塔池化模块和用于特征融合的特征金字塔自下而上的路径增强模块。实验结果表明，提出的网络模型的性能优于 YOLOv3、YOLOv4、SSD 和更快的 R-CNN 等网络模型。

2023 年，张政超^[39]针对工业场景检测中存在资源受限的问题，提出了一种轻量化 YOLOv5 的钢材表面缺陷检测算法。首先，使用 Shuffle Netv2 代替 YOLOv5 的主干网络，有效减少了模型的参数量并提高了运行速度；然后在网络中添加了轻量级上采样 CARAFE（Contentaware Reassembly Of Features）算子，在降低参数和计算量的同时提高了感受野的范围；最后，将 Conv 卷积层替换为 GSConv 卷积层，平衡模型的准确性与检测速度。通过实验，改进的 YOLOv5 模型相较于原始模型在平均精度均值上提升了 1.4%，达到了 78.5%，计算量和参数量分别降低 31%和 15.4%，改进后的算法模型能够对带钢表面缺陷进行实时检测。

2024 年, Zhu 等人^[40]针对传统的钢材表面缺陷检测效率低, 速度慢的问题, 提出了一种轻量高效的 LE-YOLOv5 算法, 应用于计算量低的工业场景中。首先, 使用 ShuffleNetv2 替换主干网络, 减少网络模型的参数量; 然后, 提出了一个 CBMM 模块, 在初始下采样阶段扩展模型的全局感受野, 提高模型捕获全局信息的能力; 最后, 设计了一个全局协调注意力机制, 有效地连接特征图中的位置和空间信息。通过实验, 算法的平均检测精度达到了 79.1%。

2024 年, Tie 等人^[41]针对现有基于深度学习的钢材表面缺陷检测模型在计算能力有限的终端设备中难以部署的问题, 基于 YOLOv8 算法模型, 提出了一种轻量化钢材表面缺陷检测模型 LSKA-YOLOv8。首先, 使用 KWConv 卷积, 减少网络模型的计算量; 其次, 使用 BIFPN 增强上下文信息的捕获, 并减少模型的参数数量; 然后, 使用感受野阻滞模块替换空间金字塔池化快速模块, 扩大模型的感受野; 最后, 引入大可分离核注意力 (LSKAttention) 模块, 增强网络对目标特征的理解和捕获能力。通过实验, 在 NEU-DET 数据集上 LSKA-YOLOv8 的平均检测精度提高了 4.4%, 模型的参数量和计算量分别下降了 26.7%和 50%。

1.3 研究现状分析

在工业钢材表面缺陷检测算法的研究中, 不同场景下的应用需求多种多样, 因此使用的算法也在不断改进。从近几年来国内外目标检测算法的研究现状中可以看出, 相关的算法研究主要表现在两方面: 提升算法的平均检测精度和算法在设备部署前进行模型轻量化处理。虽然各国学者们在对基于深度学习的工业钢材表面缺陷检测算法的改进中取得了一些较好的成果, 但是在基于深度学习的工业钢材表面缺陷检测中一直存在着一些关键性问题:

(1) 工业钢材表面缺陷种类复杂, 尺度变化大

在通过大量的调查研究后发现, 算法模型对于某种特定单一的缺陷进行检测时效果良好, 但是当检测种类复杂且尺度变化大的多种类型缺陷时, 算法模型并不能取得较好的检测结果。然而在工业生产中, 钢材表面缺陷包含夹杂、裂纹、斑块等多种类型缺陷, 而且不同类型缺陷的大小和形状千奇百怪, 这就导致不同缺陷之间差异不够明显, 相同缺陷之间的差异又比较大, 导致算法模型在提取各种缺陷特征时会出现混淆, 从而导致在缺陷分类时出现误检、漏检

等情况，使算法模型对各类缺陷的平均检测精度较低。

（2）缺陷的特征提取受到图片背景的干扰

许多钢材表面的缺陷与钢材的颜色相同，并且在拍摄照片时，图片的质量也会受到光照强度、相机像素分辨率的高低等一些客观因素影响，使得缺陷的特征信息不够丰富，这就导致算法在提取特征信息时，有些缺陷的特征信息和图片的背景特征信息不容易区分，从而导致缺陷的平均检测精度较低。

（3）检测精度与模型轻量化难以同时改善。

目前，许多研究者在对算法的检测精度进行提高时，都会使用复杂的结构对网络进行优化。虽然算法的检测精度得到了提高，但是算法模型也会变得更加庞大，算法在检测的过程中会需要更长时间来响应，这会导致算法的实时性变差。然而应用在工业场景中的目标检测算法，需要考虑算法的实时性和嵌入式设备的硬件计算资源问题，而与算法实时性相关的是算法模型的大小，学术界中通常使用计算量和参数量作为评价模型大小的主要指标。计算量是模型所需的计算次数，反映了模型对硬件计算单元的需求；参数量是模型中的参数总和，跟模型在磁盘中所需的空间大小直接相关。当研究者们使用复杂结构优化算法时就会增加算法的计算量和参数量，因此算法模型在设备上部署时，需要同时考虑模型的大小和检测精度。

1.4 主要研究内容与章节安排

本研究以提升工业钢材表面缺陷检测精度为目标，以 YOLOv8n 算法模型为基础，实验中使用 NEU-DET 数据集，开展工业钢材表面缺陷检测算法研究。针对工业钢材表面缺陷检测中，存在钢材表面缺陷种类复杂和尺度变化大、缺陷的特征提取受到图片背景的干扰影响检测精度的问题，提出了检测能力更强的 YOLOv8n-Improved 算法模型。以此为基础，提出轻量化 YOLOv8n-Simple 算法模型，可以运用在计算资源有限的工业嵌入式设备和边缘设备中，进行工业钢材表面缺陷的实时检测，并结合 PySide6 设计一个工业钢材表面缺陷检测系统，实现对工业钢材表面缺陷的检测任务。论文具体内容如下：

第一章：绪论。介绍工业钢材表面缺陷检测的背景及意义，引出开展基于深度学习的工业钢材表面缺陷检测算法的重要性；分析了基于深度学习的目标检测算法的国内外研究现状，以及这些目标检测算法在工业钢材表面缺陷检测

中的研究现状；最后，总结了现阶段使用深度学习的目标检测算法，在工业钢材表面缺陷检测时始终存在的一些问题。

第二章：深度学习目标检测理论。详细介绍了深度学习主流目标检测算法中卷积神经网络模型结构；然后，分别介绍了深度学习中，两阶段目标检测算法 R-CNN 系列和单阶段目标检测算法 YOLO 系列，选出目前主流的 YOLOv8n 网络模型，依据影响工业钢材表面缺陷检测精度的问题，对算法进行改进，并介绍了一些评价算法优劣的通用指标。

第三章：YOLOv8n 模型改进。针对影响检测精度的问题，从各个方面对 YOLOv8n 模型进行改进。首先对损失函数进行介绍，选择可以更好降低预测结果与真实标签之间差异的损失函数，保证模型的检测性能；然后针对缺陷的特征提取受到图片背景的干扰问题，对注意力机制模块进行介绍，对比实验不同注意力机制在网络不同位置，对缺陷特征提取的影响程度；针对工业钢材表面缺陷种类复杂和尺度变化大的问题，对可变形卷积进行介绍，将可变形卷积 v2 与 C2f 模块结合构建新的 C2f_DCNv2 模块，提高网络模型对缺陷特征的学习能力。

第四章：基于改进 YOLOv8n 的工业钢材表面缺陷检测算法。将各种改进的最佳方法融入于 YOLOv8n 网络模型中，形成最终的改进算法模型；随后，对 NEU-DET 数据集进行介绍；然后，介绍了检测的实验环境和使用的参数；最后，验证对比改进算法的实验结果。

第五章：工业钢材表面缺陷检测算法的轻量化及系统设计。在实际应用中，工业钢材表面缺陷检测算法在运用到嵌入式设备时，需要考虑实时性、资源限制、成本效益和移动性等方面。分别介绍三种轻量化卷积模块并对算法进行改进，对比分析轻量化后的模型检测结果，将最佳的轻量化算法与检测需求相结合，使用 PySide6 与改进的算法搭建一个工业钢材表面缺陷检测系统，可以对工业钢材表面缺陷图片进行检测。通过检测钢材表面缺陷图片，对系统的检测功能进行测试。

第六章：总结与展望。对本文所研究的工作内容进行总结，然后探讨工作中的一些不足之处并对未来进行展望。

本文的主要流程结构如图 1-2 所示。

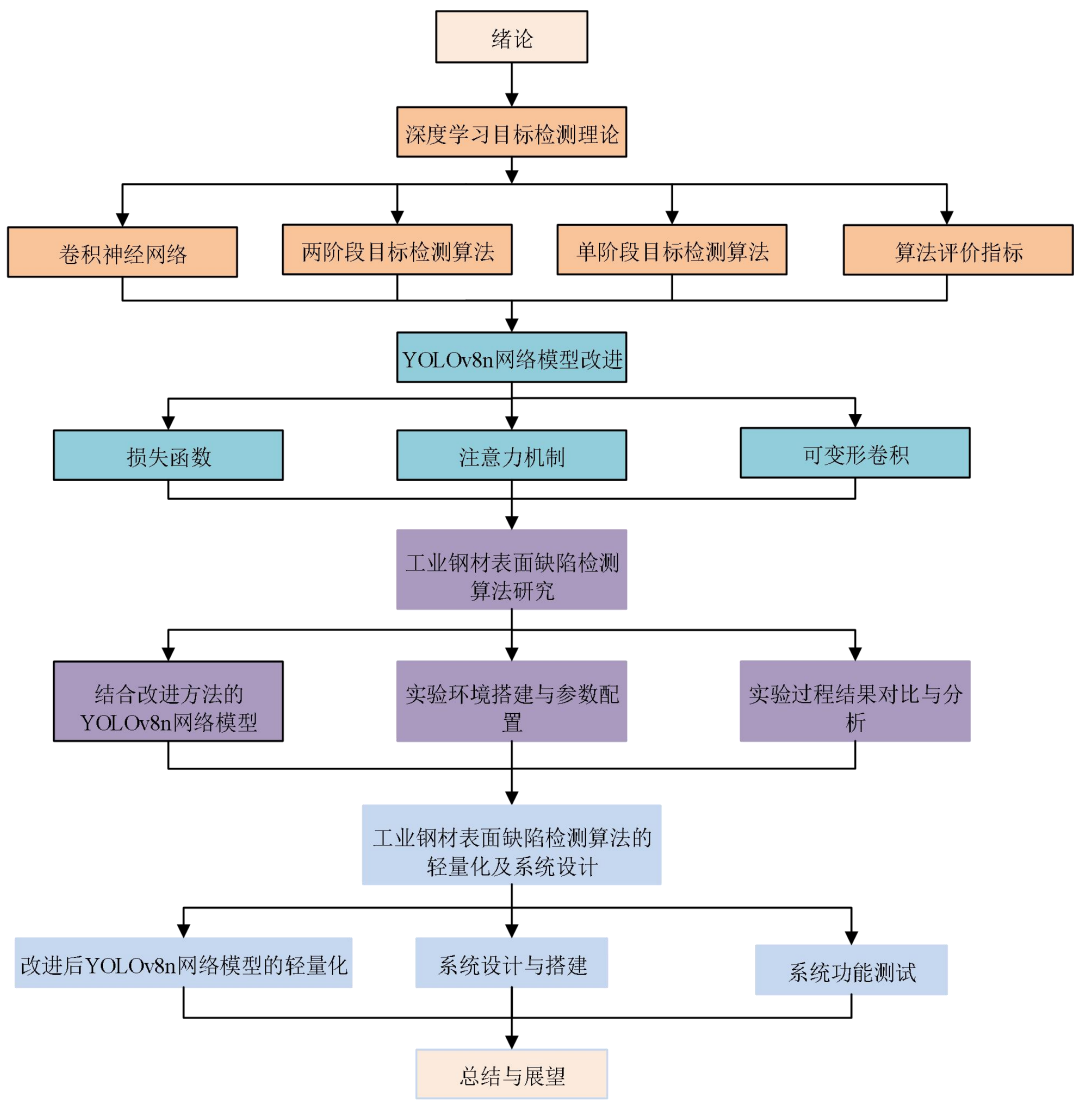


图 1-2 论文主要结构流程

第 2 章 深度学习目标检测理论

深度学习中的目标检测技术在计算机视觉领域的研究中扮演着重要角色，随着深度学习的发展，特别是卷积神经网络的出现，目标检测进入了新时代。卷积神经网络在图形的特征提取上展现了强大的能力，于是研究人员把卷积神经网络应用在目标检测领域后，相继出现了一系列优秀的目标检测算法，这些算法在对目标的平均检测精度以及检测速度上有了很大的提升，并应用在了对车辆和人员的实时监控、工业生产线上对瑕疵产品的检测等。在本章节中，将深入介绍卷积神经网络以及深度学习中的目标检测算法。

2.1 卷积神经网络

1998 年，LeCun 等人提出了一个非常成功的卷积神经网络模型 LeNet-5^[42]，开启了利用卷积神经网络自动提取图像特征的先河。在 20 世纪 90 年代，许多美国银行都使用基于 LeNet-5 的手写数字识别系统来识别支票上面的手写数字。然而，直到 2012 年的 ImageNet 图像分类竞赛中，Krizhevsky 等人提出了 AlexNet 网络模型并获取了比赛的冠军，当提出的卷积神经网络在图像的目标特征提取上展现出了强大的能力后，卷积神经网络模型才被外界广泛关注。如图 2-1 为卷积神经网络的整体架构，其中包括卷积层、池化层和全连接层。

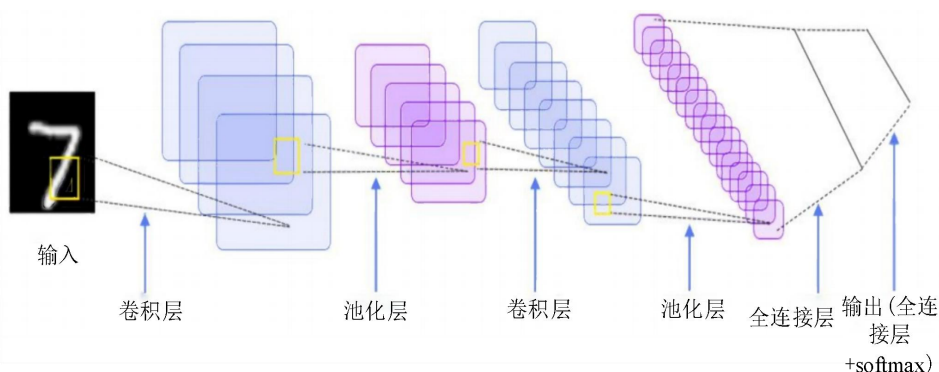


图 2-1 卷积神经网络整体架构

在图片输入后，卷积层对图片进行一些特征的提取，从而得到特征图，池化层会在获得卷积层的特征图后，进行特征选择和降维，最后全连接层对特征进行整合分类和预测，为后面的输出做准备。

2.1.1 卷积层

卷积层是卷积神经网络的重要组成部分，主要通过卷积运算来提取输入数据的特征。卷积操作可以用图 2-2 表示，卷积运算是输入数据与卷积核之间的乘积运算，当输入数据经过卷积层时，卷积核会在输入的 4×4 数据矩阵上以一个特定的步幅前进。在前进过程中，卷积核会与部分输入的 4×4 数据矩阵进行相乘，相乘后的数据再进行相加得到输出特征图的一个像素值。当卷积核在输入的 4×4 数据矩阵遍历之后，就会得到整张输出特征图。

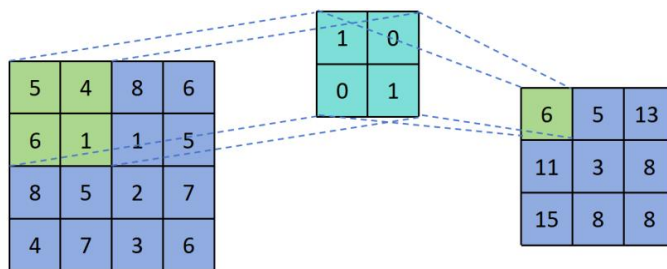


图 2-2 卷积操作示意图

卷积运算的公式可以表示为：

$$C(i, j) = \sum_m \sum_n I(i + m, j + n) \cdot K(m, n) \quad (2-1)$$

式中， C 是输出特征图， I 是输入图像， (i, j) 是输出特征图坐标， K 是卷积核， (m, n) 是卷积核坐标。

卷积核可以提取一个局部区域的特征，而在卷积层中存在多个不同的卷积核，这样卷积层就可以提取更多的特征信息。通常，在输入端会使用一些较小的卷积核来提取图像边缘以及图像颜色等信息，在输出端使用较大的卷积核对更深层的语义信息进行提取。

2.1.2 池化层

卷积层的后面通常使用池化层，对卷积层输出的特征图进行下采样和保持特征不变性。在下采样的过程中会对特征图进行降维，减少特征图的数量和计算量，同时也可以预防模型过拟合，增强模型的泛化能力。在特征不变性中，可以保留特征的关键信息，使模型依旧能够识别图像的微小位移、旋转和缩放等变换。

池化操作通常分为最大池化和平均池化。在最大池化操作过程中，特征图

分为大小相同的数块区域，每块区域中都含有前面卷积层生成的像素。在这些像素中选择数值最大的一个作为新生成特征图的一个像素，其他的舍弃，将每块区域都取最大值后就可以生成新的特征图，如图 2-3 所示。这样可以减少卷积过程中噪声对特征图的影响，使图像中最显著的特征得到保留，有助于提取图像中的纹理和形状。

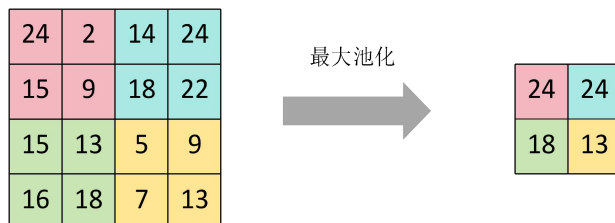


图 2-3 最大池化操作

平均池化操作也是先将特征图分为大小相同的数块区域，每块区域中都含有前面卷积层生成的像素，将每块区域中的像素值相加后取平均值得到新的像素，完成每块区域的运算后，就可以得到平均池化后新的特征图，如图 2-4 所示。平均池化可以提取该区域的主要特征，减少噪声的干扰，同时使模型依旧可以识别输入数据的小范围平移。

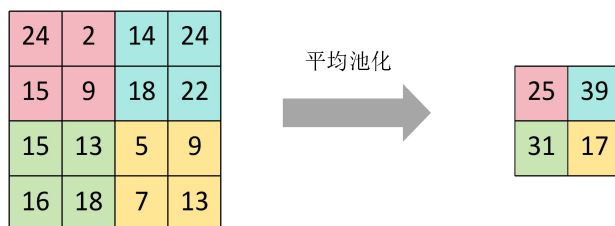


图 2-4 平均池化操作

2.1.3 全连接层

全连接层位于卷积和池化之后，一般在卷积神经网络的最后几层。在全连接层中，每一层神经元与下一层的神经元全部相连。在进行全连接时，会将前面得到的二维特征图转换为一维的向量，这些一维向量中含有各个特征信息，整合起来后就含有了全局信息，最后通过学习的权重和偏置参数实现对输入目标特征的学习和分类，全连接层示意图如图 2-5 所示。

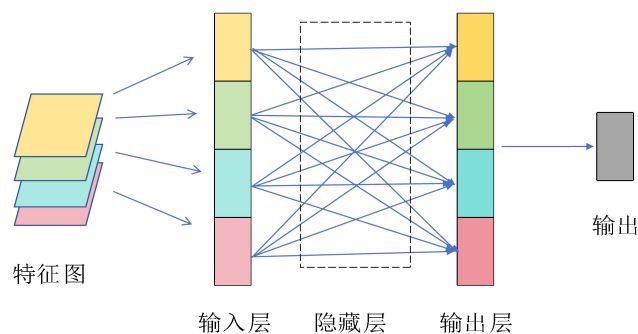


图 2-5 全连接层结构

2.1.4 激活函数

卷积运算过后得到的特征数据是线性的，而现实中许多图像信息是非线性的，如果只进行一些线性变换和运算，那么卷积神经网络只能简单的学习一些图像的基本特征信息，无法得到更全面的特征信息。因此，可以利用激活函数的非线性特性，把线性特征数据转化为非线性的数据，这样网络就可以学习到图像的一些非线性的特征信息，从而提高模型的学习和表达能力。在激活函数中，常用的饱和激活函数有 Sigmoid^[43]和 Tanh^[44]等。当输入的值达到某一界限后，在饱和激活函数中，其后所有数值在函数的导数输出中都趋于零。在激活函数中，常用的非饱和激活函数有 ReLU^[45]和 Leaky ReLU 等，其函数的导数不会趋于零。如表 2-1 所示为它们的函数表达式和基本形状。

表 2-1 激活函数表达式和基本形状

激活函数	表达式	图形
Sigmoid	$f(x) = \frac{1}{1 + e^{-x}}$	
Tanh	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	
ReLU	$f(x) = \max\{0, x\}$	
Leaky ReLU	$f(x) = \begin{cases} x & x \geq 0 \\ ax & x < 0 \end{cases}$	

2.2 两阶段目标检测算法

两阶段目标检测算法分为两个处理过程，第一个过程会使用候选框在图像中框选一系列需要检测的目标，第二个过程对产生的候选框位置坐标进行修正，优化目标在候选框中的位置并且对候选框中的目标进行分类，识别出要检测目标的类别。常见的两阶段目标检测算法有：R-CNN^[46]、Fast R-CNN^[47]和 Faster R-CNN^[48] 等。

2014 年，Rshickr 等人在继承传统目标检测的思想提出了 R-CNN 算法。作为 R-CNN 系列的第一代算法，R-CNN 算法过程如图 2-6 所示。首先，在输入的图片中生成一系列的候选框；然后，对候选框区域使用卷积来提取特征图；最后，通过线性支持向量机对特征进行分类，确定候选框中目标所属的类别。同时，使用线性回归调整候选框的位置，更准确的定位目标。候选框和卷积网络的结合，使模型的准确率大大提高。但是在生成候选框阶段，通常会产生 2000 个候选框，这些候选框都会进行特征提取。因此，网络中的计算量会增加，从而造成资源的浪费和时间的大量耗费，使模型检测速率变慢。

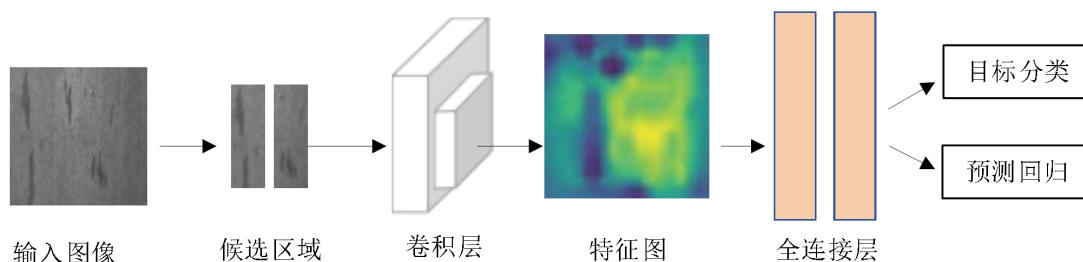


图 2-6 R-CNN 算法结构图

2015 年，Girshickr 提出了 Fast R-CNN，借鉴 SPP-Net 模型，将处于全连接层的空间金字塔池化层（SPP）简化为 ROI 池化层，ROI 池化层将不同大小的目标区域转化为固定尺寸，方便后续的全连接层处理，使网络可以更好的对目标进行分类和定位。同时，以 Soft-max 代替 SVM 进行分类回归，将所有特征向量存储在 GPU 中，这在很大程度上节省了磁盘的空间。Fast R-CNN 算法过程如图 2-7 所示。

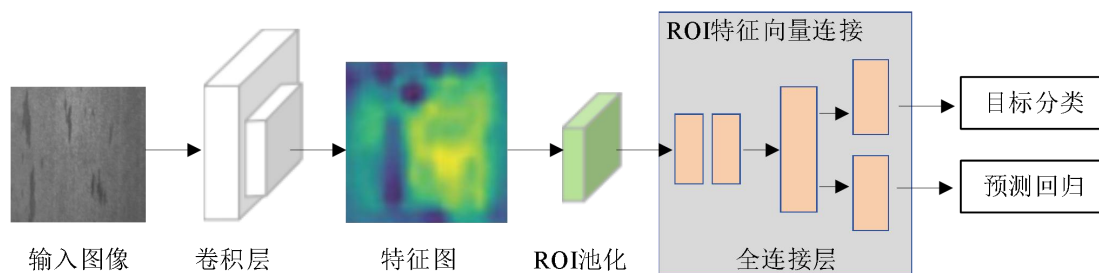


图 2-7 Fast R-CNN 算法结构图

2015 年，Rens 等人在 Fast R-CNN 的基础上提出了 Faster R-CNN，在特征提取时使用区域预测网络（Region Proposal Network，RPN），不再使用候选框进行特征提取，其 Faster R-CNN 算法过程如图 2-8 所示。首先，将输入图像通过特征提取网络进行特征提取；然后，将提取得到的特征图通过区域预测网络生成候选框，并判断候选框中是否含有目标；随后，对候选框进行修正，ROI 池化层收集区域预测网络产生的候选框坐标，并依据特征提取网络中提取出的位置进行目标提取，最后在全连接层进行分类和回归。

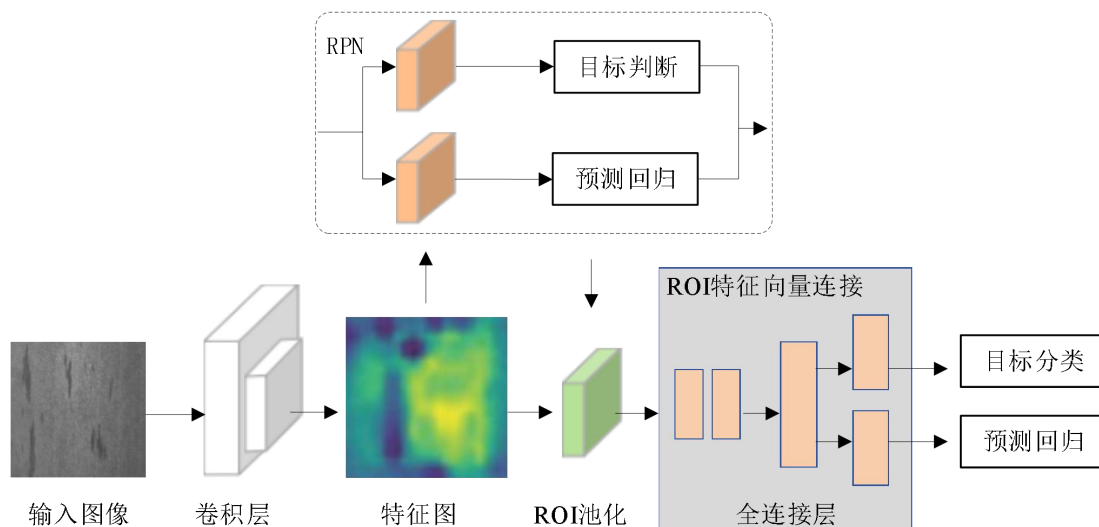


图 2-8 Faster R-CNN 算法结构图

2.3 单阶段目标检测算法

在单阶段目标检测算法中，舍去了生成候选框的过程，在图像上直接进行目标的定位和分类，这极大程度上降低了模型网络的复杂结构，为设备节省了许多计算资源，因此算法的检测速度非常快。如今十分热门的单阶段目标检测算法 YOLO 系列经过多个版本的更新迭代，检测精度逐渐提高。如今，许多版本被用在工业场景的检测中。

2015 年, Redmon 等人^[49]提出 YOLO 算法。YOLO 算法的检测过程结构如图 2-9 所示。它通过使用卷积神经网络, 在整张图片上提取特征, 然后直接进行分类与回归, 这大大提高了目标检测的速度。在检测过程中, 图像被分成若干个相等大小的网格, 当目标物体的中心位于网格中时, 这个网格就负责检测此目标, 而且每个网格中会有若干个边界框用来包含检测的目标物体。但是, 当有一个网格中的预测框中有多个物体时, 算法很难进行判断, 不适合小目标检测。

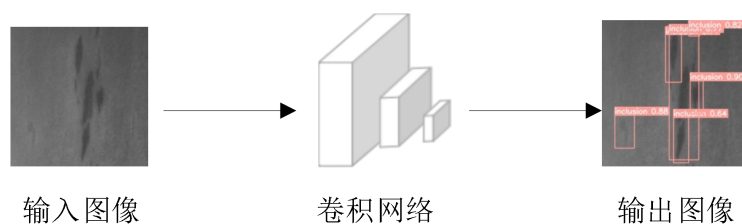


图 2-9 YOLO 算法结构图

2017 年, Redmon 等人^[50]对 YOLO 算法做出了改进, 提出了 YOLOv2。YOLOv2 在训练集中对锚的位置和形状进行聚类, 形成可以更好包括物体大小的边界框, 并且增加了预测框的数量, 使模型对密集小目标的检测有了很大的提升。同时, 在卷积层中使用了批量归一化, 因为在神经网络训练过程中, 每层输入的数据分布会不断变化, 这会导致训练时间增加甚至出现梯度消失等问题。批量归一化就是为了解决这一问题, 它在每一层的激活函数前使一小批数据的均值接近 0, 方差接近 1, 然后进行归一化来改善网络的收敛性, 使网络训练的更快、更稳定。YOLOv2 还取消了卷积网络中的全连接层, 这使图片可以输入更多的尺寸大小。

2018 年, Redmon 等人^[51]在 YOLOv2 的基础上提出了 YOLOv3。在 YOLOv3 中, 引入了残差网络结构, 增加了多尺度特征图的融合, 这有效提高了算法的检测精度。YOLOv3 算法将骨干架构网络 Darknet-19 改为 Darknet-53, 增加了更多的卷积层用来提取特征, 在卷积的同时引入残差网络结构, 使网络更容易优化目标函数, 提高网络检测的准确率。当网络获得更多的特征图信息后, 网络可以在三个不同尺度的特征图上进行预测, 并且在输出时采用逻辑分类器, 将输出作为多标签来处理。例如, 检测出来的物体既可以是一只小猫, 也可以是一只狸花猫。此时, YOLO 算法的网络结构开始被分为三部分: 特征提取层 (Backbone), 特征融合网络 (Neck), 和输出端 (Head)。YOLOv3 网

络结构如图 2-10 所示。

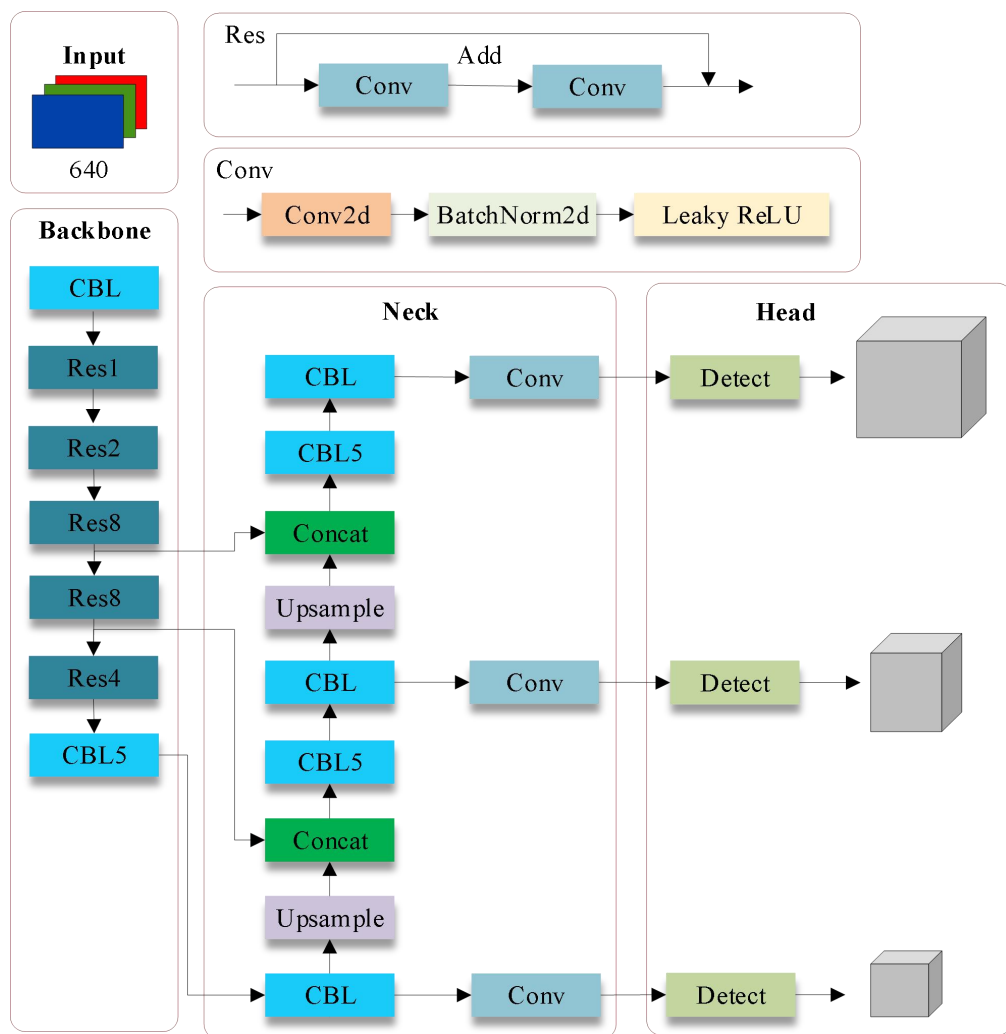


图 2-10 YOLOv3 算法结构图

2020 年，Bochkovskiy 等人对 YOLOv3 做了优化，提出了 YOLOv4。YOLOv4 算法的本质是将当时的多种先进方法融合到 YOLOv3 算法中，其网络结构如图 2-11 所示。在图像增强方面，使用了 Mosaic 数据增强方法，Mosaic 数据增强通过随机缩放、随机裁剪、随机翻转和旋转的方式将 4 张图片拼接在一起，这种增强方法极大的丰富了训练时训练集的数量。在网络结构方面，添加了 CSP 结构和 PAN 结构，并利用 CSPNet 构建了 CSPDarknet-53 网络。在卷积网络中使用 Mish 激活函数，进一步提升了网络的特征提取能力。在损失函数方面，引入了优化能力更强的 CIoU 损失函数，在加快预测框回归速度的同时提高了检测精度。

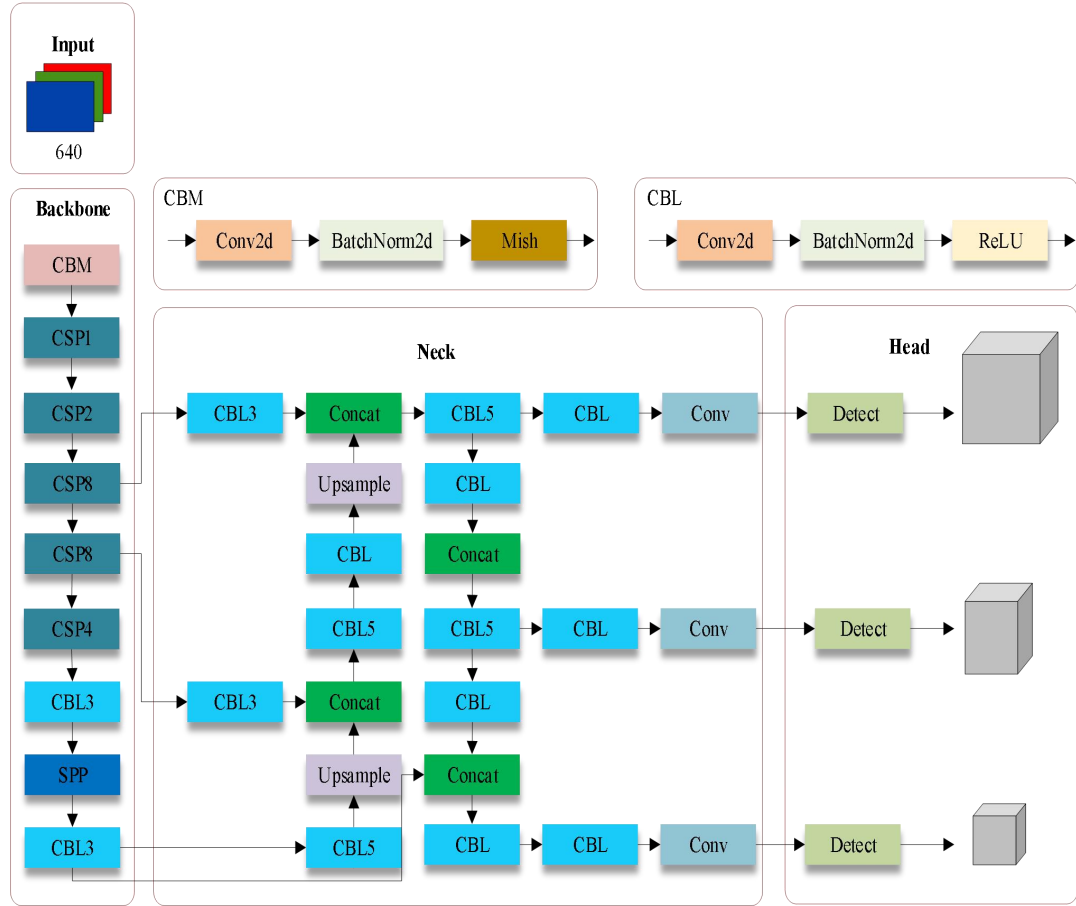


图 2-11 YOLOv4 算法结构图

2020 年，Glenn 等人^[52]公开了 YOLOv5 代码，在 YOLOv4 的基础上，增加了自适应锚框、加权非极大值抑制等一些训练技巧，还添加了 Focus 结构优化参数量用来提升检测速度。根据模型的卷积深度，YOLOv5 推出了 s、m、l、x 四种不同大小的网络模型，可以根据不同的任务选择不同的模型。因此，YOLOv5 也是目前业界公认综合表现较好的目标检测算法之一，其结构如图 2-12 所示。

2023 年，YOLOv5 的发布者 ultralytics 将 YOLOv8 代码开源，YOLOv8 在 YOLOv5 的模型基础上全面提升改进。此次算法将 YOLOv5 的 C3 结构换成了梯度流更丰富的 C2f 结构，在 head 部分换成了目前主流的解耦头结构。改进后的算法在准确率方面比 YOLOv5 有了一些提升。同时，作者也在不断更新 YOLOv8 版本，本文使用的是 YOLOv8 的 1.0 版本，其网络结构如图 2-13 所示。

在 Input 输入端，算法会对数据进行预处理，分别使用 Mosaic、RandomAffine 等数据增强技术、锚框自适应计算和图片自适应缩放操作。在主干网络 Backbone 中，首先使用两次卷积进行下采样，然后在 C2f 模块中进行残

差特征学习。输入特征在 C2f 模块主要分为两部分：一部分进行少量卷积操作，另一部分经过多个卷积层进行堆叠。在 Backbone 的末端是 SPPF 模块，该模块主要进行融合操作，将上层的输入通过并行方式经过多个不同的 MaxPool，这在一定程度上可以解决目标尺度不同的问题。在 Neck 中，主要是将特征进行融合。它结合了 FPN（Feature Pyramid Network）和 PANet（Path Aggregation Network）的思想，将主干网络中的浅层信息和颈部中的深层信息进行融合，让模型可以学习到更多目标的信息。Head 分为 3 个检测接头，检测头为解耦头，将分类和回归任务进行解耦，并采用基于中心检测算法的无锚框（Anchor-Free）检测方式对物体中心和边缘信息进行检测，有效减小了锚框的手动设计，提高目标检测的准确性。

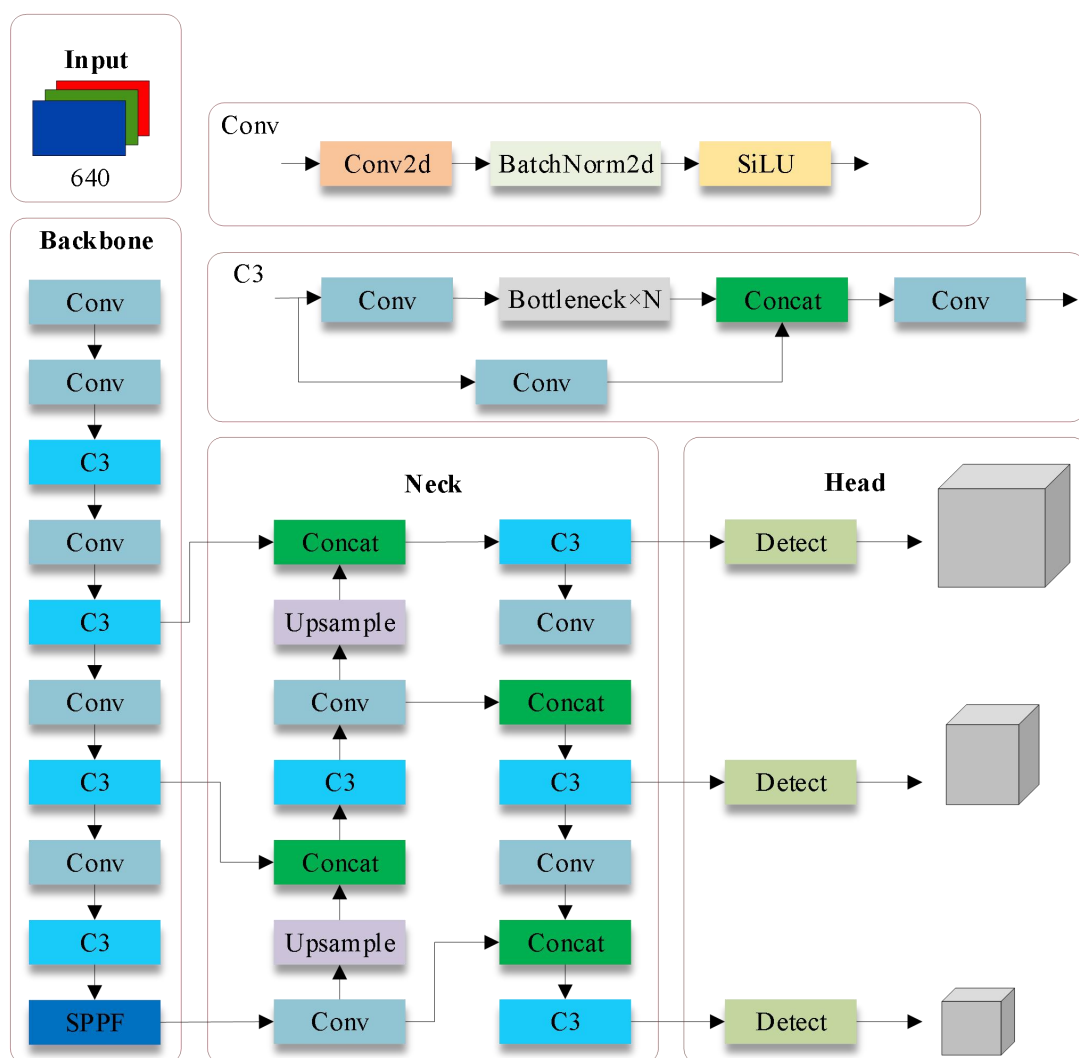


图 2-12 YOLOv5 算法结构图

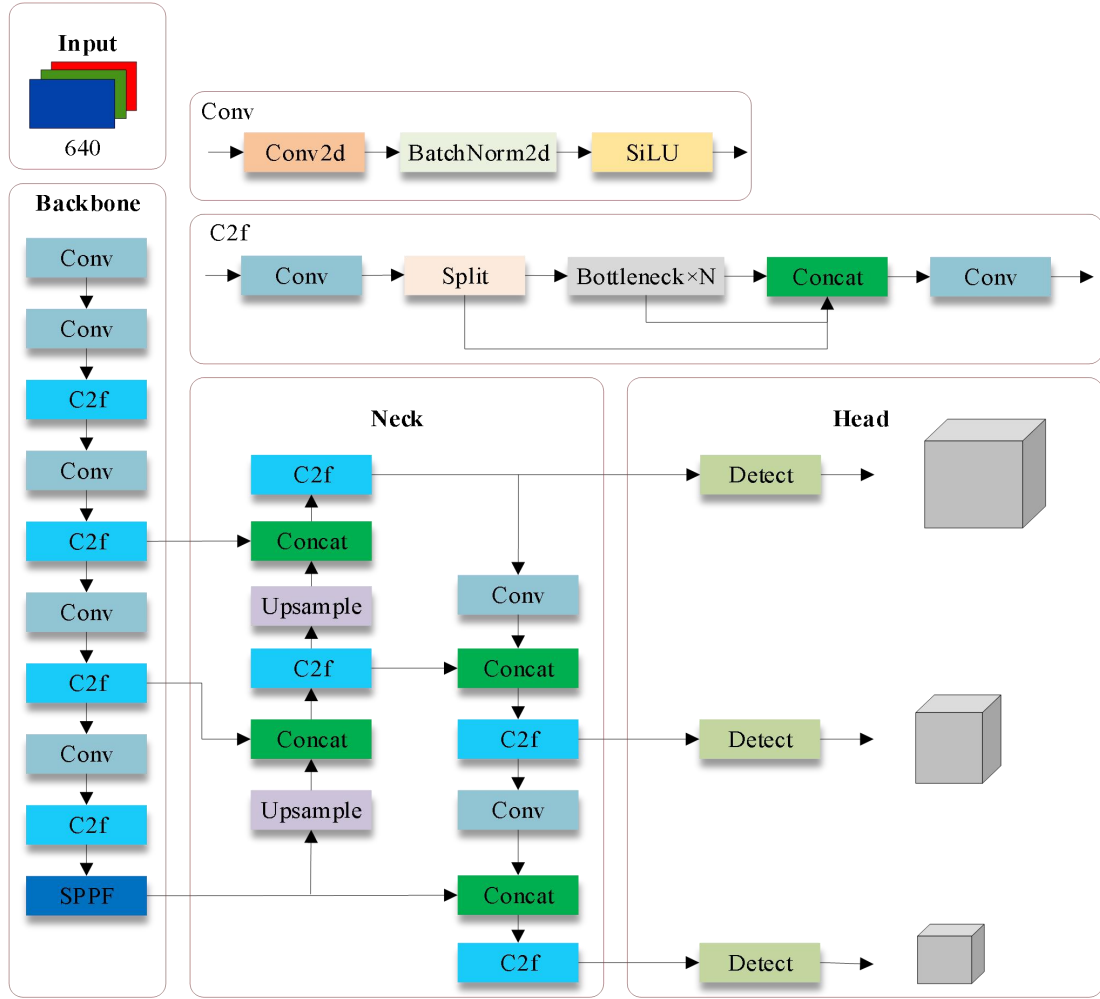


图 2-13 YOLOv8 算法结构图

2.4 评价指标

目标检测算法中模型众多，若没有评价标准会导致无法客观地判断哪个模型的检测效果更好、速度更快，以及在不同场景下哪个模型的适用性更好。因此，需要借助一些统一评价标准，直观地反应模型的效果，通过分析评价标准的各项指标来明确模型的优势和不足。本研究分别使用了如下几个指标对模型的性能进行评价：

(1) 精准度 (Precision)：预测为正样本中，预测正确的个数与预测为正样本个数的比例，其公式定义为：

$$Precision = \frac{TP}{TP + FP} \quad (2-2)$$

式中， TP 表示标记为正样本在预测时也为正样本的数量， FP 表示标记为负样本但预测时为正样本的数量。精准度的值域为(0,1)，当精准度的值越大，模型

对正样本的识别率越高，可以用来衡量模型对目标识别的准确性。

(2) 召回率 (Recall): 预测为正样本中预测正确的个数与标记为正样本个数的比例，其公式定义为：

$$Recall = \frac{TP}{TP + FN} \quad (2-3)$$

式中， FN 表示将标记为正样本但预测为负样本的个数。召回率的值域为(0,1)，当召回率的值越大，模型对所有目标的查全率越高，可以用来评价模型对查找目标的全面性和覆盖率。

(3) 平均精度均值 (mean Average Precision, mAP): 通过对多个类别的平均精度取平均值得到平均精度均值。平均精度均值是精准度和召回率曲线下的面积，其公式定义分别为：

$$AP = \int_0^1 P(R) dR \quad (2-4)$$

$$mAP = \frac{\sum_{i=0}^n AP(i)}{n} \quad (2-5)$$

mAP 是综合评价模型在不同类别上检测精度的指标，在目标检测中一般使用 mAP 来评价模型识别物体的效果。

(4) 参数量(Params): 模型中需要学习的参数总量，是整个网络中所有权重和偏置参数数量的总和。在一定程度上参数量反应了模型的复杂度，参数量越多，模型的复杂程度也越高。

(5) 计算量(FLOPs): 计算量是网络模型中涉及到的乘法和加法操作的数量。计算量的大小通常会影响模型的训练和应用速度，计算量大的模型通常需要设备提供更多的计算资源。

2.5 本章小结

本章节在开头引入了深度学习中的一种重要神经网络架构卷积神经网络，并围绕卷积神经网络的结构组成进行详细介绍。在深度学习的目标检测中，分别介绍了使用卷积神经网络架构的两阶段目标检测算法和单阶段目标检测算法的发展和网络结构，引出本文要使用的 YOLOv8 算法。最后对评价算法优劣的通用指标进行了介绍，方便后面章节对改进后的算法进行评价。

第 3 章 YOLOv8n 网络模型改进

使用 YOLOv8 算法对工业钢材表面缺陷检测时，会遇到第一章提到的基于深度学习的工业钢材表面缺陷检测中存在的一些关键性问题，导致模型对工业钢材表面缺陷检测的精度低。因此，本章以 YOLOv8n 模型为基础，针对遇到的关键性问题分别对模型进行改进。首先，选择适合工业钢材表面缺陷检测的损失函数，减少模型在预测结果与真实标签之间的差异，同时使模型在检测缺陷时可以快速达到收敛状态；其次，针对缺陷的特征提取受到图片背景干扰的问题，提出添加注意力机制的措施，对比分析不同注意力在网络不同位置的提升效果；最后，针对钢材表面缺陷种类复杂，尺度变化大的问题，提出引用可变形卷积与模型的模块结合构建新模块的措施。通过这些改进，改善检测过程中遇到的关键性问题，提高钢材表面缺陷平均检测精度。

3.1 选取损失函数

在深度学习的目标检测中，通常也使用损失函数的值来评价模型的性能。当损失函数的值越小，那么模型的预测结果与真实标签之间的差异越小，模型预测的结果就越准确。因此，在模型训练过程中，我们需要将损失函数值朝着更小的方向去优化，使模型的目标检测性能更好。损失函数在很大程度上决定了模型的性能，因此在算法改进的初始阶段就需要选择合适的损失函数来改善训练过程中的回归情况，得到更优的模型^[53]。常见的损失函数有 GIoU、DIOU 和 CIOU，本章在检测精度和损失值上进行 GIoU、DIOU 和 CIOU 损失函数的比较，选取更适合模型检测的损失函数。

3.1.1 GIoU 损失函数

GIoU 损失函数在 IoU 损失函数的基础上进行改进，加入了预测框与真实框的最小闭包区域，有效解决了预测框和真实框没有重叠时，IoU 损失函数无法提供足够的梯度信息进行优化模型的问题。通常用于处理遮挡、相邻目标等复杂场景下的目标检测。GIoU 损失函数的计算公式可以表示为：

$$L_{GloU} = 1 - GloU = 1 - IoU + |(C - (A \cup B)) / C| \quad (3-1)$$

式中， IoU 是预测框与真实框的交集和并集的比值， A 是预测框， B 是真实框， C 是他们的最小闭包区域。

通过公式可以分析出， $GIoU$ 可以同时关注重叠区域和一些非重合区域，可以比较好的反映两个框在闭包区域中的相交情况。

3.1.2 DIoU 损失函数

$DIoU$ 损失函数在 IoU 损失函数的基础上多考虑了预测框和真实框中心点之间的距离，多用于动态目标检测场景。当目标处于非重叠或者小重叠时， $DIoU$ 损失函数可以快速收敛。 $DIoU$ 损失函数的计算公式可以表示为：

$$L_{DIoU} = 1 - DIoU = 1 - IoU + (\rho^2(d, d^{gt})) / k^2 \quad (3-2)$$

式中， ρ 是中心点之间的距离， d 是预测框的中心点， d^{gt} 是真实框的中心点， k 是包含两个框的最小对角线长度。

通过公式可以分析出，当两框距离较远时，利用两距离的比值，可以有效优化 Loss 值较大的问题。

3.1.3 CIoU 损失函数

$CIoU$ 损失函数在 $DIoU$ 的基础上又考虑了宽高比的因素，多用于小目标和不规则目标的检测中。 $CIoU$ 损失函数可以帮助模型更好地学习框的定位。 $CIoU$ 损失函数的计算公式可以表示为：

$$L_{CIoU} = 1 - CIOU = 1 - DIoU + (\alpha v / (1 - IoU) + v) \quad (3-3)$$

$$\alpha = v / (1 - IoU + v) \quad (3-4)$$

$$v = 4(\arctan(w^{gt}/h^{gt}) - \arctan(w/h))^2 / \pi^2 \quad (3-5)$$

式中， h 和 w 分别表示预测框的高和宽， h^{gt} 和 w^{gt} 分别表示真实框的高和宽。

通过公式可以分析出， $CIoU$ 损失函数考虑到了边界框回归的重叠面积、中心点距离及长宽比。

3.1.4 实验结果对比与分析

将三种损失函数依次加入 YOLOv8n 网络模型中，然后对工业钢材表面缺

陷进行检测，其检测结果如表 3-1 所示。通过分析表中数据可知，CIoU 损失函数在精准度上比其它两种损失函数要高，达到了 75.1%，而且 mAP 值也比表中其它两种损失函数的值要高，达到了 77.3%。这表明模型在使用 CIoU 损失函数时，在工业钢材表面缺陷上检测的结果更好。

表 3-1 三种损失函数结果

损失函数	P/%	R/%	mAP/%
GIoU	73.3	70.8	76.7
DIoU	73.9	71.4	77.0
CIoU	75.1	70.1	77.3

如图 3-1 所示，为三种损失函数加入到 YOLOv8n 模型中的损失函数曲线对比图。从曲线图中可以看出，CIoU 的损失函数曲线在 25Epochs 之后与其它两个损失函数的曲线开始分开，并且下降速度开始加快。在 300Epochs 左右时，最终的损失值比其它两个损失函数值要低。由损失函数值的评价标准可知，使用 CIoU 损失函数更有助于减少预测结果与真实标签的差异，使模型的目标检测性能更好。

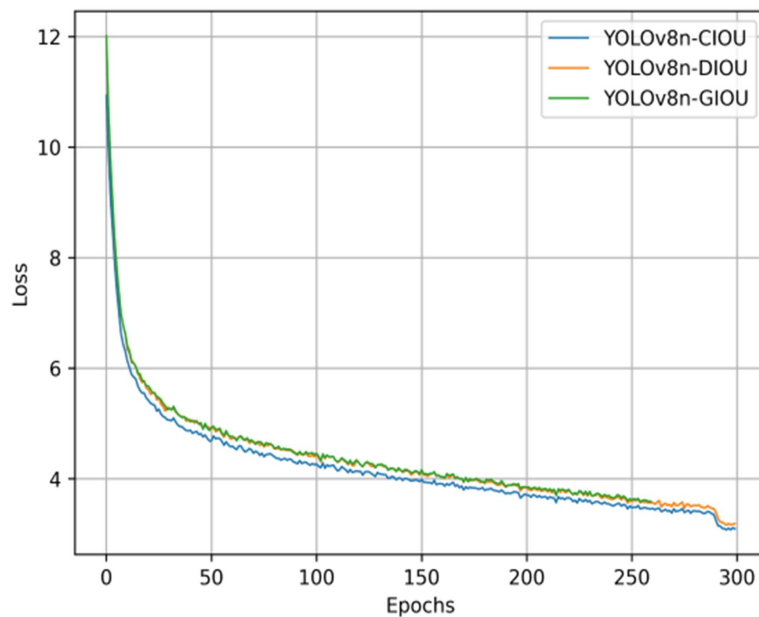


图 3-1 三种损失函数对比曲线图

3.2 引入注意力机制模块

注意力机制模块在深度学习的目标检测中能够增强特征的提取，从而提高检测精度和模型效率。在特征的提取中，它可以从通道、空间等多维度提取和

整合特征，丰富特征内容，使模型可以更好的了解目标特征信息。在提升检测精度方面，能够使模型更准确的确定目标的位置和特征信息，使模型更好的识别目标。在提高模型效率中，有一些注意力机制可以减少计算量并优化资源分配，这样模型可以把大量资源用在特征的学习上，减少不必要的资源占用。因此，在目标检测中，能够帮助模型学习到数据集中更多的目标特征的注意力机制至关重要，而注意力机制按照作用域可以分为空间注意力机制、通道注意力机制和混合注意力机制。本章使用 SE 注意力机制模块对算法模型进行改进，并与最近流行的 CBAM、ECA、SK 和 EMA 注意力机制模块进行对比分析。

3.2.1 CBAM 注意力机制

CBAM (Convolutional Block Attention Module) ^[54]属于混合注意力机制，包含空间注意力和通道注意力，其结构如图 3-2 所示。

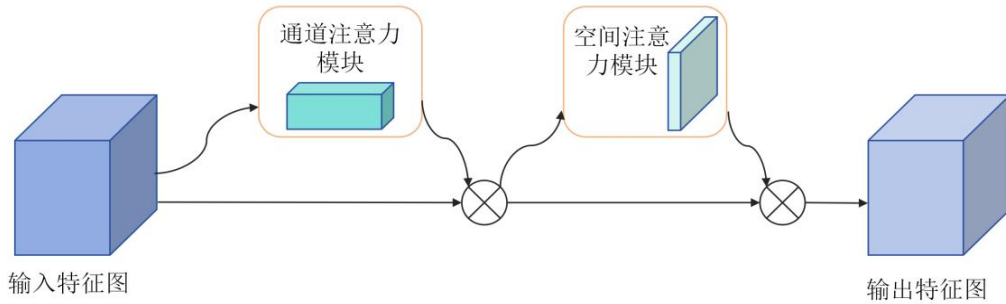


图 3-2 CBAM 注意力机制结构图

在通道注意力模块中，对输入的特征图进行全局平均池化和全局最大池化，得到两个通道特征。将两个特征使用多层感知机处理后相加，再经过 Sigmoid 激活函数，最后与原输入的特征图相乘得到新的特征图。通道注意力模块的过程可以表示为：

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (3-6)$$

式中， $M_c(F)$ 表示一维通道特征图， $AvgPool(F)$ 表示平均池化， $MaxPool(F)$ 表示最大池化， MLP 为多层感知机， σ 为 Sigmoid 激活函数。

在空间注意力模块中，输入为通道注意力模块中得到的特征图，分别进行全局平均池化和全局最大池化。然后拼接池化后的特征图，并与 7×7 的卷积核进行卷积操作。最后，将卷积操作后的特征图经过 Sigmoid 激活函数，与输入的特征图相乘得到最终的输出特征图。其中空间注意力模块的过程可以表示为：

$$M_c(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \quad (3-7)$$

式中， $f^{7 \times 7}$ 表示卷积核为 7×7 的卷积操作。

3.2.2 ECA 注意力机制

ECA (Efficient Channel Attention) [55] 是一种在不增加模型复杂度的情况下使模型性能提升的通道注意力机制，其结构如图 3-3 所示

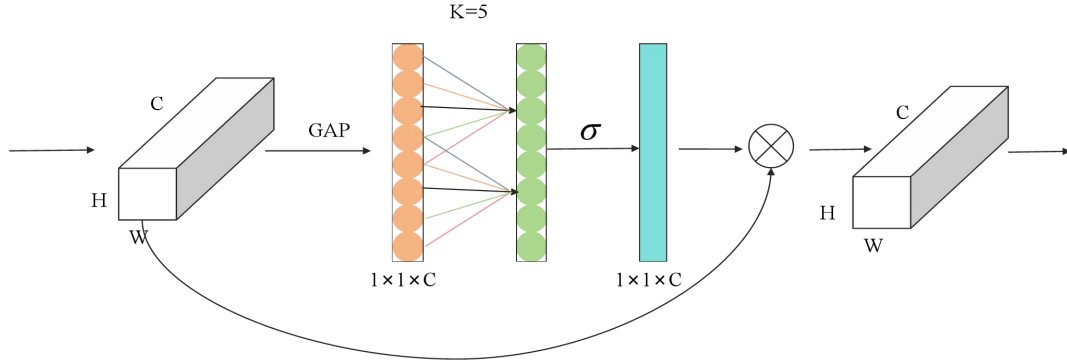


图 3-3 ECA 注意力机制结构图

首先，对特征图进行全局平均池化操作；然后，根据通道数自适应的确定卷积核 K 的大小；随后每个通道都会与 K 个相邻通道进行学习；最后，将通过 Sigmoid 激活函数的权重赋给输出的特征图，得到最终的输出特征图。

其中卷积核 K 的公式可以表示为：

$$C = \phi(K) = 2^{(\gamma \times k - b)} \quad (3-8)$$

$$k = \phi(C) = \lceil (b + \log_2(C)) / \gamma \rceil_{odd} \quad (3-9)$$

式中， C 表示通道维度； $|x|_{odd}$ 表示最接近 x 的奇数； b 取 1， γ 取 2。

3.2.3 SK 注意力机制

SK (Selective Kernel) [56] 是能够让网络输入的特征图，自适应地选择合适大小的卷积核，进行提取特征的通道注意力机制。SK 注意力机制主要由三部分组成，分别为 Split、Fuse 和 Select，其结构如图 3-4 所示。

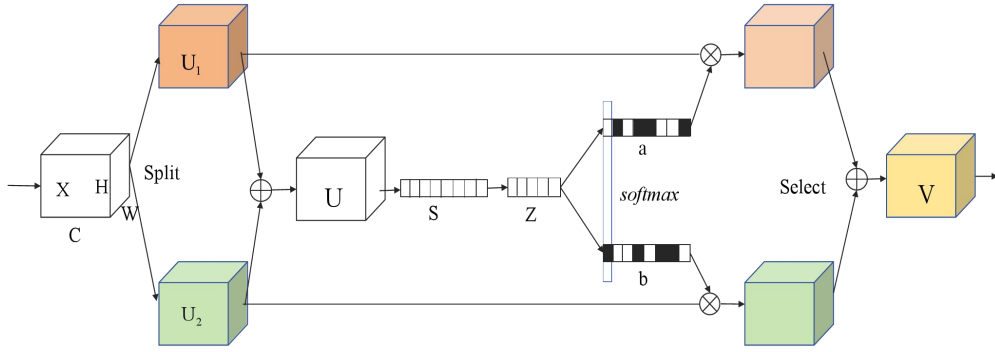


图 3-4 SK 注意力机制结构图

Split 部分会使用不同的卷积核对输入的特征图进行卷积，得到不同感受野的特征图。

Fuse 部分将得到的不同感受野的特征图进行融合，然后进行全局平均池化和全连接层，得到不同卷积核的权重，最后通过 Softmax 函数进行归一化。

Select 部分将得到的权重与输入后得到的不同感受野的特征图进行赋值，赋值后的特征图相加得到最终的输出特征图。

3.2.4 EMA 注意力机制

EMA (Efficient Multi-scale Attention) [57] 是通过对通道和空间两个维度的信息进行加权计算来突出特征信息的混合注意力模块，其结构如图 3-5 所示。

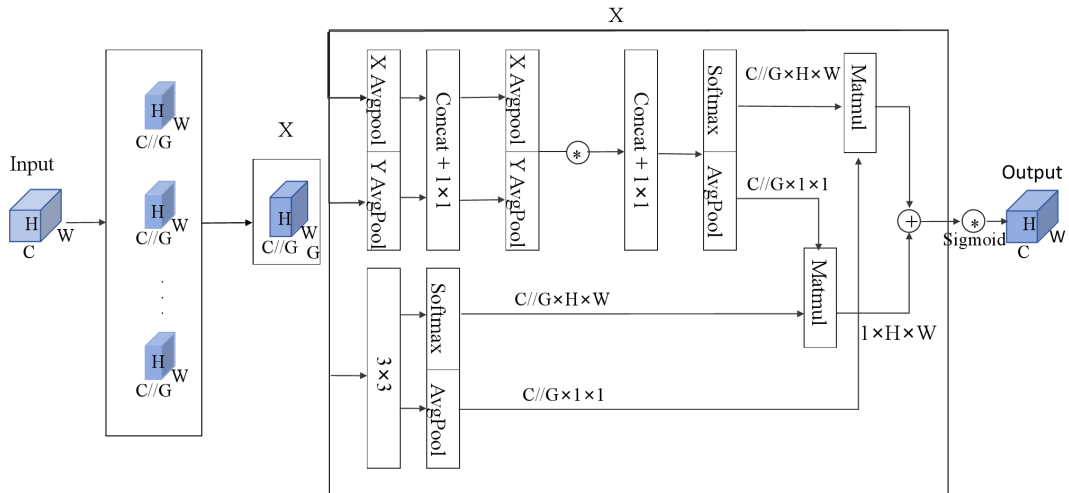


图 3-5 EMA 注意力机制结构图

在 EMA 注意力机制模块中，首先对输入的特征图分为 g 组，然后对一部分进行一维水平和垂直全局池化操作，另一部分用 3×3 的卷积提取特征。在两部分输出特征后，通过 Sigmoid 函数和归一化处理，最终通过跨纬度交互模块进一步聚合两部分的输出特征，以捕捉像素级成对关系。

3.2.5 SE 注意力机制

SE (Squeeze-and-Excitation) 属于通道注意力机制, 在 2017 年由 Hu 等人^[58]提出, 是一种通过对卷积网络中的通道关系, 进行建模来提高网络性能的方法。它通过重新分配获取的权重, 增强有用的特征并抑制不重要的特征, 从而提高获取需要信息的效率。SE 注意力机制模块分为四个步骤: 转换 (Transformation)、压缩 (Squeeze)、激励 (Excitation) 和重新加权 (Scale)。其网络结构如图 3-6 所示。

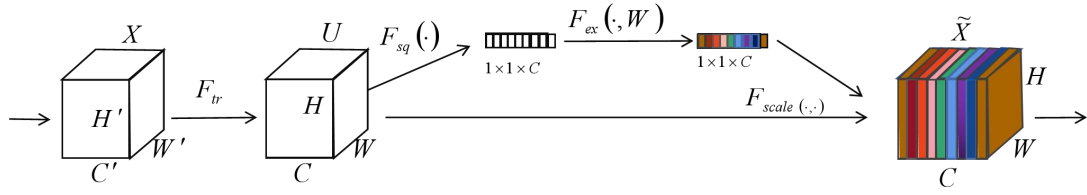


图 3-6 SE 注意力机制结构图

转换步骤对应图中的 F_{tr} 操作, 它将输入特征图 X 通过卷积操作输出特征图 U 。在卷积操作中, 输入特征图 X 的每一层都经过一个 2D 空间核的卷积, 最终得到 C 个输出的特征图 U 。

压缩步骤对应图中的 F_{sq} 操作。通过全局平均池化, 将包含全局信息的 $W \times H \times C$ 特征图直接压缩成一个 $1 \times 1 \times C$ 的特征向量 Z , 通过此操作来生成一个描述整个通道信息的向量, 这样就缓解了通道依赖性的问题。其公式可以表示为:

$$Z_C = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (3-10)$$

激励步骤对应图中的 F_{ex} 操作。通过两层全连接层, 对生成的整个通道信息的 $1 \times 1 \times C$ 向量进行操作。第一层中的全连接层将通道数降低, 第二层中的全连接层将通道数恢复。通过 Sigmoid 激活函数得到每个通道的权重 S , 其公式可以表示为:

$$S = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (3-11)$$

重新加权步骤对应图中的 F_{scale} 操作。将前面得到的注意力权重 S 加权到每个通道的特征上, 得到最终输出特征图 $\tilde{X}$ 。其公式可以表示为:

$$\tilde{X}_c = F_{scale}(u_c, s_c) = s_c u_c \quad (3-12)$$

3.2.6 实验结果对比与分析

注意力机制在目标检测模型中的不同位置时，其检测效果不相同。在主干网络中加入注意力机制，能够让网络更好的聚焦于图像中与目标相关的区域和特征，忽略不相关的背景等信息，从而使特征的提取更加高效；另一方面可以帮助模型在早期捕捉目标的全局信息，为后续目标的准确定位和分类提供帮助。在颈部网络中加入注意力机制，可以为融合不同层次的特征图进行优化，在融合过程中提供合理的权重来突出主要特征，使特征的融合更加合理；另一方面可以增强不同层次特征的有效传播，避免在融合过程中某些特征的丢失，进而保证特征的完整性和有效性。根据位置不同，设计了 3 种注意力机制的添加方案，并将最近流行的 CBAM、ECA、SK、EMA 和 SE 注意力机制模块进行对比分析。

(1) 如图 3-7 所示，将五种注意力机制分别加入到主干网络中三个 Concat 前，提高网络对缺陷特征的提取能力。并且，在此位置添加注意力机制，网络可以将不同视野下经过 C2f 模块提取过后的缺陷特征，直接与颈部网络的缺陷特征进行融合，从而增加缺陷的特征信息。

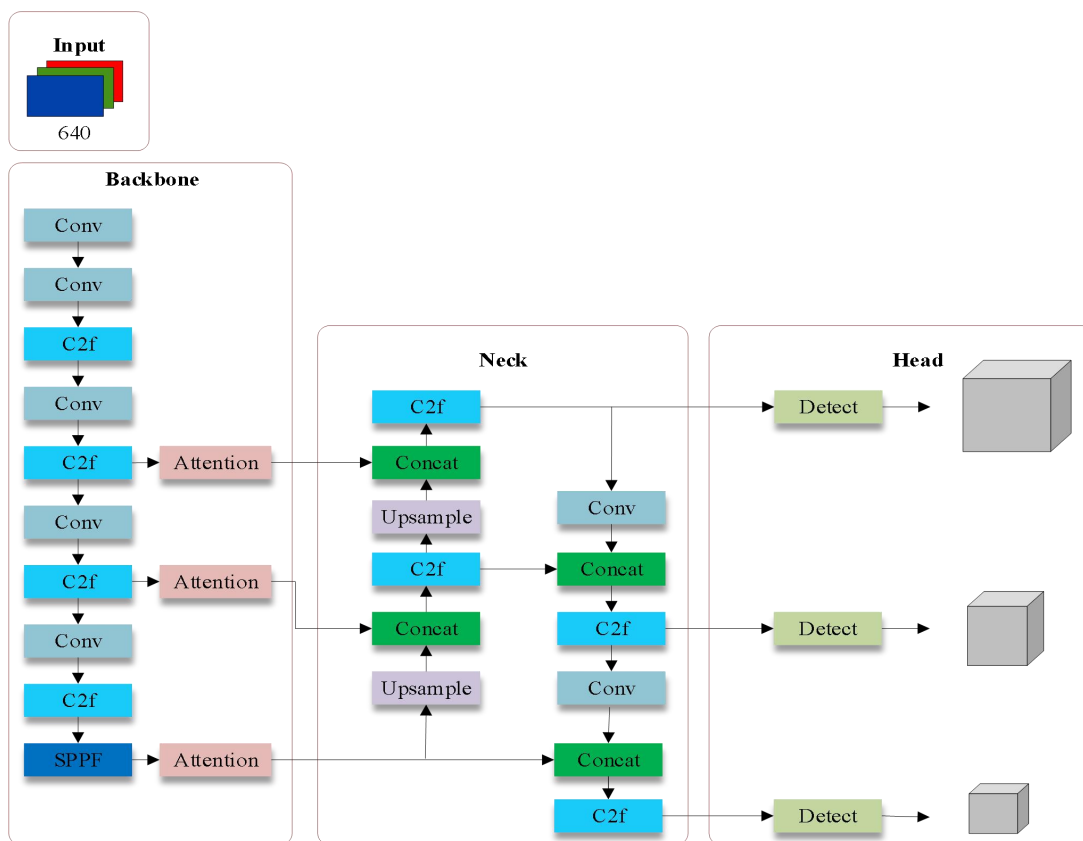


图 3-7 主干网络 Concat 前加入注意力机制结构图

在进行工业钢材表面缺陷检测后，其检测结果如表 3-2 所示。通过分析表中数据可知，SE 注意力机制在五个注意力机制中 mAP 值最高，达到了 78.6%。由 SE 注意力机制的特性可知，在此位置添加 SE 注意力机制，SE 注意力机制可以为网络提供更好的特征值权重，提高模型对于各种缺陷的特征提取能力。

表 3-2 主干网络 Concat 前注意力机制实验数据

网络模型	Gflops	Params($\times 10^6$)	mAP (%)
YOLOv8n	8.2	3.0	77.3
CBAM	8.2	3.0	78.0
ECA	8.2	3.0	77.6
SK	21.5	10.3	77.2
EMA	8.4	3.0	77.6
SE	8.2	3.0	78.6

(2) 如图 3-8 所示，在主干网络中三个 C2f 后，分别加入五种注意力机制。这种添加方式可以使网络更好的聚焦于上层网络提取到的特征信息，同时还可以帮助模型在早期捕捉目标的全局信息。

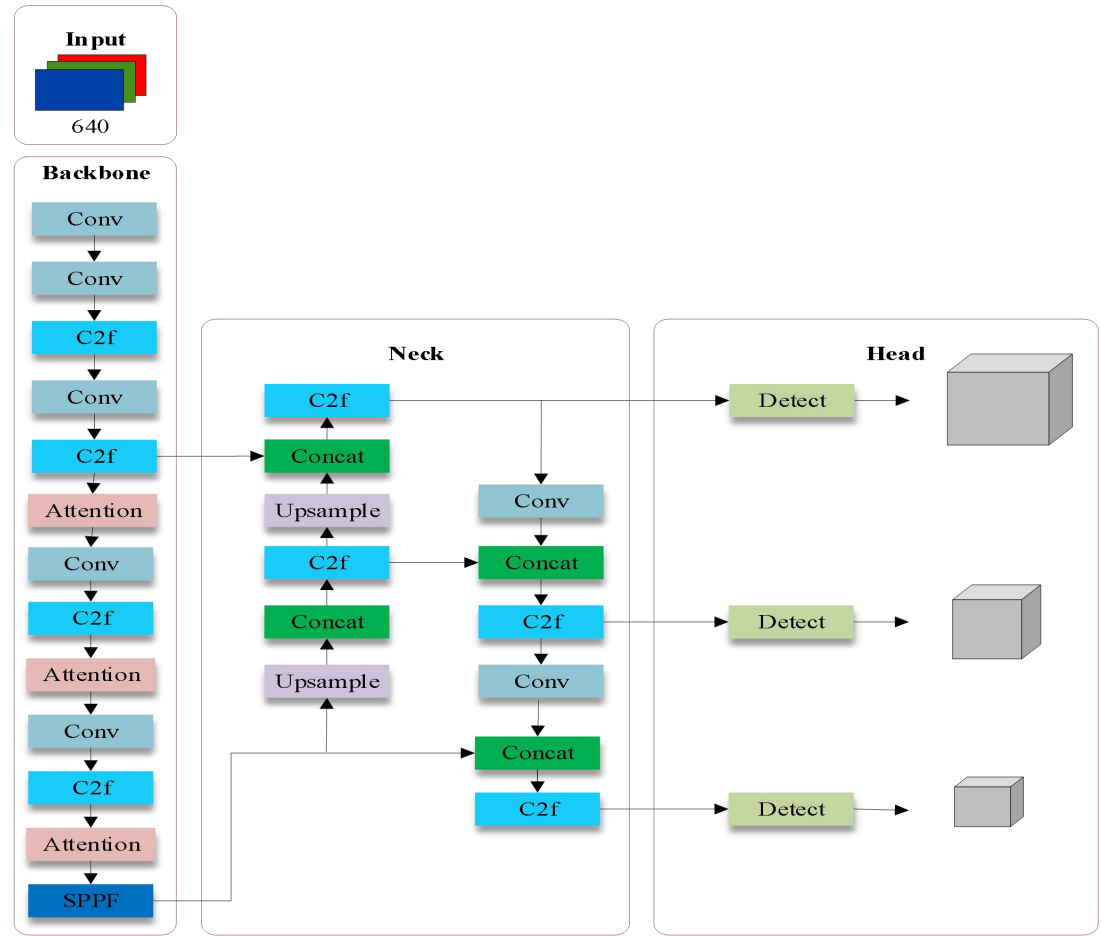


图 3-8 主干网络 C2f 后加入注意力机制结构图

在进行工业钢材表面缺陷检测后，其检测结果如表 3-3 所示。通过分析表中数据可知，ECA 注意力机制和 EMA 注意力机制的 mAP 比表中其它注意力机制高，达到了 78.5%。通过 ECA 注意力机制和 EMA 注意力机制的特性可知，在此位置添加 ECA 注意力机制和 EMA 注意力机制，可以利用注意力机制多次聚焦特征提取过程，从而增加缺陷的特征图，使网络提取到更多的缺陷特征。因此，ECA 注意力机制和 EMA 注意力机制在此位置的检测结果比其它注意力机制更有优势。

表 3-3 主干网络 C2f 后注意力机制实验数据

网络模型	Gflops	Params($\times 10^6$)	mAP (%)
YOLOv8n	8.2	3.0	77.3
CBAM	9.2	4.7	77.5
ECA	9.1	4.7	78.5
SK	22.4	12.0	77.5
EMA	9.4	4.7	78.5
SE	9.2	4.7	77.9

(3) 如图 3-9 所示，在颈部网络中三个 C2f 后，分别加入五种注意力机制，可以使网络对不同层次融合后的缺陷特征再次提取，并利用注意力机制获取的特征权重优化缺陷的主要特征，使模型更好地提取不同缺陷的特征，抑制不重要的特征。

表 3-4 颈部网络 C2f 后注意力机制实验数据

网络模型	Gflops	Params($\times 10^6$)	mAP (%)
YOLOv8n	8.2	3.0	77.3
CBAM	9.2	4.7	76.6
ECA	9.2	4.7	76.8
SK	68.8	23.4	78.3
EMA	10.0	4.7	77.6
SE	9.2	4.7	78.3

在进行工业钢材表面缺陷检测后，其检测结果如表 3.4 所示。通过分析表中数据可知，SK 注意力机制和 SE 注意力机制的 mAP 比表中其它注意力机制高，达到了 78.3%。通过 SK 注意力机制和 SE 注意力机制的特性可知，在此位置，SK 注意力机制和 SE 注意力机制获取的特征权重在赋值给获得的特征图后，可以提高网络对缺陷特征的识别。

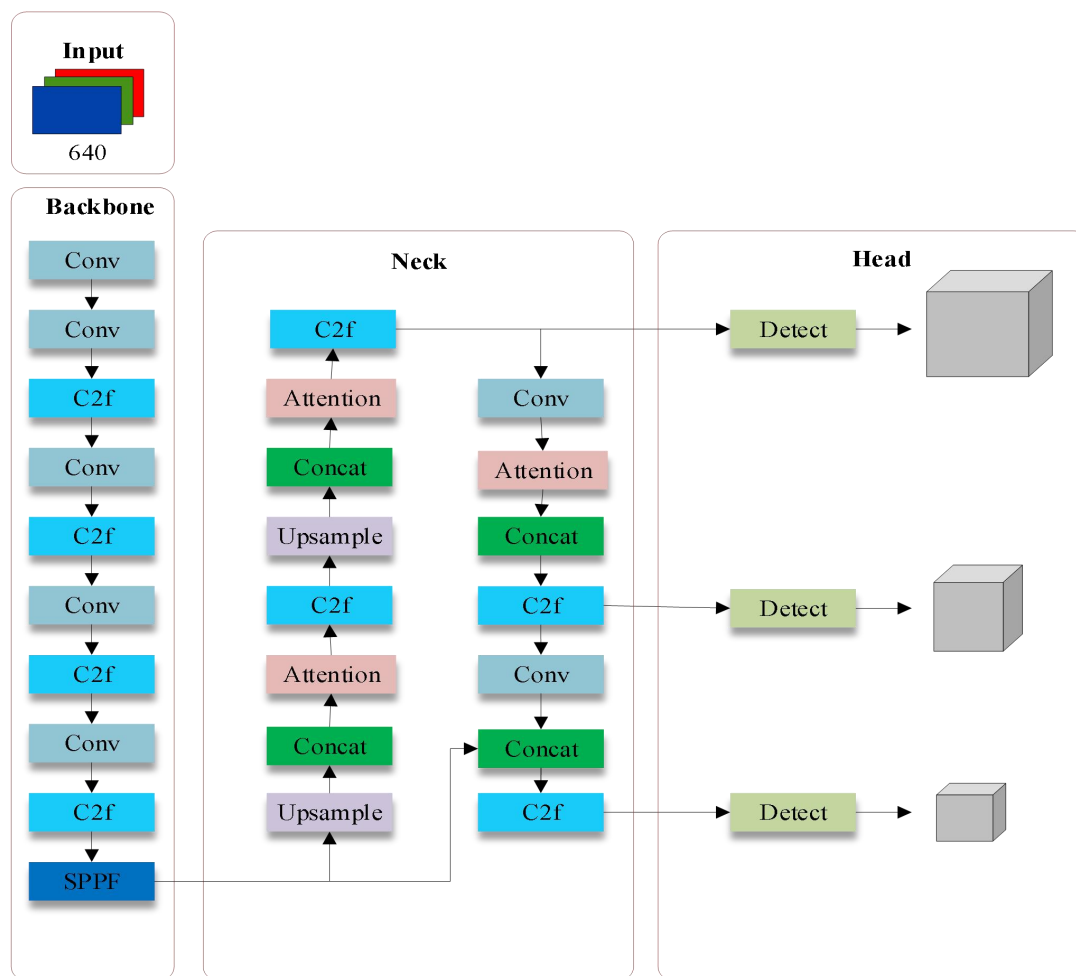


图 3-9 颈部网络 C2f 后加入注意力机制结构图

通过分析对比三种不同添加方案的检测结果表格发现，在主干网络中三个 Concat 前，加入 SE 注意力机制的方案中，网络模型对于各种缺陷特征的提取效果，比表格中其它注意力机制以及不同的添加位置更好，模型检测的 mAP 值最高为 78.6%。同时也证明了 SE 注意力机制在此位置添加后，网络利用 SE 注意力机制提取到的特征权重重新赋值给特征图，网络可以提取到更多各类缺陷特征，从而提高最后的检测结果。通过此次实验，为 YOLOv8n 引入注意力机制找到一个效果最好的方法。

3.3 可变形卷积改进

在目标检测任务中，卷积用来提取特征。在传统的卷积操作中，通常使用相同大小或比例的卷积核，以固定模式进行滑动来提取特征。这种方式对于形状规则的物体检测效果较好，但在处理形状复杂的物体时却会受到限制。而在工业钢材表面缺陷检测中，钢材表面缺陷通常呈现出不规则的形状。因此想要

在工业钢材表面缺陷检测中取得较好的检测效果，需要对传统的卷积进行改进。

3.3.1 可变形卷积 v1

2017 年，Dai 等^[59]提出了可变形卷积网络（Deformable Convolution Network, DCN），可变形卷积通过卷积核的采样点，引入可学习的卷积核偏移量，这样卷积核可以根据输入特征图的内容，自适应的调节形状，与物体的形状等特征更加贴合，从而有效的学习到物体的特征。传统卷积与可变形卷积采样点示意图如图 3-10 所示。

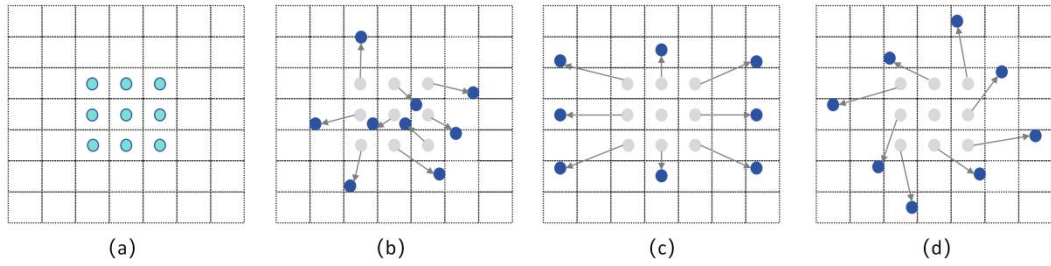


图 3-10 传统卷积与可变形卷积采样点示意图：(a)传统卷积采样位置；

(b)蓝色部分展示了传统卷积增强后偏移的采样位置；

(c)可变形卷积的尺度、各向异性的纵横比变换；(d)可变形卷积旋转的变换

在传统卷积中，输出特征图 y 上的每个位置 p_0 可以表示为：

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n) \quad (3-13)$$

式中， p_n 枚举了 R 中的位置。

在可变形卷积中，规则网格 R 上增加了偏移量 $\{\Delta p_n | n = 1, \dots, N\}$ ， $N = |R|$ ，那么对于输出特征图 y 的每个位置 p_0 可以表示为：

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (3-14)$$

因为采样是在不规则和偏移位置 $p_n + \Delta p_n$ 上进行的，而偏移量 Δp_n 通常为分数，不方便处理，因此方程要通过双线性插值实现，可以表示为：

$$x(p) = \sum_q G(q, p) \cdot x(q) \quad (3-15)$$

式中， p 表示任意位置， q 枚举了特征图中 x 的所有整数空间位置， $G(\cdot, \cdot)$ 是双线性插值核， G 是二维的，它被分为两个一维核，可以表示为：

$$G(q, p) = g(q_x, p_x) \cdot g(q_y, p_y) \quad (3-16)$$

$$g(a,b)=\max(0,1-|a-b|) \quad (3-17)$$

可变形卷积的过程如图 3-11 所示。首先，对特征图上通过卷积得到偏移量，此时卷积核具有与当前卷积层相同的扩张度和空间分辨率，因此，输出偏移量与输入特征图的空间分辨率相同。通道 $2N$ 对应于 N 个 2D 偏移量，生成输出特征的卷积核和偏移量，都是在训练期间同时进行学习的，其中偏移量的学习通过双线性插值实现。

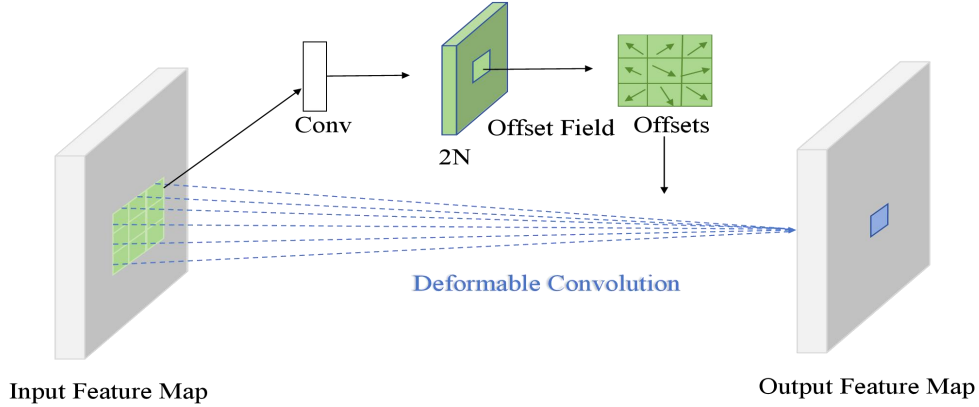


图 3-11 可变形卷积示意图

3.3.2 可变形卷积 v2

2018 年，Zhu 等^[60]的研究发现，可变形卷积网络非常适用于物体的几何变化。通过实验观测，他们发现可变形卷积特征的空间支持比常规卷积网络更符合物体结构，但是这种支持有时会远远超出感兴趣的范围，这时特征就会受到不相关图像内容的影响。因此，Zhu 等人提出了可变形卷积网络的新版本 DCNv2 算法，它具有增强学习可变形卷积的建模能力。首先，在网络中堆叠更多的可变形卷积，为更多卷积层配备偏移学习能力，使 DCNv2 能够在更广泛的特征级别上控制采样；然后是可变形卷积模块中的调制机制，可变形卷积不仅可以调整输入特征的偏移量，还可以调制来自不同空间位置的输入特征幅度。因此，DCNv2 改变了样本的相对影响和空间分布。这种调制过程可以表示为：

$$y(p)=\sum_{k=1}^K w_k \cdot x(p+p_k+\Delta p_k) \cdot \Delta m_k \quad (3-18)$$

其中 w_k 和 p_k 分别表示权重和预先指定的偏移量的第 k 个位置。 Δp_k 和 Δm_k 分别是第 k 个位置可学习到的偏移量和调制标量， $\Delta m_k \in [0,1]$ 。

如图 3-12 所示为 DCNv2 的过程。与 DCNv1 相比，卷积分支在预测采样点

偏移量的同时也预测了采样点的权重。通过调制权重，使网络可以更加关注感兴趣的区域，减少了不相关图形的干扰。

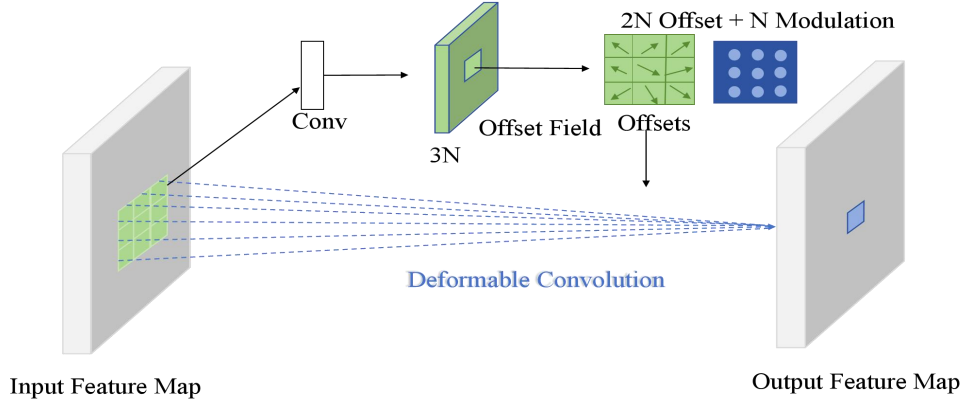


图 3-12 可变形卷积 v2 示意图

3.3.3 可变形卷积改进模型结构

DCNv2 有效改善了 DCNv1 中对不重要特征的学习，在卷积过程中可以更加专注目标特征进行学习，因此将 DCNv2 模块引入到 YOLOv8n 网络模型中，让网络专注学习各种缺陷的特征，从而网络可以更好地处理钢材表面缺陷种类复杂，尺度变化大的问题。在将 DCNv2 模块引入到 YOLOv8n 网络模型之前，将 DCNv2 模块替换 C2f 中的 Conv 模块形成新的 C2f_DCNv2 模块，并将 C2f_DCNv2 模块替换 YOLOv8n 网络中的部分 C2f 模块，然后进行钢材表面缺陷的检测，如图 3-13 所示为构成的 C2f_DCNv2 模块示意图，如图 3-14 所示为 C2f_DCNv2 模块在 YOLOv8n 网络中替换的位置。

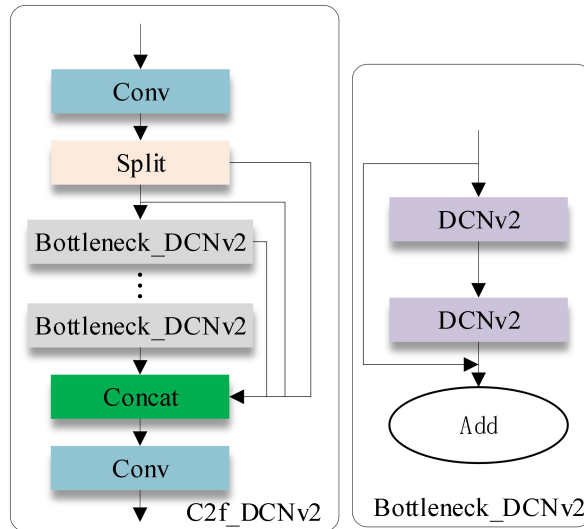


图 3-13 C2f_DCNv2 模块结构

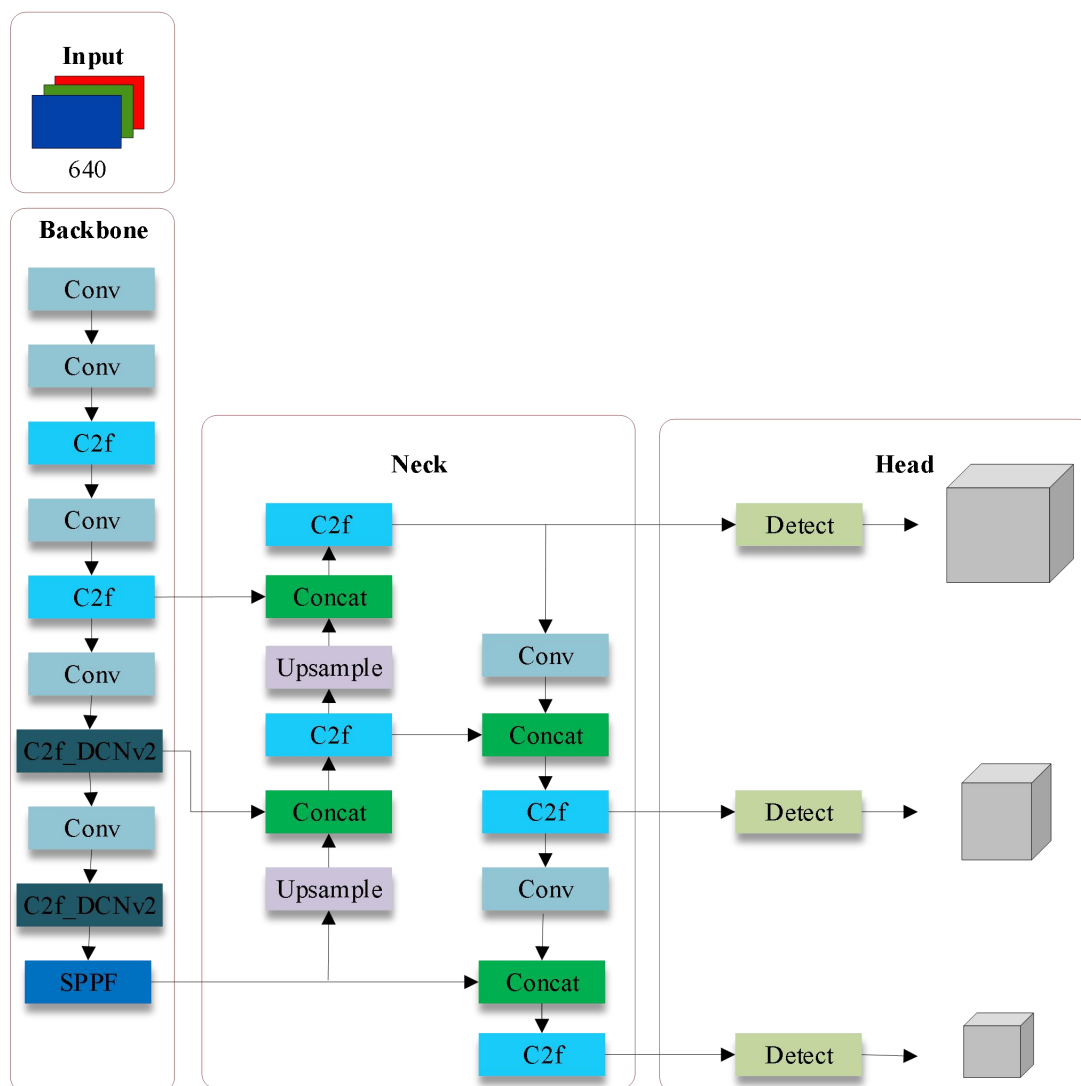


图 3-14 C2f_DCNv2 模块添加位置结构图

将 C2f_DCNv2 模块加入到算法模型中，其检测结果与原始算法的检测结果对比如表 3-5 所示。分析表中数据可知，加入 C2f_DCNv2 模块后，算法检测的 mAP 比之前提高了 0.5%，可见 C2f_DCNv2 模块在一定程度上提高了模型对各类缺陷特征的学习。

表 3-5 添加 C2f_DCNv2 模块算法结果对比

网络模型	Gflops	Params($\times 10^6$)	mAP (%)
YOLOv8n	8.2	3.0	77.3
C2f_DCNv2	7.8	3.1	77.8

3.4 本章小结

本章通过对 YOLOv8n 算法进行改进，改善工业钢材表面缺陷检测过程中

遇到的一些难题，从而提高算法对工业钢材表面缺陷的检测精度。首先,将 GIoU、DIOU 和 CIOU 三种损失函数分别引入到 YOLOv8n 算法中，在衡量模型预测结果与真实值之间差异上进行性能上的比较，根据表中数据分析出 CIOU 损失函数可以使 YOLOv8n 算法模型达到最佳性能。然后，针对工业钢材表面缺陷的颜色与钢材背景相似，很容易出现误检、漏检的难题，在 YOLOv8n 模型的不同位置设计了三种注意力添加方案，依次添加 CBAM、ECA、SK、EMA、SE 五种注意力机制模块并对比分析。在主干网络中三个 Concat 前，加入 SE 注意力机制的方案中，SE 注意力机制可以使模型充分发挥将提取特征图的权重重新赋值功能，最大程度地提高了网络模型对缺陷特征的提取能力，可以更好地提取到关键的缺陷特征信息，抑制了其它不重要的特征信息。针对钢材表面缺陷的种类复杂，大小不同的问题，提出了引用 DCNv2 模块与 C2f 模块结合，替换 C2f 模块中的普通卷积，构成新的 C2f_DCNv2 模块，增强网络对工业钢材表面各种形状缺陷特征图的学习能力，从而使网络模型可以获得更准确的缺陷形状，提高对缺陷的检测精度。

第 4 章 基于改进 YOLOv8n 的工业钢材表面缺陷检测算法

在第三章中提到的各种改进方法都可以有效提高模型的检测精度。因此在本章中，将第三章对比分析出的最优改进方法，整合到 YOLOv8n 网络模型中。在工业钢材表面缺陷数据集上进行训练验证，获得最终的工业钢材表面缺陷检测算法。通过消融实验和主流算法对比实验，分析改进后 YOLOv8n 算法的性能。

4.1 改进 YOLOv8n 的算法模型

改进后的 YOLOv8n 算法网络整体结构如图 4-1 所示，将其命名为 YOLOv8n-Improved 算法，以便后续介绍。

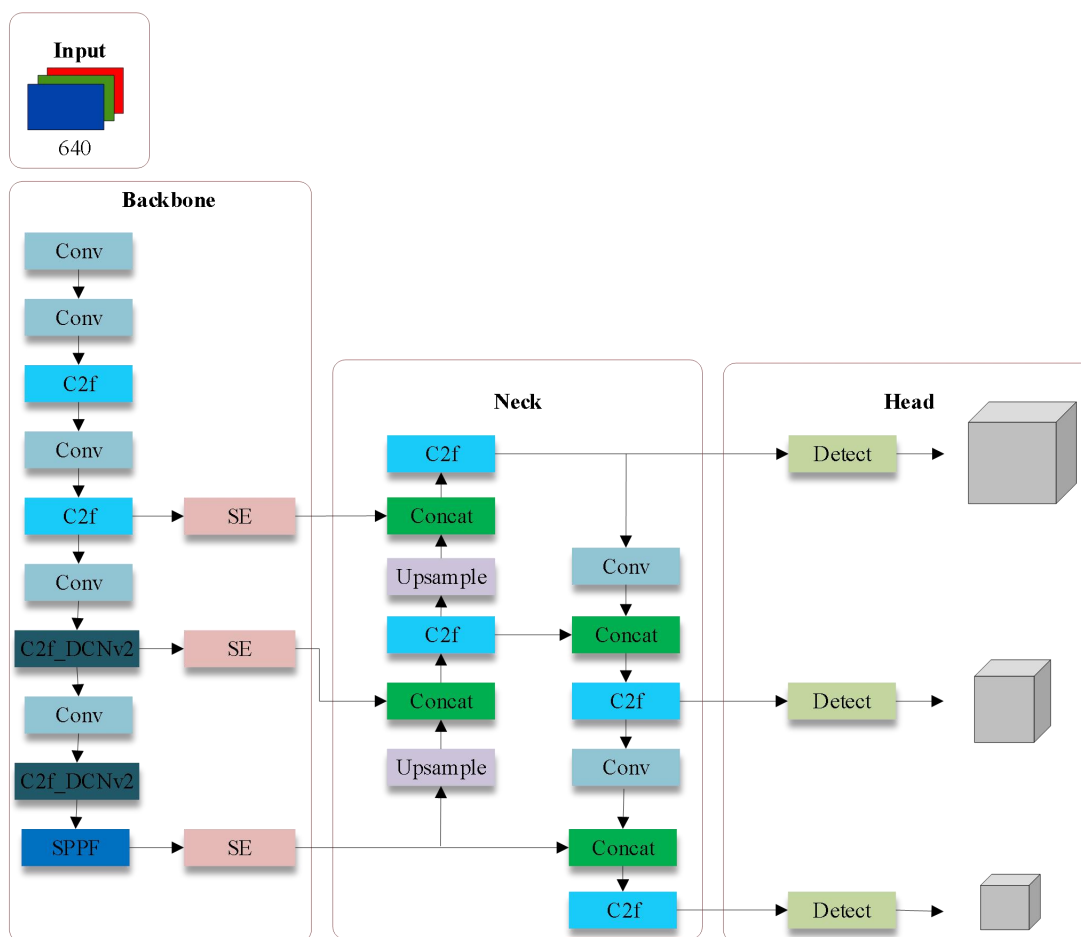


图 4-1 YOLOv8n-Improved 网络整体结构图

在 YOLOv8n-Improved 算法网络结构中，首先，使用 CIoU 损失函数使模

型在检测过程中减少模型预测结果与真实标签之间的差异，使模型达到最佳的检测性能，快速得到各种钢材表面缺陷的检测结果；其次，在 YOLOv8n 主干网络中三个 Concat 前，加入 SE 注意力机制，改善缺陷的特征提取受到图片背景的干扰问题，提高网络模型对缺陷的特征提取能力；最后，使用构建的 C2f_DCNv2 模块替换 YOLOv8n 主干网络中部分 C2f 模块，改善钢材表面缺陷种类复杂，尺度变化大的问题，提高卷积模块对形状各异缺陷特征的学习能力。

4.2 实验过程及分析

4.2.1 工业钢材表面缺陷数据集

为了评估改进算法模型的可行性，本文采用东北大学宋克臣副教授个人主页中的带钢 NEU-DET 数据集进行实验。在 NEU-DET 数据集中共有 1800 张灰度图片，其中含有划痕、裂纹、斑块、夹杂、麻点以及压入氧化铁皮 6 种缺陷，每种缺陷各 300 张，其中各类缺陷如图 4-2 所示。

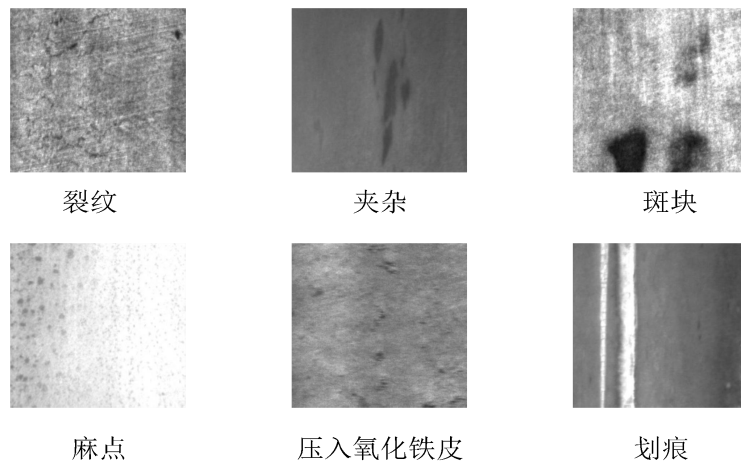


图 4-2 钢材表面各类缺陷

裂纹(Crazing): 在钢材生产过程中如果钢材受热速度太快或受热不均匀，导致热应力过大，表面会产生裂纹缺陷。

夹杂(Inclusion): 在钢材生产过程中，外部环境的非金属夹杂物附着在钢材表面或嵌入到钢材里不能脱落，这样就形成了夹杂缺陷。

斑块(Patches): 在钢材进行冷却加工时表面发生了化学反应，产生一些氯化亚铁，而氯化亚铁中的铁元素呈现出黑色，这样就形成了斑块缺陷。

麻点(Pitted Surface):在钢材的轧制过程中，如果轧辊的表面出现了损失或者

沾有金属屑等异物，那么钢材表面就会出现凹痕从而出现麻点缺陷。

压入氧化铁皮(Rolled-in Scale): 在钢材的轧制过程中，有些氧化铁皮被压入钢材中，而在冷却后经过酸性物质清洗，这些氧化铁皮容易脱离，从而形成压入氧化铁皮缺陷。

划痕(Scratches): 钢材表面出现磨损或者刮伤痕迹形成划痕缺陷。

由于公开的数据集中有些钢材表面缺陷没有被完整的标注，因此在实验前使用 LabelImg 工具对数据集中的缺陷进行查看和手动标注，其标注界面如图 4-3 所示。本次实验将处理后的数据集，按照 8:2 的比例分为训练集和验证集。

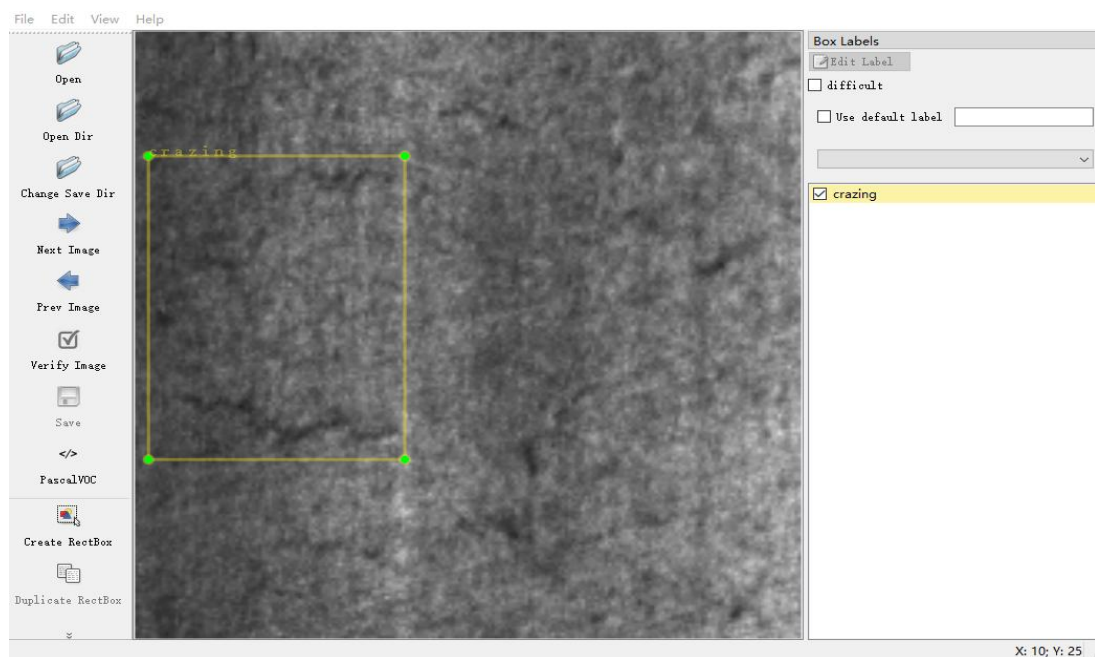


图 4-3 LabelImg 操作界面

4.2.2 实验环境

本次实验使用 Anaconda 安装 Pytorch 环境，并且使用 Pycharm 运行深度学习算法，其具体的环境配置如表 4-1 所示。

表 4-1 实验环境

名称	版本型号
CPU	Intel(R)Core(TM)i9-10900X CPU @3.70GHz
GPU	NVIDIA GeForce RTX3080
操作系统	Windows 10
框架	PyTorch 2.0.0+Cuda 11.8

在训练算法时，我们使用“yolov8n.pt”作为预训练模型权重，其中 Epochs

值设为 300，Batch Size 值设为 16。当 Epochs 值设为 1 时，完整地遍历 1 次整个训练数据集，此次实验需要完整的遍历 300 次整个训练数据集，在每次遍历完后，预训练模型权重都会更改。Batch Size 是指由于训练的数据集很大，不能一次性将全部的数据集放入计算资源（GPU）中进行训练，因此会将数据集分成若干小批次。在运行算法时，YOLOv8n-Improved 的网络参数如表 4-2 所示。

表 4-2 YOLOv8n-Improved 网络参数

序号	From	Parameters	Module	Arguments
0	-1	464	Conv	[3,16,3,2]
1	-1	4672	Conv	[16,32,3,2]
2	-1	7360	C2f	[32,32,1,True]
3	-1	18560	Conv	[32,64,3,2]
4	-1	49664	C2f	[64,64,2,True]
5	-1	73984	Conv	[64,128,3,2]
6	-1	260204	C2f_DCN2	[128,128,2,True]
7	-1	295424	Conv	[128,256,3,2]
8	-1	522806	C2f_DCN2	[256,256,1,True]
9	-1	164608	SPPF	[256,256,5]
10	4	512	SE	[64]
11	6	2048	SE	[128]
12	9	8192	SE	[256]
13	-1	0	Upsample	[None,2]
14	[-1,11]	0	Concat	[1]
15	-1	148224	C2f	[384,128,1]
16	-1	0	Upsample	[None,2]
17	[-1,10]	0	Concat	[1]
18	-1	37248	C2f	[192,64,1]
19	-1	36992	Conv	[64,64,3,2]
20	[-1,15]	0	Concat	[1]
21	-1	123648	C2f	[192,128,1]
22	-1	147712	Conv	[128,128,3,2]
23	[-1,12]	0	Concat	[1]
24	-1	493056	C2f	[384,256,1]
25	[18,21,24]	752482	Detect	[6,[64,128,256]]

表中，From 列代表此层的输入来源，-1 表示来自上一层输入；Parameters 列代表模型中的参数量；Module 列代表模型各模块的名称；Arguments 列代表模块中的参数信息。

4.2.3 消融实验

在得到最佳的改进方案后，为了证明各改进模块的有效性，对改进后的算法进行消融实验，本次实验分为 3 组，其中“√”表示该组采用此模型结构，实验结果如表 4-3 所示。

此次消融实验以 YOLOv8n 模型并使用 CIOU 损失函数作为基组，依次使用上述提到的方法对 YOLOv8n 进行改进。首先，在 YOLOv8n 模型中使用 CIOU 损失函数使模型达到最佳性能并快速给出检测结果，其检测结果如组 1 所示，其中 mAP 代表检测的平均检测精度为 77.3%，GFLOPs 是模型在运行过程中的计算量为 3.0，Params 代表模型中需要学习的参数总量为 8.2；将 SE 注意力机制添加到 YOLOv8n 模型中，从表中第 2 组与第 1 组对比数据可知，添加 SE 注意力机制使得 mAP 增加了 1.3%，达到了 78.6%，SE 注意力机制提取缺陷特征的能力和重新分配提取特征权重的功能，提高了模型对缺陷特征的提取，抑制了不相关特征图的干扰；将第 3 组和第 2 组进行对比可知，在网络主干中使用 DCN2 卷积与 C2f 模块中的 Conv 卷积结合组成新的 C2f_DCN2 模块后，网络对不同类型的缺陷特征学习能力得到提高，并且有效降低了模型复杂度，使改进后模型的 mAP 提高了 1.0%，但模型的参数量略有微升高。最终，改进的算法模型的检测精度比原始算法提高了 1.9%，改善了检测精度的问题。

表 4-3 消融实验数据

添加模块	组 1	组 2	组 3
YOLOv8n(CIOU)	√	√	√
SE	—	√	√
C2f_DCN2	—	—	√
mAP(%)	77.3	78.6	79.2
GFLOPs	8.2	8.2	7.8
Params($\times 10^6$)	3.0	3.0	3.1

4.2.4 实验结果对比与分析

为了更直观地感受到 YOLOv8n-Improved 相较于 YOLOv8n 的优势，本文给出了两种算法在损失曲线、性能曲线上的比较曲线图。

如图 4-4 所示为 YOLOv8n-Improved 算法与 YOLOv8n 算法相比较的损失曲

线图。图(a)展示了训练过程中的定位损失曲线，在经过 100Epochs 后，YOLOv8n-Improved 算法定位损失的下降速度明显开始快于 YOLOv8n 算法，这表明 YOLOv8n-Improved 算法在预测目标位置时的正确性开始高于 YOLOv8n 算法。图(b)展示了训练过程中的分类损失曲线，在 25Epochs 后，YOLOv8n-Improved 算法分类损失的下降速度明显快于 YOLOv8n 算法，这表明 YOLOv8n-Improved 算法对缺陷的分类的正确性开始高于 YOLOv8n 算法。同理，图(c)和图(d)是验证过程中的定位损失曲线和分类损失曲线，从图中也可以看出 YOLOv8n-Improved 算法的优越性。

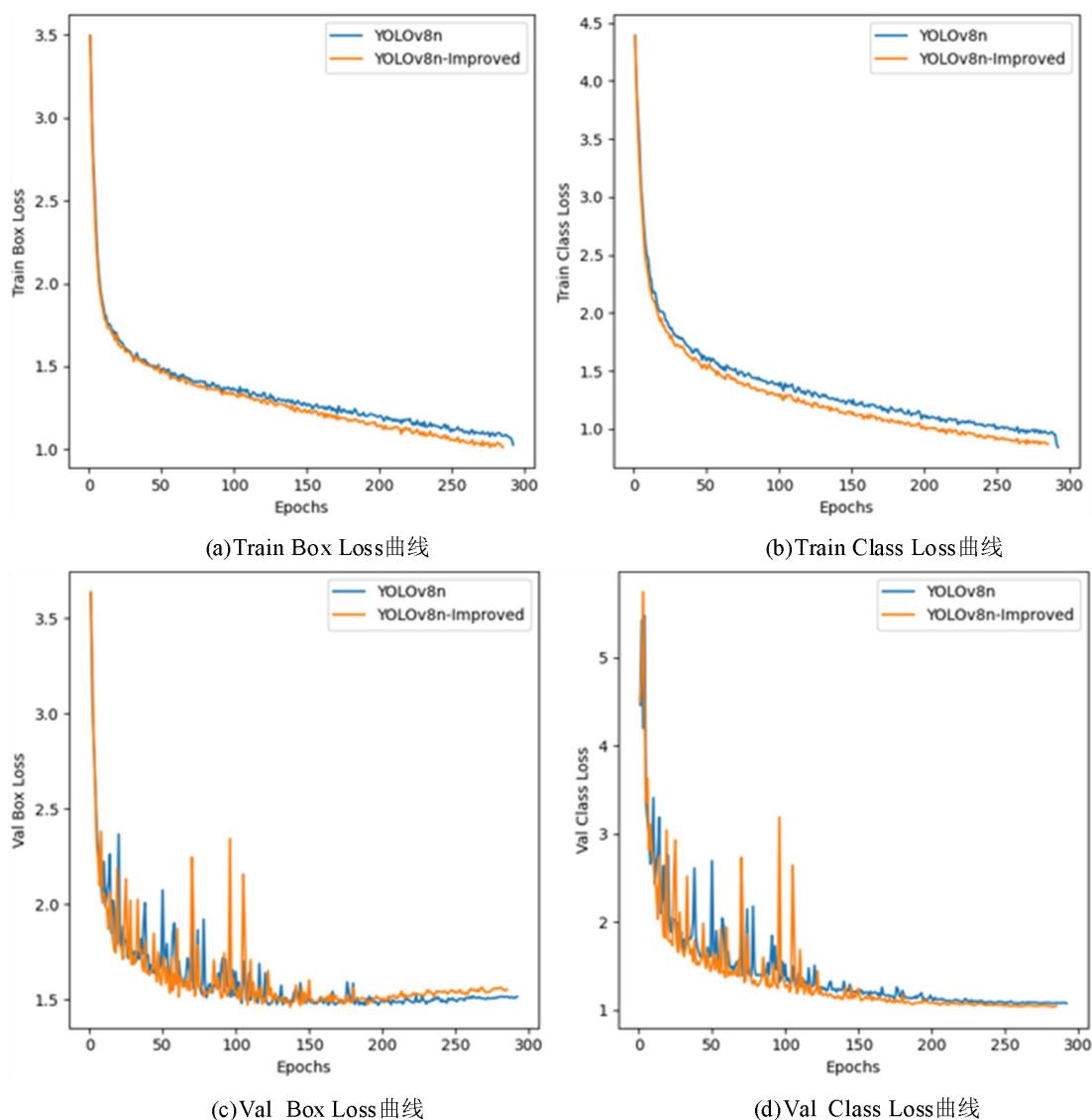


图 4-4 YOLOv8n-Improved 与 YOLOv8n 算法各类损失函数曲线对比图

如图 4-5 所示为 YOLOv8n-Improved 算法与 YOLOv8n 算法相比较的性能曲线图。图(a)展示了训练过程中精度变化的曲线图，在经过 120Epochs 后，

YOLOv8n-Improved 算法的精度曲线开始高于 YOLOv8n 算法，这表明 YOLOv8n-Improved 算法预测的正例样本开始高于 YOLOv8n 算法。图(b)展示了训练过程中召回率变化的曲线图，在 240Epochs 后，YOLOv8n-Improved 算法的召回率曲线开始高于 YOLOv8n 算法，这表明 YOLOv8n-Improved 算法在找出真正正例的能力开始高于 YOLOv8n 算法。图(c)展示了训练过程中 IoU 阈值取 0.5 时平均精度均值的曲线图，在经过 120Epochs 后，YOLOv8n-Improved 算法的平均精度均值曲线开始高于 YOLOv8n 算法。图(d)展示了训练过程中 IoU 阈值在 0.5 到 0.95 之间时平均精度的曲线图，在经过 120Epochs 后，YOLOv8n-Improved 算法的平均精度均值曲线开始高于 YOLOv8n 算法。因此 YOLOv8n-Improved 算法在性能上要优于 YOLOv8n 算法。

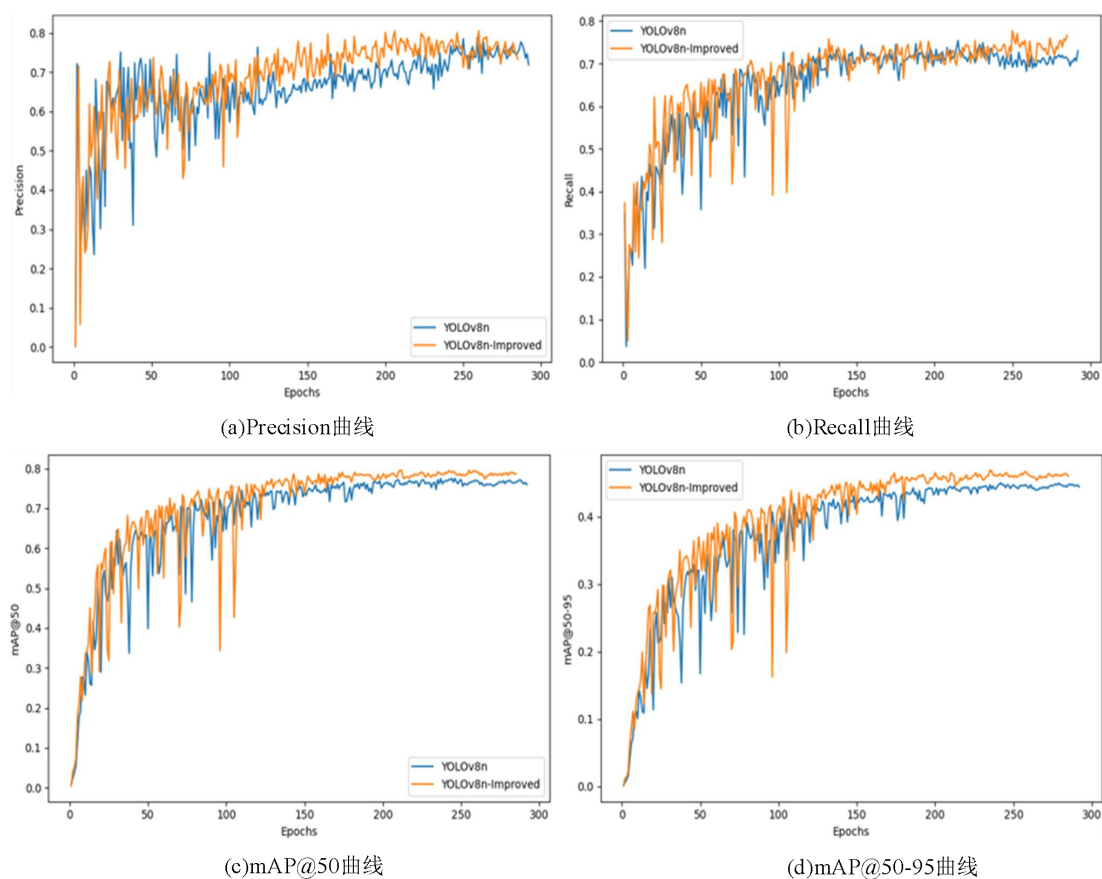


图 4-5 YOLOv8n-Improved 与 YOLOv8n 算法各性能对比曲线图

YOLOv8n-Improved 算法的优越性不仅仅体现在损失曲线和性能曲线上，通过 PR 曲线图，可以进一步了解 YOLOv8n-Improved 算法在各种缺陷上的检测结果。YOLOv8n 算法和 YOLOv8n-Improved 算法在各种缺陷上的检测结果的 PR 曲线图如图 4-6、如图 4-7 所示。

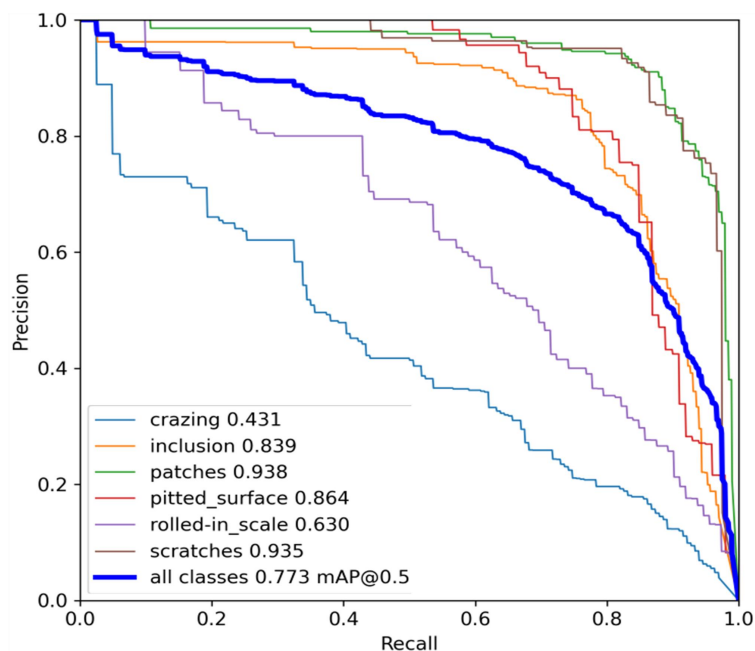


图 4-6 YOLOv8n 检测结果的 PR 曲线

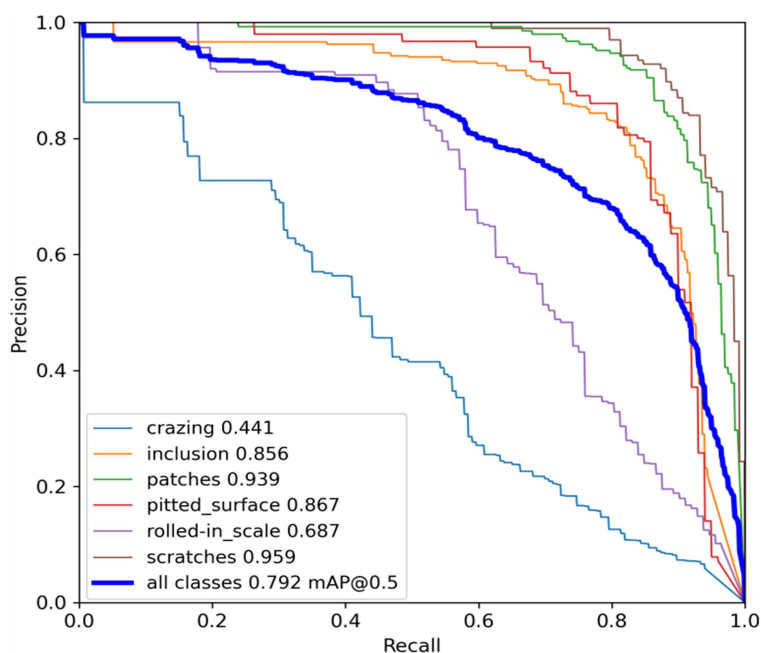


图 4-7 YOLOv8n-Improved 检测结果的 PR 曲线

从整体围成面积来分析，YOLOv8n-Improved 模型的 PR 曲线面积相较于 YOLOv8n 模型的 PR 曲线面积大，因而对缺陷的检测性能更好一些。再从对各种缺陷的检测准确度分析，YOLOv8n-Improved 模型中 **crazing** 类型缺陷的准确度从 43.1%提升到了 44.1%，**inclusion** 类型缺陷的准确度从 83.9%提升到了 85.6%，**patches** 类型缺陷从 93.8%提升到了 93.9%，**pitted_surface** 类型缺陷的准确度从 86.4%提升到了 86.7%，**rolled-in_scale** 类型缺陷的准确度从 63.0%提升到了 68.7%，**scratches** 类型缺陷从 93.5%提升到了 95.9%。

如图 4-8 所示，可以直接了解 YOLOv8n-Improved 算法在工业钢材表面缺陷的难点检测上的改善情况。YOLOv8n-Improved 算法在检测划痕缺陷时，检测到了 YOLOv8n 算法未检测到的较小划痕缺陷，这表明 YOLOv8n-Improved 算法有效改善了 YOLOv8n 算法在工业钢材表面种类尺度变化大的情况下，出现漏检的情况。

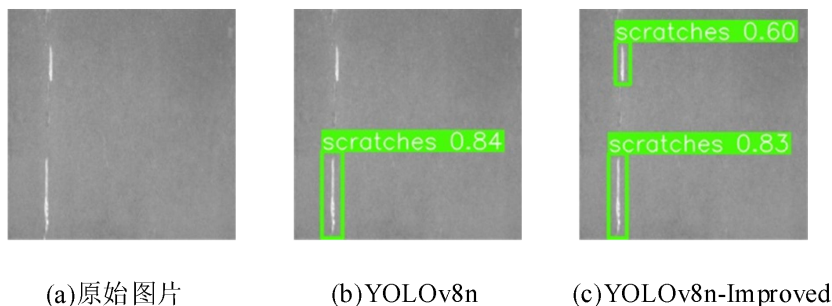


图 4-8 划痕缺陷漏检检测图

如图 4-9 所示，YOLOv8n-Improved 算法有效的解决了缺陷误检的情况。在检测两种相似的斑块缺陷和压入氧化铁皮缺陷时，YOLOv8n-Improved 算法相较于 YOLOv8n 算法正确识别出了压入氧化铁皮缺陷。

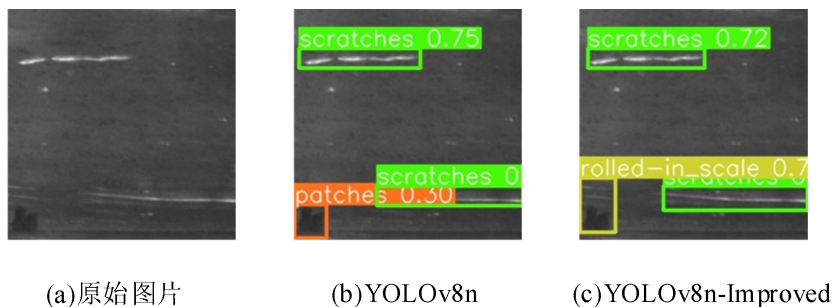


图 4-9 划痕缺陷误检检测图

如图 4-10 所示，YOLOv8n-Improved 算法相较于 YOLOv8n 算法在对斑块缺陷的检测不仅改善了漏检问题，在检测精度上也得到了较高的提升。

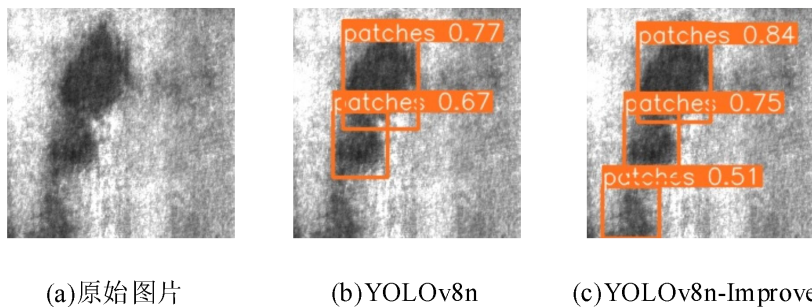


图 4-10 斑块缺陷检测图

为了进一步查看 YOLOv8n-Improved 算法在各种缺陷上的检测精度提升情

况，如图 4-11 所示，给出了 YOLOv8n 与 YOLOv8n-Improved 在各种缺陷上检测的部分结果图。从图中可以看出，这 6 种工业钢材表面缺陷的检测精度都得到了有效提升，其中夹杂缺陷和压入氧化铁皮缺陷的检测精度提升效果显著。划痕缺陷和斑块缺陷的检测精度提升效果一般。

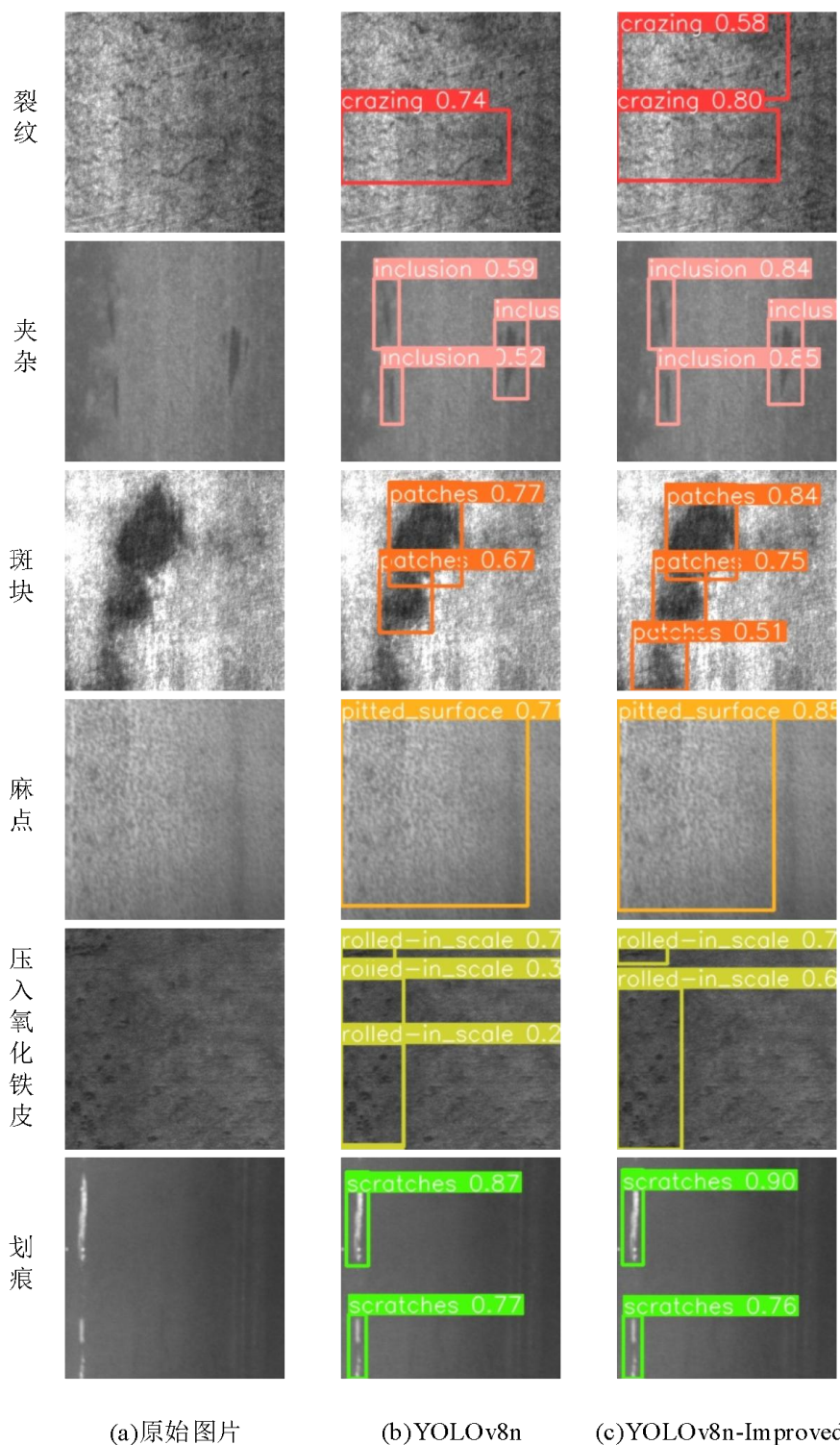


图 4-11 改进前后效果对比

为了更直观地了解 YOLOv8n-Improved 算法在各种缺陷检测上提升效果，

结合 YOLOv8n 算法和 YOLOv8n-Improved 算法在各种缺陷上的检测结果的 PR 曲线图绘制了柱状图，如图 4-12 所示。

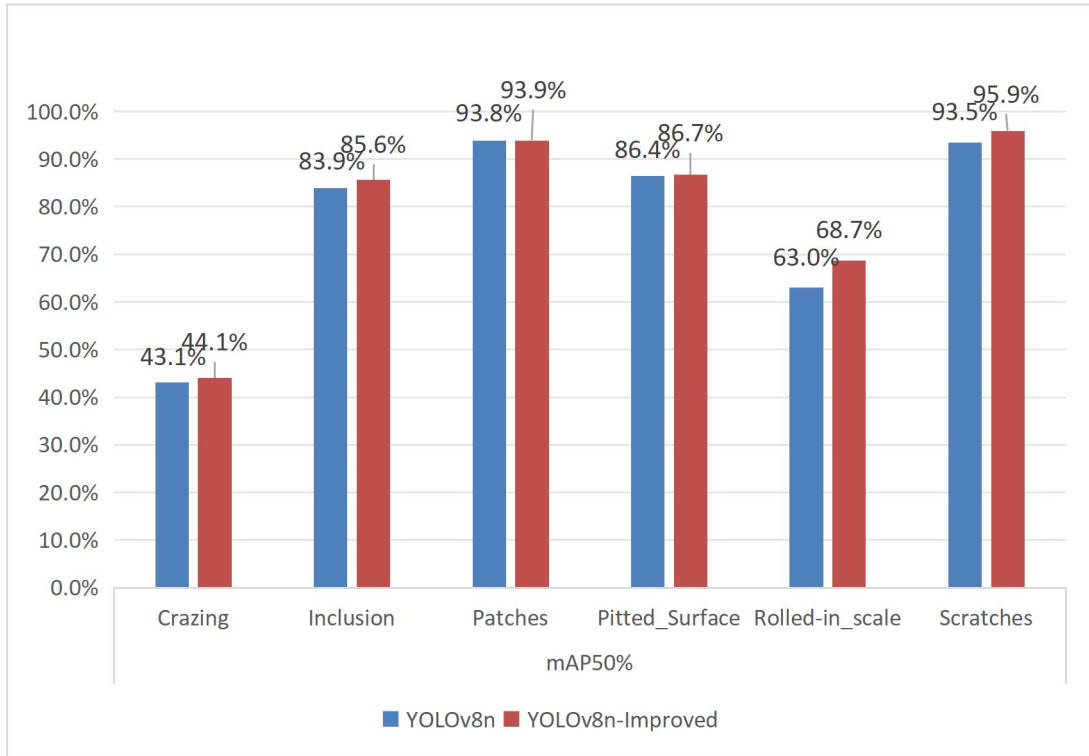


图 4-12 YOLOv8n 与 YOLOv8n-Improved 各类缺陷精度对比图

4.2.5 主流算法对比与分析

如今，主流算法已经成为行业标准，为了验证改进算法的合理性和应用价值，与一些流行的目标检测算法和一些近期研究者在 YOLOv8n 的基础上改进的算法，在 NEU-DET 数据集上检测，进行检测结果上的对比，如表 4-4 所示。

表 4-4 各改进算法数据

算法	Params($\times 10^6$)	GFLOPs	mAP (%)
Faster-RCNN	56.8	-	77.2
YOLOv3	61.5	153.4	75.9
YOLOv5s	7.0	16.0	76.7
YOLOV5m	20.8	45.2	77.7
YOLOv8n	3.0	8.2	77.3
文献[61]	3.3	7.8	76.9
文献[62]	2.5	6.2	77.1
文献[63]	3.1	8.7	79.3
YOLOv8n-Improved	3.1	7.3	79.1

从表中数据可以发现，主流的两阶段 Faster-RCNN 算法不仅参数量上比 YOLOv8n-Improved 算法要高出十几倍，在 mAP 的比较中也没有 YOLOv8n-Improved 算法高。在单阶段目标检测算法中，主流的 YOLOv3、YOLOv5s、YOLOv5m 算法模型在 mAP 比较中也是劣势，最高的 mAP 与 YOLOv8n-Improved 算法的 mAP 相差 1.4%，而且 YOLOv5m 算法的参数量和计算量非常大，不利于算法的部署。在与近期研究者们提出的改进算法相比较时，最高的 mAP 也只比 YOLOv8n-Improved 算法的 mAP 高 0.2%，但是在计算量上比 YOLOv8n-Improved 算法多了 1.4GFLOPs，会增加设备的计算资源。因此，YOLOv8n-Improved 算法具备替代一些现有方案的潜力，是一个优秀的改进算法。

4.3 本章小结

本章通过将第三章提出的最佳改进方法整合到 YOLOv8n 算法中，提出了 YOLOv8n-Improved 算法。在进行消融实验中，使用 CIoU 损失函数的 YOLOv8n 算法检测精度达到了 77.3%；在 YOLOv8n 主干网络中三个 Concat 前加入 SE 注意力机制后，算法的检测精度提高了 1.3%，达到了 78.6%，提高了算法对缺陷特征提取能力；最后将构建的 C2f_DCNv2 模块替换 YOLOv8n 主干网络中部分 C2f 模块后，算法的检测精度又提高了 0.6%，达到了 79.2%，算法对各种类型缺陷的特征学习得到了提高。通过对比改进前后算法对缺陷检测的结果，YOLOv8n-Improved 算法有效改善了 YOLOv8n 算法误检、漏检的情况。在与其他主流算法进行比较时，也可以看出 YOLOv8n-Improved 算法在平均检测精度上也具有优势。

第 5 章 工业钢材表面缺陷检测算法的轻量化及系统设计

工业钢材表面缺陷检测是工业钢材生产过程中的一项重要环节，轻量化的算法在工业钢材表面缺陷检测中具有重要意义。首先，轻量化的算法具有更快的应用速度，可以保证实时检测的要求。其次，轻量化的算法可以使设备在提供少量的计算资源的情况下就可以高效运行，特别适用于嵌入式系统或边缘计算设备，为中小型企业提供更好的设备选择，降低设备对硬件的需求。因此，在第 4 章 YOLOv8n-Improved 算法的基础上，使用三种轻量化卷积依次对网络模型进行轻量化。将在检测精度和模型大小上平衡效果最好的轻量化模型，与 PySide6 联合设计了一个工业钢材表面缺陷检测系统，并检验系统运用到工业嵌入式设备之前的检测功能。

5.1 模型轻量化

YOLOv8 采用了基于 EfficientNet 的缩放模型方式^[64]，该方法通过对系数进行缩放来调整卷积层的通道数和层数，即通过改变网络的宽度和深度来获得不同规模的模型参数，达到减少网络的复杂性的目的。这种缩放模型的方法是在不改变模型整体框架的前提下，使模型适应各种检测精度的需求，YOLOv8n 就是其中一种缩放模型。虽然通过减少网络的宽度和深度进行简化网络结构的方式，会大量减少算法的计算量和参数量，但是这也会使算法模型检测精度大幅度降低。通常，深度学习模型压缩的方式有量化、轻量化模型架构、剪枝以及知识蒸馏。为了使工业钢材表面缺陷检测可以在实时性、资源限制、成本效益和移动性等方面取得更好的平衡，轻量化模型构架成为了更好的选择。轻量化模型的核心是在尽量保持精度的前提下，对网络模型的参数量和计算量进行减少。本章中，分别使用 GhostNet 卷积、深度可分离卷积（Depthwise Separable Convolution, DWConv）和 DSConv 卷积（Distribution Shifting Convolution）三种轻量化模块，对 YOLOv8n-Improved 网络模型进行轻量化，从而得到更加切合工业检测任务的算法模型。

5.1.1 GhostNet 卷积

2020 年，华为诺亚方舟实验室提出了一种轻量级卷积模块 GhostNet^[65]，在进行得到特征图操作时，他们发现许多特征图可以不需要进行复杂的非线性变换，只要对其中一个特征图通过简单线性变换操作就可以得到另一特征图，从而实现减少特征图的冗余，使得模型更加轻量化。如图 5-1 所示展示了 GhostNet 卷积的基本原理：首先，使用部分卷积核用来卷积操作生成特征图；然后，对生成的特征图进行简单的线性运算生成更多的特征图；最后，将全部的特征图用来组成新的输出。

假设输出通道和卷积核的数量分别为 c 和 n ，特征图的高和宽分别为 h 和 w ，线性变换的卷积核大小为 d ，变换数量为 s 。在经过卷积后，其卷积核大小为 k ，输出的特征图高和宽分别为 h' 和 w' 。那么普通卷积和 GhostNet 卷积的参数数量之比为：

$$rs = \frac{n \cdot h' \cdot w' \cdot c \cdot k \cdot k}{\frac{n}{s} \cdot h' \cdot w' \cdot c \cdot k \cdot k + (s-1) \cdot \frac{n}{s} \cdot h' \cdot w' \cdot d \cdot d}$$

$$= \frac{c \cdot k \cdot k}{\frac{1}{s} \cdot c \cdot k \cdot k + \frac{s-1}{s} \cdot d \cdot d} = \frac{s \cdot c}{s + c - 1} = s \quad (5.1)$$

在得到的普通卷积和 GhostNet 卷积的参数数量之比中我们可以看出，普通卷积的参数数量是 GhostNet 卷积参数数量的 s 倍。因此采用 GhostNet 卷积可以使模型轻量化。

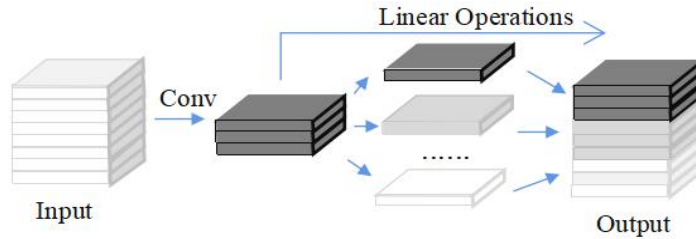


图 5-1 GhostNet 卷积结构

5.1.2 深度可分离卷积

标准卷积，利用若干个多通道卷积核对输入的多通道图像进行处理，输出的 feature map 既提取了通道特征，又提取了空间特征，其卷积过程如图 5-2 所

示。以 $5 \times 5 \times 3$ 的输入为例，卷积核大小为 3×3 ，输出通道数为 4。那么常规卷积中得到 $4 \times 3 \times 3 \times 3 = 108$ 个参数，计算量为 $3 \times 3 \times (5-2) \times (5-2) \times 3 \times 4 = 972$ 。

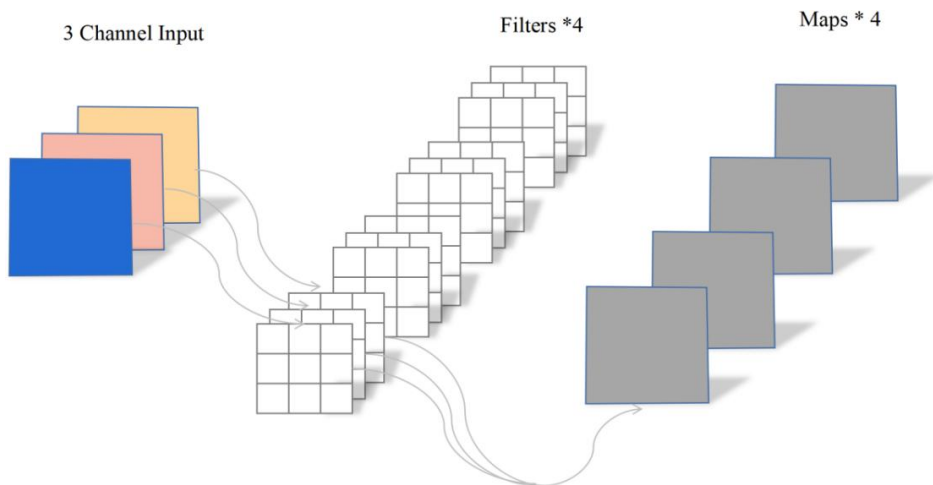
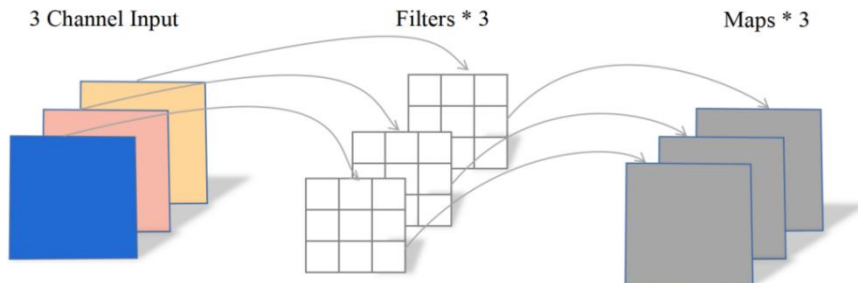
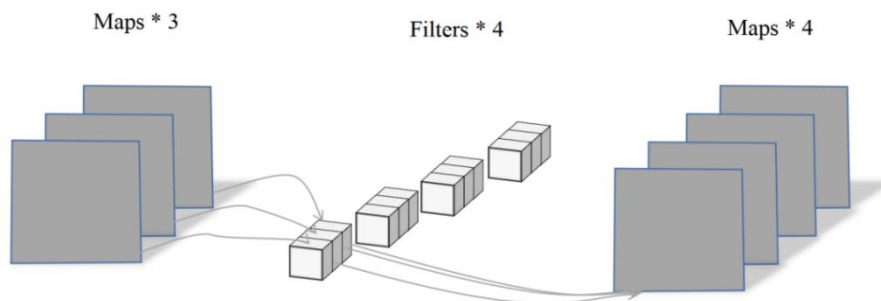


图 5-2 标准卷积示意图

在深度可分离卷积过程中，分别使用逐通道卷积和逐点卷积来提取图形的空间特征与通道特征。深度可分离卷积在特征维度上分组卷积，对每个 **channel** 进行独立的逐通道卷积，并在输出前使用一个 1×1 卷积将所有通道进行聚合，从而达到了减少模型的计算量和参数量，提高了计算效率。



(a) 逐通道卷积



(b) 逐点卷积

图 5-3 深度卷积可分离示意图

深度可分离卷积的过程如图 5-3 所示。逐通道卷积过程中每个通道各自对应一个不同的卷积核进行卷积，然后将生成的特征图再进行逐点卷积。逐点卷

积的运算过程和普通卷积类似，在进行逐点卷积过后就可以得到最终的特征图。以 $5 \times 5 \times 3$ 的输入为例，卷积核大小为 3×3 ，输出通道数为 4。深度可分离卷积中逐通道卷积中有 $3 \times 3 \times 3 = 27$ 个参数，计算量为 $3 \times 3 \times (5-2) \times (5-2) \times 3 = 243$ 。在逐点卷积中有 $1 \times 1 \times 3 \times 4 = 12$ 个参数，计算量为 $1 \times 1 \times 3 \times 3 \times 3 \times 4 = 108$ 。所以深度可分离卷积有 $27 + 12 = 39$ 个参数，总计算量为 $243 + 108 = 351$ ，可以看出深度可分离卷积是常规卷积的 $1/3$ ，因此采用深度可分离卷积可以使网络的参数量和计算量得到减少，从而实现模型轻量化的目的。

5.1.3 DSConv 卷积

2019 年，Marcelo 等人在传统卷积的基础上提出了一种变体为 DSConv，DSConv 可以很轻松的替换到传统的神经网络框架中，并减少内存的使用量，加快网络的计算速度。DSConv 卷积主要分为可变量化内核（VQK）和分布偏移两个部分。VQK 为 DSConv 卷积的量化组件，可以实现更快且内存高效的乘法。它用来保存可变位长整数值，其大小与传统卷积相同。分布偏移与传统卷积操作中卷积核的作用类似，它用来偏移 VQK 的分布。偏移过程主要通过内核分布移位器（Kernel Distribution Shifter，KDS）张量和通道分布移位器（Channel Distribution Shifter，CDS）张量来实现，其中 KDS 负责移动每个内核切片中的分布，而 CDS 负责调整每个通道切片中的分布位置。当传统卷积的张量大小为 (ch_o, ch_i, k, k) 时， ch_o 为下一层的通道数， ch_i 为当前层数的通道数， k 是内核的高度和宽度，其过程如图 5-4 所示。

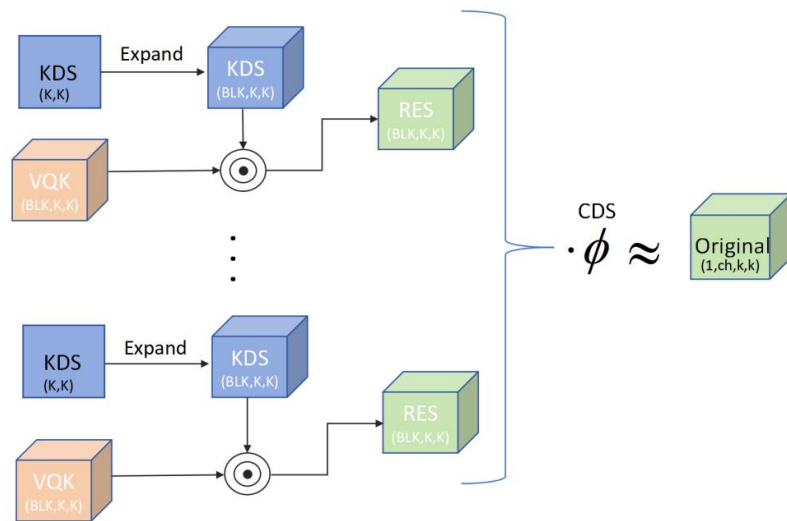


图 5-4 DSConv 卷积示意图

5.1.4 实验结果对比与分析

将三种轻量化卷积模块分别替换 YOLOv8n-Improved 模型主干网络的普通卷积模块，替换位置如图 5-5 所示，减少模型的计算量和参数量，得到更适用于工业中嵌入式设备的轻量化模型。

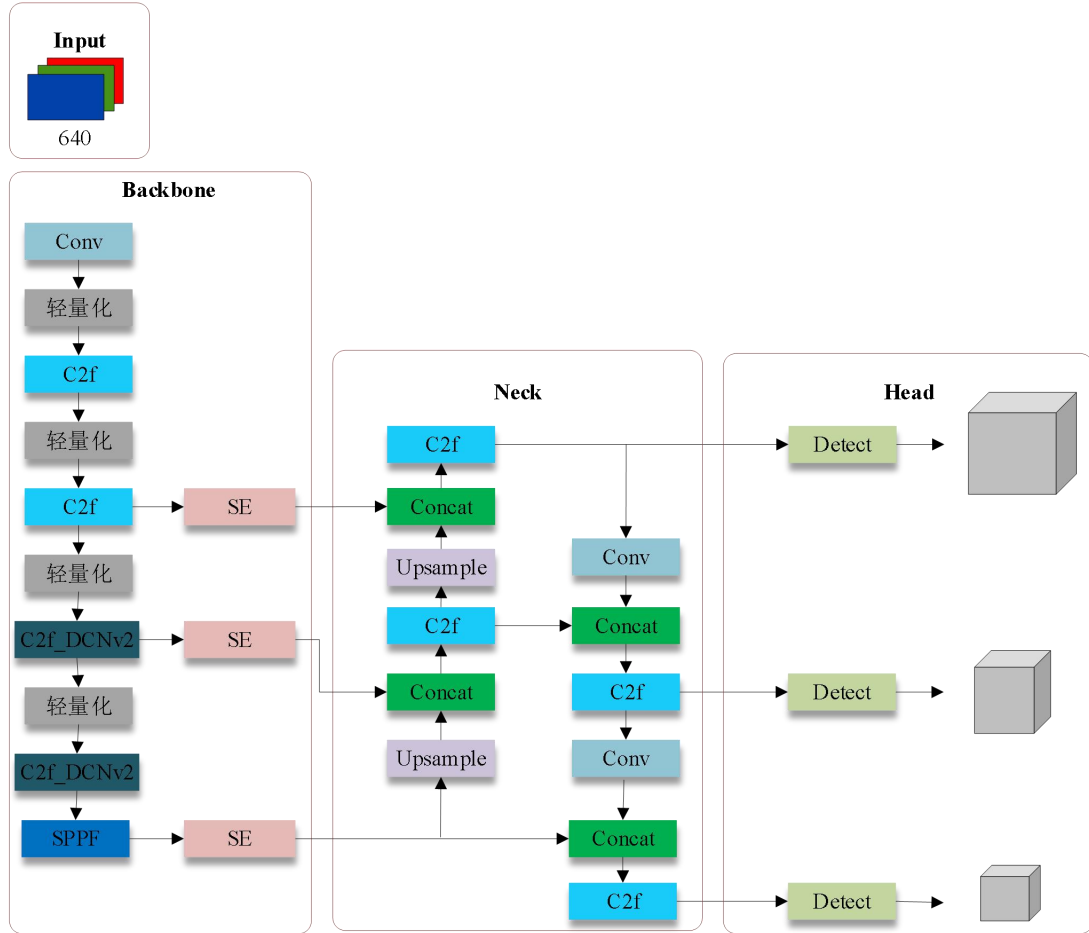


图 5-5 三种轻量化卷积模块添加位置

实验结果如表 5-1 所示。从表中数据可以分析出，GhostNet 轻量级卷积在使用部分特征图进行线性变换后，网络模型参数量个数从 3.1×10^6 降低到了 2.7×10^6 ，这也使网络模型对于缺陷检测的平均检测精度从 79.2% 降低到了 78.2%。深度可分离卷积模块在进行逐通道卷积和逐点卷积后，不仅使参数量降低到了 2.7×10^6 ，而且算法所需的计算量也因为卷积的特性大幅度减少，降低到了 6.7Gflops。DSConv 卷积通过使用 VQK 和分布偏移后，使模型的计算量也得到了大幅度的降低，并且在平均检测精度上为 78.4%，比其他两种轻量化方式更高。从整体上对数据进行分析可以看出，DSConv 卷积在平衡模型复杂度与检测精度方面比其他卷积更具有优势，因此将使用 DSConv 卷积轻量化后的

算法作为最终的改进算法，并命名为 YOLOv8n-Simple，为下面设计工业钢材表面缺陷检测系统提供算法上的支持。

表 5-1 各卷积模块数据

卷积模块	Gflops	Parameters	mAP(%)
Conv	7.8	3.1×10^6	79.2
GhostNet	7.2	2.7×10^6	78.2
DWConv	6.7	2.7×10^6	77.9
DSCConv	6.8	2.8×10^6	78.4

5.2 工业钢材表面缺陷检测系统设计

5.2.1 系统方案设计

本文设计的工业钢材表面缺陷检测系统，通过使用 PySide6 创建一个基本的 GUI（Graphical User Interface）界面，并结合 YOLOv8n-Simple 算法实现一些常用的功能。在输入模块中，使用者可以选择图片，视频或者摄像头三种检测模式；然后在系统模块中，选择合适的 YOLOv8 模型权重文件进行加载；最后在输出模块中会输出缺陷的类别，缺陷的置信度以及缺陷的个数。如图 5-6 是工业钢材表面缺陷检测系统架构图。

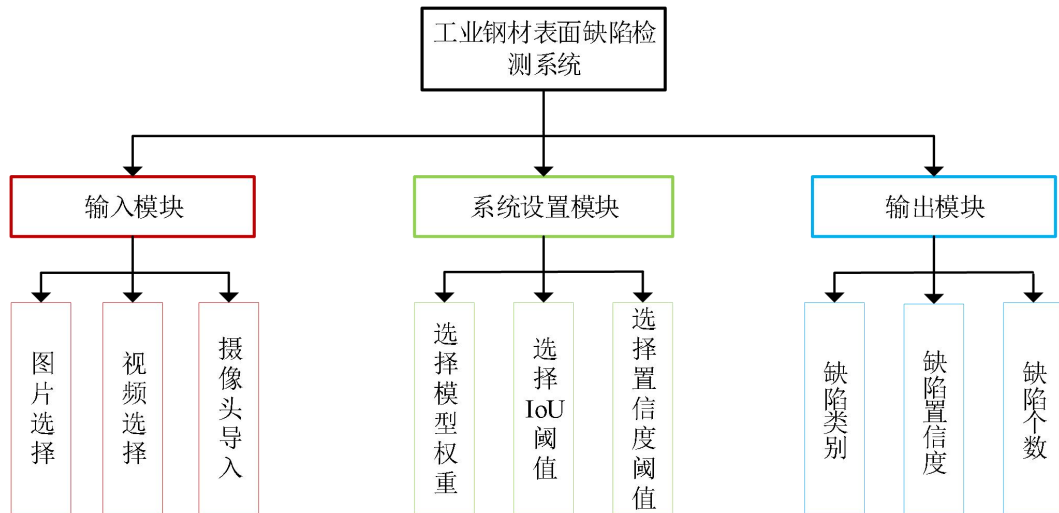


图 5-6 工业钢材表面缺陷检测系统架构图

5.2.2 系统开发环境

工业钢材表面缺陷检测系统运用 PyTorch 深度学习框架以及 PySide6 可视

化界面框架，其中各配置软件、硬件环境和开发使用的工具如表 5-2 所示。

表 5-2 系统开发环境

开发环境	工具
操作系统	Windows 10
CPU	Intel(R)Core(TM)i9-10900X CPU @3.70GHz
GPU	NVIDIA GeForce RTX3080
界面框架	PySide6
深度学习框架	PyTorch 2.0.0+Cuda 11.8
开发工具	Pycharm
开发语言	python

5.2.3 PySide6 简介

PySide6 是 Python 的一个 GUI 框架，是 Qt6 应用框架的 Python 绑定。它允许开发者使用 Python 语言来创建强大、跨平台的，实现与后端算法交互的界面。

PySide6 的特点包括：

(1) 跨平台性：通过编写一次代码就可以在 Windows、Linux 等多个主流操作系统上运行。

(2) 功能丰富：提供了众多的 GUI 组件，如窗口、按钮、菜单、对话框等，以及布局管理等工具，使开发者能够轻松创建复杂的用户界面。

(3) 与 Qt Designer 集成：开发者可以利用 Qt Designer 可视化的设计界面，并通过生成相应的 Python 代码来加速开发过程。

在 Qt Designer 的可视化界面中，可以选择一些按钮、窗口等进行工业钢材表面缺陷检测系统的界面搭建，将搭建好的界面保存并生成.ui 文件。在.ui 文件里，使用的控件会转化为 python 语言，避免了使用 python 进行编程来生成和控制按键的大小的复杂过程。然后在.ui 的文件中导入轻量化后的算法模型，并在按钮的槽函数中实现图像的获取和算法的检测逻辑，从而实现工业钢材表面缺陷检测系统的设计。

Qt Designer 设计界面如图 5-7 所示：

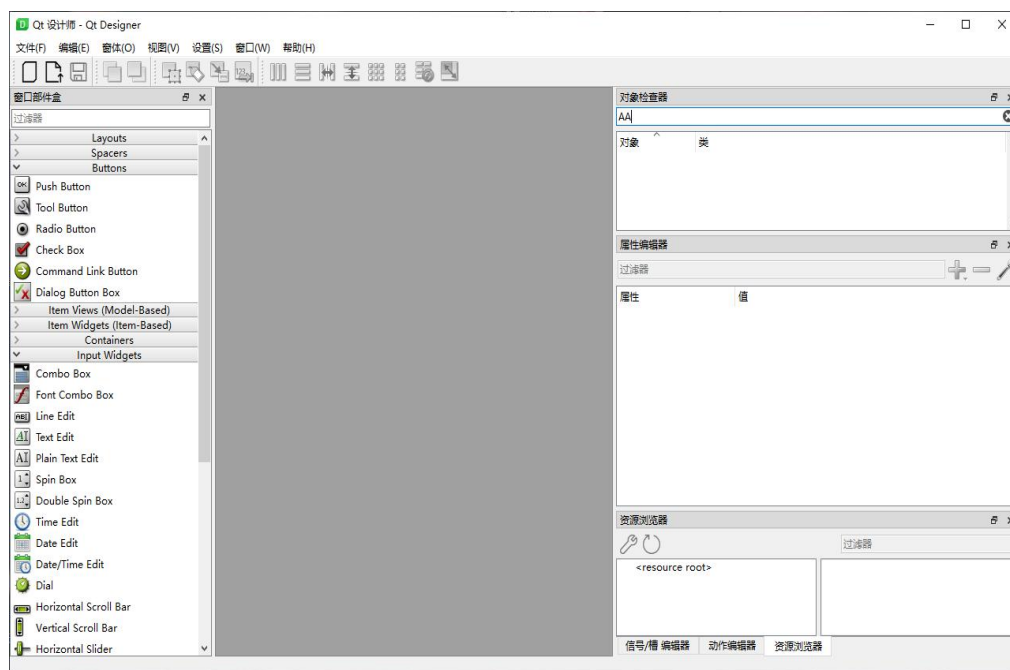


图 5-7 Qt Designer 设计界面

5.3 系统主界面设计

工业钢材表面缺陷检测系统完成后，在 Pycharm 平台上运行编写的代码，电脑屏幕中就会显示工业钢材表面缺陷检测的界面。工业钢材表面缺陷检测系统在检测时，可以直接通过调用 YOLOv8n-Simple 算法在训练过程中产生的最好权重文件进行缺陷检测，不需要再进行训练。当检测完成后，检测的结果图片会保存，然后显示到工业钢材表面缺陷检测的界面上，其工业钢材表面缺陷检测系统的运行界面如图 5-8 所示。

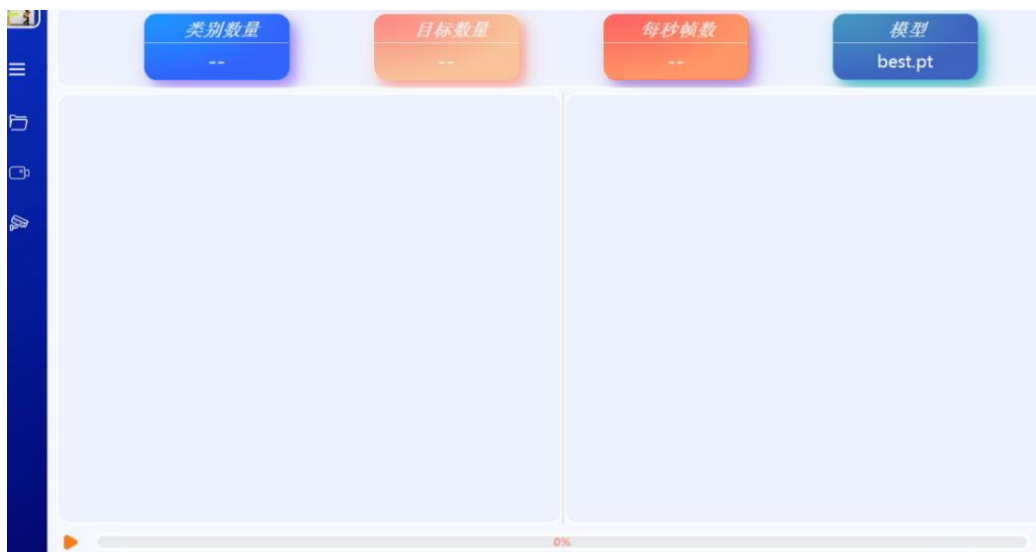


图 5-8 钢材表面缺陷检测系统界面

工业钢材表面缺陷检测系统界面的左侧是一个菜单栏，系统界面会提供三个选项，分别为打开上传的图片文件选项，打开上传的视频文件选项和打开摄像头选项。在工业钢材表面缺陷检测界面右上侧，可以选择模型的权重文件，如图 5-9 所示，这里使用的是 YOLOv8n-Simple 算法在训练过程中产生的最佳权重文件。



图 5-9 模型权重选择



图 5-10 裂纹缺陷检测

当检测钢材表面缺陷时，检测系统会直接使这个权重文件对钢材表面进行检测。同时设置 IoU 阈值为 0.45，当对钢材表面缺陷检测时，大于 IoU 阈值的预测框作为正样本，帮助模型学习到准确的特征。设置置信度阈值为 0.25，当对钢材表面缺陷检测时，检测精度高于 0.25 的缺陷会被标记出来。在进行检测后，系统界面上面就会显示检测到的工业钢材表面缺陷的类别个数，同时可以

显示检测到的工业钢材表面缺陷的个数，最后在工业钢材表面缺陷检测界面的中间空白处，会显示检测的图片以及检测后的结果图片。对其中的一张工业钢材表面图片进行检测，其中检测结果如图 5-10 所示。由检测的结果图片可以看出，检测到的缺陷类别有 1 种，其缺陷的名称为裂纹，检测到了两个裂纹缺陷，平均检测精度分别为 0.80 和 0.66。

5.4 系统功能测试

为了验证本章设计的工业钢材表面缺陷检测系统对其他缺陷检测的有效性，依次对各种缺陷也进行了检测，其检测结果如图 5-11 所示

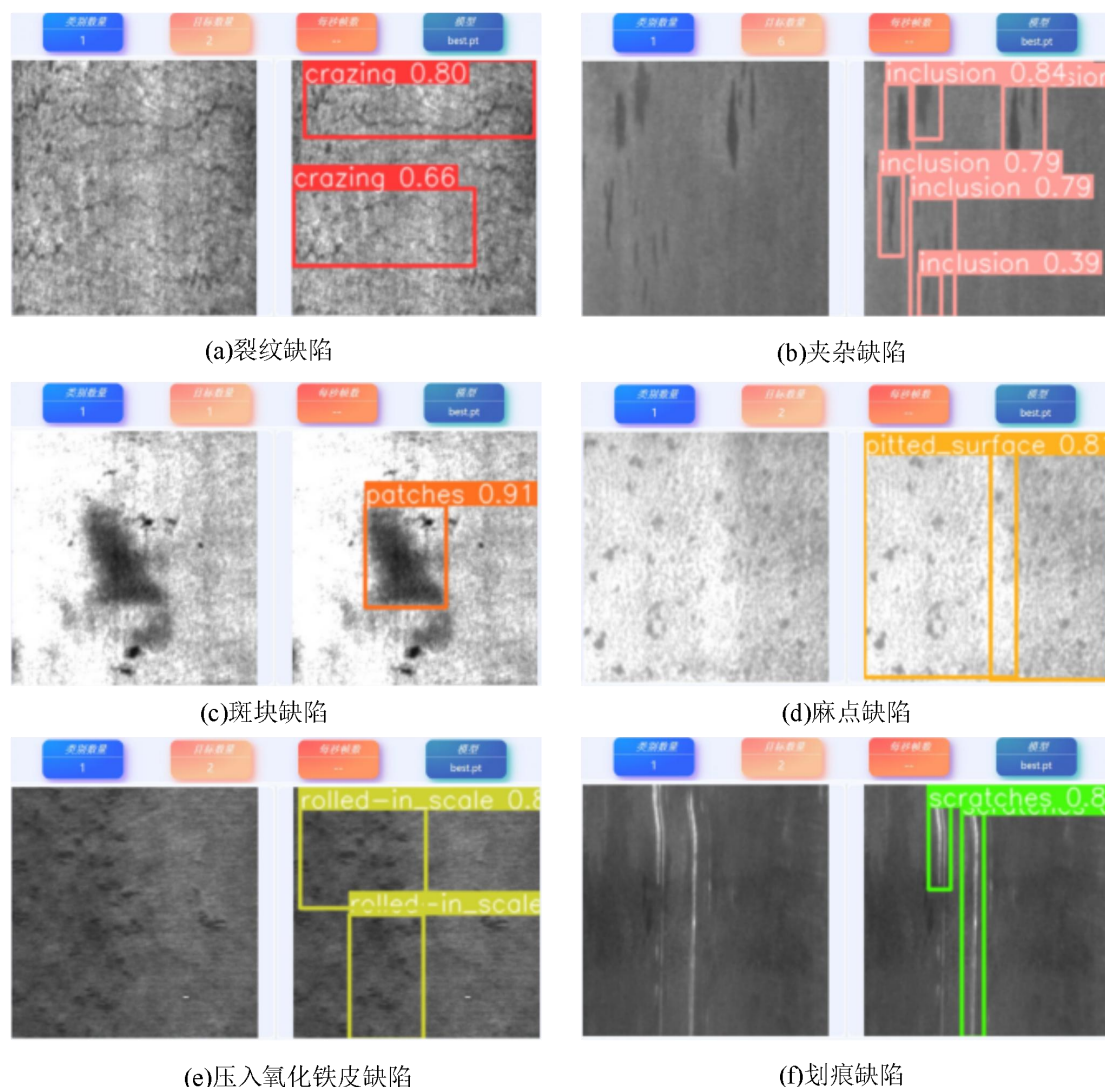


图 5-11 各缺陷类别检测结果

从图中可以看出，所有的钢材表面缺陷都可以被设计的工业钢材表面缺陷检测系统有效的检测，并且把工业钢材表面缺陷的类别数量、目标数量以及检

测前后缺陷的图片显示到界面中，这证明了设计的工业钢材表面缺陷检测系统是可行的。

5.5 本章小结

本章分别介绍了三种轻量化卷积，并分别添加到 YOLOv8n-Improved 算法网络模型中进行轻量化，然后进行了对比实验。通过对比实验发现，在使用 DSConv 卷积对 YOLOv8n-Improved 算法进行轻量化时，DSConv 卷积中的可变量化内核和分布偏移在降低网络模型的参数量和计算量时发挥了巨大作用，使得轻量化后的模型，在三种轻量化的方案中，对网络模型的参数量、计算量和平均检测精度平衡效果最好。使用 PySide6 与 YOLOv8n-Simple 算法模型，开发了一个工业钢材表面缺陷检测系统。系统包含数据输入模块、系统设置模块和输出显示模块。其中输入模块进行图片检测、视频检测或者摄像头检测，系统设置模块对需要使用的权重文件进行选择。输出模块会显示检测缺陷的类别数量和数目。最后，使用工业钢材表面检测系统对各类钢材表面缺陷进行检测，验证设计系统的可行性。

第 6 章 总结与展望

6.1 总结

在“中国制造 2025”战略中，制造业与新一代信息技术深度融合，工业钢材表面缺陷检测与深度学习技术融合到了一起。然而在融合过程中，基于深度学习的工业钢材表面缺陷检测，却一直面临钢材表面缺陷尺度变化大、缺陷与背景相似难以提取等难题。因此，本文以深度学习中的 YOLOv8n 算法模型为基础进行研究，针对性的提出了改进的 YOLOv8n 算法模型，提高了算法模型的检测性能。同时，为了将改进的算法运用到一些计算资源有限的工业嵌入式设备中，对算法进行了轻量化，并设计了工业钢材表面缺陷检测系统。具体工作内容如下：

(1) 对 YOLOv8n 模型进行各方面改进。首先，为了减少算法在训练过程中计算的预测结果与真实标注的差异，介绍了 GIoU、DIOU 和 CIoU 三种损失函数，对比分析了平均检测精度，并比较了三种损失函数运用在模型中的损失曲线，其中 CIoU 损失函数性能最好；针对缺陷的特征提取容易受到图片背景的干扰问题，介绍了 CBAM、ECA、SK、EMA、SE 五种注意力机制模块，分别添加到 YOLOv8n 网络模型中，并设计了不同位置的 3 种添加方案。在检测过程中，SE 注意力机制的重新分配权重，抑制不重要特征的功能在工业钢材表面缺陷检测中提升的检测精度最佳；针对工业钢材表面缺陷种类复杂，尺度变化大的问题，引入 DCNv2 模块，并与 YOLOv8n 算法模型中的 C2f 模块结合，提出新的 C2f_DCN2 模块，提高网络模型对工业钢材表面各种缺陷特征的提取。

(2) 提出基于改进 YOLOv8n 工业钢材表面缺陷检测算法研究。将各种改进方案的最佳改进方法添加到 YOLOv8n 算法模型中，确定最终改进的 YOLOv8n 算法模型；使用 LabelImg 工具检验 NEU-DET 数据集，建立工业钢材表面缺陷检测数据集；搭建工业钢材表面缺陷检测实验环境，对比分析改进的 YOLOv8n 算法模型与原算法模型的检测效果，改进的算法模型相较于 YOLOv8n 算法模型在 mAP 上提升了 1.9%，达到了 79.2%，并在一定程度上解决了原始模型的漏检、误检情况。

(3) 提出使用 GhostNet 卷积, 深度可分离卷积和 DSConv 卷积三种轻量化卷积, 分别对网络模型进行轻量化, 对比分析检测结果。其中 DSConv 卷积将卷积核分解为可变量化内核和分布偏移两部分, 在降低算法计算量和参数量的同时, 更好地平衡了算法的检测精度。以 DSConv 卷积轻量化的算法为基础, 利用 PySide6 设计了工业钢材表面缺陷检测系统。系统可以对图片文件、视频文件以及使用摄像头进行检测, 通过对含有各类缺陷的图片进行检测, 检测界面中可以成功的显示出检测到的缺陷类型、缺陷个数以及检测精度等各方面信息。

6.2 展望

本文对基于深度学习的工业钢材表面缺陷检测算法进行了深入研究, 构建了改进 YOLOv8n 算法模型和轻量化模型, 并设计了工业钢材表面缺陷检测系统。但是由于本人能力和研究时间有限, 研究还存在许多不足之处, 在未来工作中仍然存在一些内容需要完善:

(1) 本文虽然提高了工业钢材表面各种缺陷的检测精度, 但是裂纹缺陷和压入氧化铁皮缺陷的检测精度与其他缺陷的检测精度相比还是略低, 在后续的研究中可以针对这两种缺陷对算法模型进行改, 从而提高其检测精度。

(2) 本文采用的是目前工业场景中流行的 YOLOv8 算法, 但是在深度学习飞速发展的时代, YOLOv9、YOLOv10 算法也即将走向目标检测的舞台, 因此可以选择这些算法在后续工作中继续开展研究。

(3) 本文设计的工业钢材表面缺陷检测系统, 其检测功能较为简单, 而且由于实验资源和时间有限, 并没有把设计的系统部署到嵌入式设备中进行检测, 在未来的工作中可以完善系统的功能, 同时将系统部署到实际设备中去, 更好的满足工业需求。

参考文献

- [1] 韩宗志. “中国制造 2025”背景下我国传统制造业面临的挑战与对策[J]. 科技资讯, 2020, 18(13): 224-226.
- [2] 杨春政,袁天祥,刘延强,等.新一代钢铁制造流程产品品牌化实践[J].中国冶金,2024,34(12):24-35.
- [3] 刘亚楠. 基于机器视觉的钢材表面缺陷识别方法分析[J]. 计量与测试技术, 2018, 45(12): 97-98.
- [4] Abou-Khousa M A, Rahman M S U, Donnell K M, et al. Detection of surface cracks in metals using microwave and millimeter-wave nondestructive testing techniques—A review[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-18.
- [5] 李翰超, 古阳, 于亚婷. 基于涡流无损检测技术的钢材表面缺陷三维模拟分析[J]. 成都: 电子科技大学学报, 2020, 49(06): 942-948.
- [6] 邱怡乐. 基于线结构光扫描的金属件表面缺陷检测[D]. 成都: 电子科技大学, 2020.
- [7] Hu Y, Zhao D, Wang Y, et al. DAnoScenE: a driving anomaly scenario extraction framework for autonomous vehicles in urban streets[J]. Journal of Intelligent Transportation Systems, 2023: 1-21.
- [8] 方力. 遥感图像飞机目标检测与分类算法研究[D]. 武汉: 华中科技大学, 2019.
- [9] Ali K B, Abdalla A N, Rifai D, et al. Review on system development in eddy current testing and technique for defect classification and characterization[J]. IET Circuits, Devices & Systems, 2017, 11(4): 338-351.
- [10] Marinetti S, Vavilov V. IR thermographic detection and characterization of hidden corrosion in metals: General analysis[J]. Corrosion Science, 2010, 52(3): 865-872.
- [11] Yang L, Wang Z, Gao S. Pipeline magnetic flux leakage image detection algorithm based on multiscale SSD network[J]. IEEE Transactions on Industrial Informatics, 2019, 16(1): 501-509.
- [12] Baillie I, Griffith P, Jian X, et al. Implementing an ultrasonic inspection system to find surface and internal defects in hot, moving steel using EMATs[J]. Insight-Non-Destructive Testing and Condition Monitoring, 2007, 49(2): 87-92.

-
- [13] Cortes C, Vapnik V. Support-vector networks[J]. Machine Learning, 1995, 20: 273-297.
 - [14] Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an application to boosting[J]. Journal of Computer and System Sciences, 1997, 55(1): 119-139.
 - [15] Zhang J, Wang H, Tian Y, et al. An accurate fuzzy measure-based detection method for various types of defects on strip steel surfaces[J]. Computers in Industry, 2020, 122: 103231.
 - [16] Chaudhari C V. Steel surface defect detection using glcm, gabor wavelet, hog, and random forest classifier[J]. Turkish Journal of Computer and Mathematics Education (TURCOMAT),2021,12(12): 263-273.
 - [17] 肖玲, 李国民. 基于图像处理的钢材表面缺陷检测方法[J]. 电子设计工程, 2024, 32(21): 167-171.
 - [18] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
 - [19] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25: 1097-1105.
 - [20] 张阳婷, 黄德启, 王东伟, 等. 基于深度学习的目标检测算法研究与应用综述[J]. 计算机工程与应用, 2023, 59(18): 1-13.
 - [21] Wang D, Shang K, Wu H, et al. Decoupled R-CNN: Sensitivity-specific detector for higher accurate localization[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(9): 6324-6336.
 - [22] Fang S, Zhang B, Hu J. Improved Mask R-CNN Multi-Target Detection and Segmentation for Autonomous Driving in Complex Scenes[J].Sensors. 2023, 23(8):38-53.
 - [23] Hu F, Hu H, Xu H, et al. Dilated Heterogeneous Convolution for Cell Detection and Segmentation Based on Mask R-CNN[J].Sensors. 2024, 24(8):2424.
 - [24] Chang K C, Wang P J, Shuai H H, et al. A DeNoising FPN with Transformer R-CNN for Tiny Object Detection[J].IEEE Transactions on Geoscience and Remote Sensing. 2024,62:1-15.
 - [25] Gong X, Xie M, Xu H. Pedestrian detection algorithm based on yolov3[C]//2022 4th International Conference on Artificial Intelligence and Advanced

-
- Manufacturing (AIAM). IEEE, 2022: 686-689.
- [26] Zhang Z B, Wang K Q, Cui M D, et al. Improved YOLOv4 Visual Detection Method for Wild Bacteria[J].Computer Engineering and Applications,2023, 59(20): 228-236.
- [27] Li Z, Bai X, He C, et al. Apple Detection and Yield Estimation Based on YOLOv5[C]//2024 7th International Conference on Advanced Algorithms and Control Engineering (ICAACE). IEEE, 2024: 754-758.
- [28] Hu J, Sun Z, Li D. Logistics distribution and inspection of unmanned vehicles based on improved YOLOv5[C]//Ninth International Conference on Electromechanical Control Technology and Transportation (ICECTT 2024). SPIE, 2024, 13251: 1245-1254.
- [29] 李泽敏. 基于深度学习的钢板表面缺陷检测研究[D]. 杭州电子科技大学, 2023.
- [30] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. IEEE, 2017: 2980-2988.
- [31] Li K, Wang X, Ji L. Application of multi-scale feature fusion and deep learning in detection of steel strip surface defect[C]//2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM). IEEE, 2019: 656-661.
- [32] Zhang Y, Xiao F, Tian P. Surface defect detection of hot rolled steel strip based on image compression[C]//2020 International Conference on Computer Network, Electronic and Automation (ICCNEA). IEEE, 2020: 149-153.
- [33] Yang N, Guo W. Application of improved yolov5 model for strip surface defect detection[C]//2022 Global Reliability and Prognostics and Health Management (PHM-Yantai). IEEE, 2022: 1-5.
- [34] Huang B, Liu J, Liu X, et al. Improved YOLOv5 network for steel surface defect detection[J]. Metals, 2023, 13(8): 1439.
- [35] Wang L, Liu X, Ma J, et al. Real-time steel surface defect detection with improved multi-scale YOLO-v5[J]. Processes, 2023, 11(5): 1357.
- [36] Han J, Cui G, Li Z, et al. DBCW-YOLO: a modified YOLOv5 for the detection of steel surface defects[J]. Applied Sciences, 2024, 14(11): 4594.
- [37] Xie W, Sun X, Ma W. A light weight multi-scale feature fusion steel surface

- hr/>
- defect detection model based on YOLOv8[J].Measurement Science and Technology, 2024, 35(5): 055017.
- [38] Hao X, Dong T, Zhang D. A highly efficient surface defect detection approach for hot rolled strip steel based on deep learning[C]//2021 6th International Conference on Robotics and Automation Engineering (ICRAE). IEEE, 2021: 318-322.
- [39] 张政超. 改进 YOLOv5 的轻量级带钢表面缺陷检测[J]. 计算机系统应用, 2023, 32(06): 278-285.
- [40] Zhu C, Sun Y, Zhang H, et al. LE-YOLOv5: A lightweight and efficient neural network for steel surface defect detection[J]. IEEE Access, 2024,12:195242-195255.
- [41] Tie J, Zhu C, Zheng L, et al. LSKA-YOLOv8: A lightweight steel surface defect detection algorithm based on YOLOv8 improvement[J]. Alexandria Engineering Journal, 2024, 109: 201-212.
- [42] Zhang J, Yu X, Lei X, et al. A novel deep LeNet-5 convolutional neural network model for image recognition[J]. Computer Science and Information Systems, 2022, 19(3): 1463-1480.
- [43] Han J, Moraga C. The influence of the sigmoid function parameters on the speed of backpropagation learning[C]//International Workshop on Artificial Neural Networks. Berlin, Heidelberg: Springer Berlin Heidelberg, 1995: 195-201.
- [44] Zheng H, Wang B, Xiao M, et al. Adaptive friction in deep learning: Enhancing optimizers with sigmoid and tanh function[C]//2024 IEEE 6th International Conference on Power, Intelligent Computing and Systems (ICPICS). IEEE, 2024: 809-813.
- [45] Gurumoorthi T, Logesh P, Ismail N M, et al. Design of a Novel Back-Propagated Relu Convolutional Neural Networks for Pest Identification and Prediction[C]//2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC). IEEE, 2024: 228-232.
- [46] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2014: 580-587.
- [47] Girshick R. Fast R-CNN[C]. Proceedings of The IEEE International Conference on Computer Vision. IEEE, 2015: 1440-1448.

-
- [48] Ren S, He K, Girshick R, et al. Faster R-cnn:Towards Real-time Object Detection with Region Proposal Networks[J]. Advances in Neural Information Processing Systems, 2015, 39(6): 1137-1149.
 - [49] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 779-788.
 - [50] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017: 7263-7271.
 - [51] Farhadi A, Redmon J. Yolov3: An incremental improvement[C]//Computer Vision and Pattern Recognition. Berlin/Heidelberg, Germany: Springer, 2018, 1804: 1-6.
 - [52] Yang G, Feng W, Jin J, et al. Face mask recognition system with YOLOV5 based on image recognition[C]//2020 IEEE 6th International Conference on Computer and Communications (ICCC). IEEE, 2020: 1398-1404.
 - [53] 赵林熔. 基于改进 YOLOv5s 算法的钢材表面缺陷检测模型研究[D]. 中北大学, 2024.
 - [54] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3-19.
 - [55] Wang Q, Wu B, Zhu P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 11534-11542.
 - [56] Li X, Wang W, Hu X, et al. Selective kernel networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 510-519.
 - [57] Ouyang D, He S, Zhang G, et al. Efficient multi-scale attention module with cross-spatial learning[C]//ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2023: 1-5.
 - [58] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 7132-7141.
 - [59] Dai J, Qi H, Xiong Y, et al. Deformable convolutional networks[C]//Proceedings

- of the IEEE International Conference on Computer Vision. IEEE, 2017: 764-773.
- [60] Zhu X, Hu H, Lin S, et al. Deformable convnets v2: More deformable, better results[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 9308-9316.
- [61] 李相垚, 侯红玲, 杨澳, 等. 面向钢材表面缺陷检测的 DCSYOLOv8 算法研究[J/OL]. 机械科学与技术, 1-10[2024-11-07].
- [62] 梁礼明, 陈康泉, 钟奕, 等. DCD-YOLOv8n: 一种高效的钢材表面缺陷检测算法[J]. 计算机工程与应用, 2025, 61(07): 117-127.
- [63] 王孟宇, 刘志强. 改进 YOLOv8 算法的钢材表面缺陷检测[J]. 机械科学与技术, 2025, 44(01): 19-29.
- [64] Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks[C]//International Conference on Machine Learning. PMLR, 2019: 6105-6114.
- [65] Han K, Wang Y, Tian Q, et al. Ghostnet: More features from cheap operations[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 1580-1589.

攻读学位期间主要研究成果

- [1] 李晓宇,瞿成明,单传辉.基于轻量化 YOLOv5s 的带钢表面缺陷检测算法[J].安徽工程大学学报,2024,39(05):48-56.

致谢

行文至此，墨迹将干，三年研究生生涯亦悄然步入终章。抚今追昔，万千思绪涌上心头，既有对过往岁月的深切眷恋，更饱含对所有帮助过我的师长、同窗与家人的诚挚谢意。

首先，谨向我的导师瞿老师致以最崇高的敬意。求学期间，瞿老师以渊博的学识、严谨的治学态度，在学术研究的道路上为我指明方向。从课题的选题立意，到研究方法的指导，再到论文的反复雕琢，瞿老师始终以精益求精的精神，给予我悉心教诲与殷切关怀。单老师在实验设备方面的鼎力支持，为研究的顺利开展提供了坚实保障，这份雪中送炭的情谊，令我感念至深。

其次，衷心感谢同窗挚友的一路相伴。特别感谢室友王坤相，在实验室的无数个日夜，我们相互切磋、共同进步；在生活中，我们彼此关照、分享悲喜。还有其他同学与朋友，在学术讨论中，我们思维碰撞、各抒己见；在面临困难时，我们携手并肩、共克时艰。这段共同奋斗的时光，已然成为我人生中最珍贵的记忆。

最后，要感谢始终未曾放弃的自己。求学之路，道阻且长，正是这份坚持与执着，支撑我跨越重重难关。更要感恩家人的无私支持，他们的理解与鼓励，是我在逐梦路上最温暖的港湾。

师恩如山，情谊似海，亲情如歌。在此，再次向所有关心、帮助过我的人致以最诚挚的谢意！未来的道路上，我将铭记这份恩情，砥砺前行，不负韶华。