

基于改进YOLOv11-EB的芯片表面缺陷检测方法

刘靛颖, 巴 鹏, 张秀珩, 刘杨子, 张译夫, 华子懿

沈阳理工大学机械工程学院, 辽宁 沈阳

收稿日期: 2026年6月29日; 录用日期: 2026年7月22日; 发布日期: 2026年7月29日

摘 要

针对芯片缺陷检测过程中缺陷特征相似、缺陷目标小、缺陷尺度差异大造成检测精度低等问题, 提出一种基于改进YOLOv11的芯片表面缺陷检测模型YOLOv11-EB。首先采用EfficientNetV2替代原YOLOv11骨干网络, 通过其更强的特征提取能力增强细粒度信息表达, 从而有效缓解缺陷特征相似、小目标易漏检以及复杂背景下误检的问题。其次将特征融合层(Neck)替换为加权双向特征金字塔网络(BiFPN), 以提升多尺度特征融合效果。实验结果表明, 该方法在复杂背景和多尺度缺陷条件下, 能够显著降低误检率与漏检率, 并有提高模型的检测精度与召回率。与原模型相比, 改进模型在精确率、召回率和平均精度均值分别提升了5.3%、11.9%和7.3%。

关键词

芯片检测, 缺陷分类, YOLOv11, 小目标检测

Chip Surface Defect Detection Method Based on Improved YOLOv11-EB

Liangying Liu, Peng Ba, Xiuheng Zhang, Yangzi Liu, Yifu Zhang, Ziyi Hua

School of Mechanical Engineering, Shenyang Ligong University, Shenyang Liaoning

Received: June 29, 2026; accepted: July 22, 2026; published: July 29, 2026

Abstract

To address the issues of low detection accuracy in chip surface defect inspection caused by highly similar defect features, small target sizes, and large scale variations, this paper proposes an improved YOLOv11-based model, termed YOLOv11-EB. First, EfficientNetV2 is adopted to replace the original backbone network of YOLOv11, leveraging its superior feature extraction capability to

文章引用: 刘靛颖, 巴鹏, 张秀珩, 刘杨子, 张译夫, 华子懿. 基于改进 YOLOv11-EB 的芯片表面缺陷检测方法[J]. 建模与仿真, 2026, 15(7): 121-130. DOI: 10.12677/mos.2026.157113

enhance fine-grained feature representation. This effectively alleviates the challenges of feature similarity, missed detections of small defects, and false detections in complex backgrounds. Second, the original feature fusion network (Neck) is replaced with a weighted Bidirectional Feature Pyramid Network (BiFPN) to improve multi-scale feature fusion. Experimental results demonstrate that the proposed method significantly reduces both false positive and false negative rates under complex background and multi-scale defect conditions, while improving overall detection performance. Compared with the original model, the proposed model achieves improvements of 5.3%, 11.9%, and 7.3% in precision, recall, and mean Average Precision, respectively.

Keywords

Chip Inspection, Defect Classification, YOLOv11, Small Object Detection

Copyright © 2026 by author(s) and Hans Publishers Inc.

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

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



Open Access

1. 引言

在芯片制造过程中，各道工艺紧密衔接，任何技术条件、材料特性、生产环境或工艺参数的细微波动，都可能引发芯片缺陷，从而降低产品良率。芯片缺陷检测是生产线质量把控的关键一环，既能反馈产品品质，也能实时管控生产工序，在半导体行业里地位越来越关键。如今封装芯片日趋微型化，产品质检标准随之不断收紧。人工目检受精度、效率和评判标准限制逐渐被淘汰，传统机器视觉易受环境干扰，需要人工手动提取特征，适用性不足。但近年来，由深度学习架构推动的目标检测技术迅速发展，取得了显著成就[1]。进一步优化深度学习模型，提高检测精度、泛化能力和计算效率仍是当前研究的重要方向[2]。文献[3]芯片表面缺陷检测在工业生产中至关重要，现有注意力机制因方向敏感性不足，难以有效获取破损、裂纹等有向线状表面缺陷。提出了一种基于 YOLOv11n 改进的双向多尺度检测网络 (DMNet) 模型，该模型设计了基于方向解耦建模的双向注意力 (DDA) 机制与多尺度特征增强 (MSFE) 模块。刘琛等提出一种基于方位环境特征的点模式匹配定位算法，快速准确定位芯片并识别引脚缺陷[4]。陈文凤等设计了一套基于 ARM-DSP 双核结构的集成芯片引脚缺陷自动检测系统[5]。李本红等采用灰度跃变检测引脚中点、中点直线拟合和引脚间距统计等方法，实现对 SOP 芯片引脚缺陷的自动检测[6]；李绎铃以连通像素区域标记法为主要算法，提出基于模板匹配的引脚缺陷识别算法[7]；金贺楠通过图像矩来对 SOP 芯片的整体位置进行定位，同时结合基于 TwoPass 的连通域分析方法来完善芯片引脚外观检测[8]。文献[9]对 YOLOv5 进行多方面改进，有效提升了微小缺陷检测精度，且模型满足工业实时检测需求；文献[10]提出的改进 YOLOv8 模型，在轻量化与检测精度上实现双重优化，但难以精准识别边缘模糊缺陷。文献[11]设计了 YOLO-RRL 模型，通过稳健特征下采样及广义重参数化策略压缩模型、优化实时性能。贺智轩等[12]将自调节特征聚合模块与 YOLOv11 中的 C3K2 模块融合形成 C3K2_FMB 模块，在空间和通道维度提高了特征表达，以及重新构造局部和非局部特征去完成高质量的输出。孙佩月[13]等提出一种基于 YOLO11n 的轻量化芯片表面缺陷检测方法，来解决芯片表面缺陷具有尺寸微小、形状不规则、类型多样等特点。尽管上述方法在芯片缺陷检测中取得了一定进展，但仍存在对小目标缺陷检测能力不足、特征相似类别易混淆等问题。为此，本文提出一种基于 YOLOv11-EB 的检测方法，使用 EfficientNetV2 骨干网络替代 YOLOv11 的骨干网络，将特征融合层替换为加权双向特征金字塔网络，针对小目标检测提升模型整体检测精度。

2. 数据集制作

2.1. 数据集构建

该研究为实现芯片表面缺陷分类任务，构建芯片缺陷类型的数据集(Chip-Classification)。数据集采集来自实验室芯片分选机，选择不同的场景，不同角度，不同曝光度条件下拍摄。图中有五种不同类型芯片，其中四种为表面缺陷类型，脏污芯片(dirty)、引脚弯曲(bent_leg)、引脚损坏(damaged_leg)、芯片划痕(scratches)；一种为正常芯片(normal_leg)。部分数据如图 1 所示。

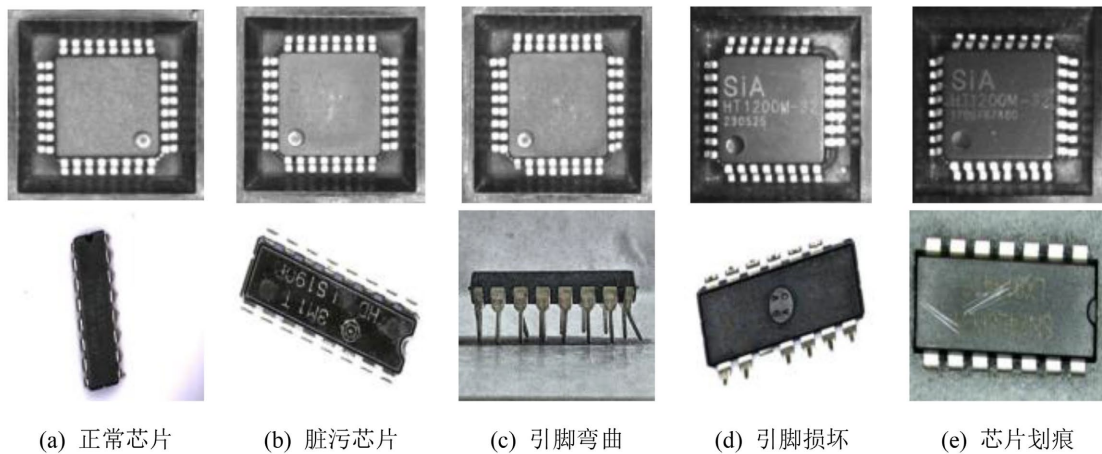


Figure 1. Dataset examples

图 1. 数据集实例

2.2. 数据预处理

经芯片分选机共采集到 535 张芯片图像数据，但其中正常芯片 305 张，引脚弯曲芯片 103 张，引脚损坏芯片 88 张，芯片划痕 23 张，脏污芯片仅有 16 张。脏污芯片和芯片划痕类别在数据上与正常芯片有着超过十倍的差距，这种极度不平衡会导致模型在训练过程中偏向于样本量较大的类别，而小样本类别会因学习不充分，模型难以提取到有效的特征表示，产生欠拟合或者过拟合，从而降低整体的识别准确率。将所有的图像分辨率统一缩放到 640×640 像素，并将数据集进行旋转、翻转、噪声、模糊等数据增强的手段进行类别平衡。经处理后的数据集得到 1268 张芯片图像。其中正常芯片 305 张，引脚弯曲芯片 309 张，引脚损坏芯片 264 张，芯片划痕 230 张，脏污芯片有 160 张。并按照 8:2 的比例划分训练集和验证集(训练集 1014 张，验证集 254 张)。随后用 labelme 进行芯片表面缺陷标注。图像标注实例如图 2 所示。

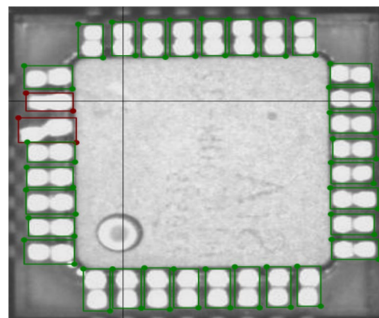


Figure 2. Image annotation examples

图 2. 图像标注实例

3. 改进 YOLOv11 检测模型

3.1. YOLOv11 网络模型结构

R. Khanam 等人[14]提出 YOLO11 模型, YOLOv11 是在 YOLOv8 基础上进一步优化得到的单阶段目标检测模型, 其整体结构由骨干网络(Backbone)、颈部网络(Neck)和检测头(Head)构成, 同时在特征表达能力与计算效率之间进行了更精细的权衡。

在骨干网络部分, 引入了改进的 C2PSA 模块以增强特征提取能力。该模块在原有 C2 结构基础上融合了部分空间注意力机制(Partial Spatial Attention, PSA), 通过对特征图进行区域划分, 并为不同区域分配独立的权重, 实现对关键空间信息的自适应建模; 通过堆叠多个 C2PSA 单元, 模型能够更加精准地捕捉目标的空间分布特征, 从而提升在复杂背景下的检测表现。

在颈部网络部分, 沿用 PANet (Path Aggregation Network)的结构, 实现不同层级特征的双向融合, 增强了模型对不同尺度目标的检测能力。采用 C3K2 模块替代传统的 C2f 模块, 以提升特征融合阶段的效率与表达能力。C3K2 在 CSP (Cross Stage Partial)结构基础上进行改进, 通过通道分离策略减少冗余计算, 同时结合双卷积结构与可变卷积核设计, 使得特征在不同尺度之间的传递更加高效。这种设计不仅增强了特征复用能力, 还在保证模型轻量化的同时提升了整体性能。

在检测头部分, 采用解耦结构设计, 对融合特征进行预测输出。将分类与定位任务分离处理, 从而降低任务间的相互干扰, 提高检测精度与收敛稳定性。YOLOv11 网络结构图如图 3 所示。

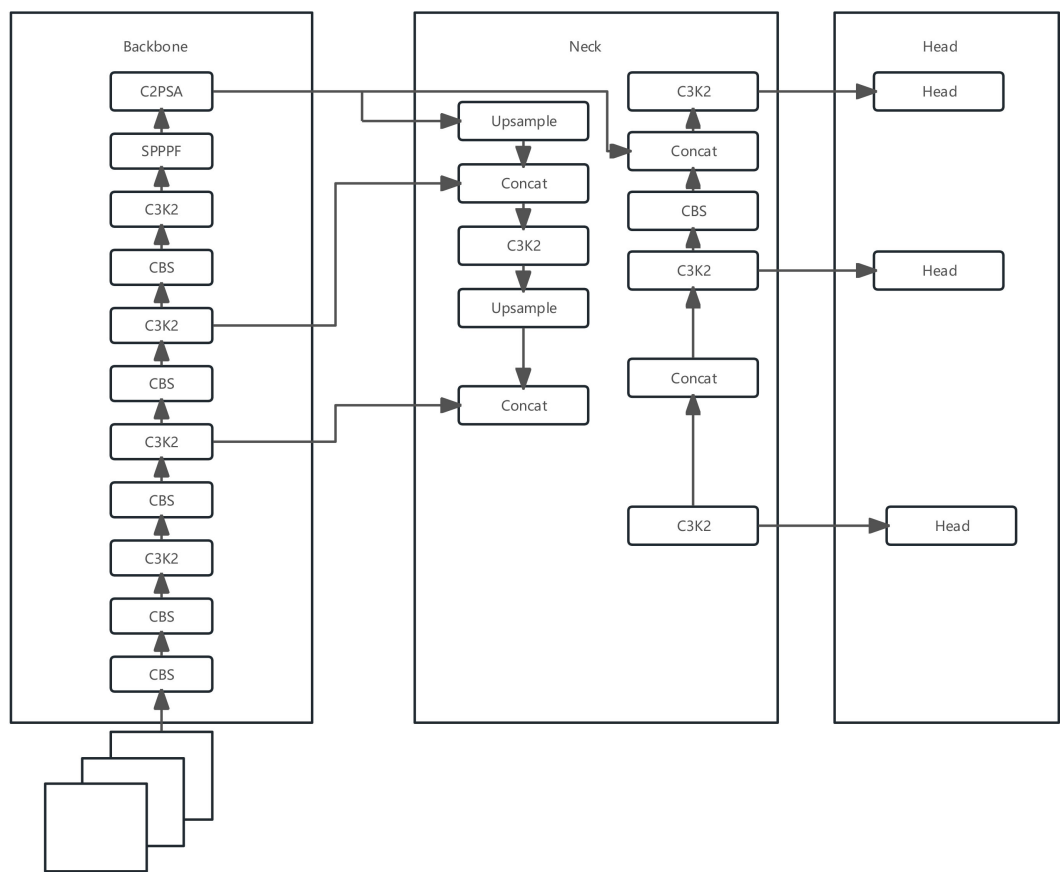


Figure 3. Network structure diagram of YOLOv11 model
图 3. 模型网络结构图

3.2. 改进模型 YOLOv11-EB

本研究在 YOLOv11 模型基础上进行改进, 将 EfficientNetV2 骨干网络替换 YOLOv11 的骨干网络, 以提升细粒度特征表达能力, 从而有效改善芯片表面缺陷检测中缺陷特征相似、小目标漏检、复杂背景误检以及小样本训练不稳定等问题; 将原有的特征融合结构替换为加权双向特征金字塔网络(BiFPN), 该结构通过引入自顶向下与自底向上的双向信息流, 并结合可学习的特征融合权重, 实现了不同尺度特征的高效整合, 有效提升了模型对微小缺陷的检测能力、缺陷尺度差异大并在复杂背景下显著降低了误检率与漏检率, 提升了准确度和召回率。改进后的模型整体结构如图 4 所示。

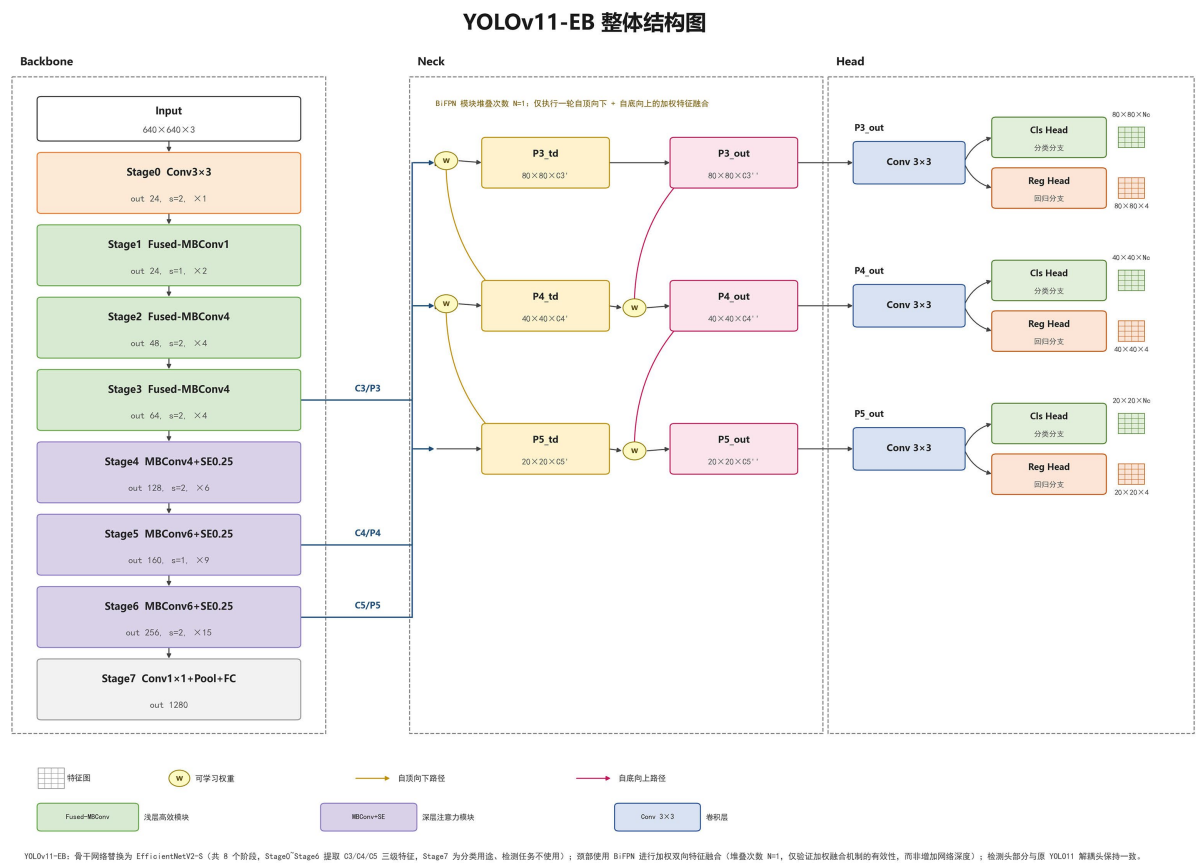


Figure 4. Structural sketch of improved YOLOv11-EB model

图 4. 构造图改进模型 YOLOv11-EB

在 YOLOv11-E 中。在 YOLOv11-B 特征融合阶段, 采用 BiFPN 替代原始 YOLO11 的 Concat 拼接操作。BiFPN 通过为每个输入特征图分配可学习的权重来实现加权特征融合。在本实验中, BiFPN 模块的堆叠次数设置为 1, 即仅执行一轮自顶向下和自底向上的双向特征融合, 随后直接送入检测头。旨在验证加权特征融合机制的有效性, 而非增加网络深度。

3.3. BiFPN 特征融合网络模块

BiFPN 通过引入可学习权重, 对不同输入特征的重要性进行自适应建模, 并多次执行自顶向下与自底向上的双向特征融合, 从而实现高效的多尺度信息交互。其结构主要包括自适应特征选择模块与自适应特征融合模块。其中, 自适应特征选择模块结合注意力机制, 有针对性地强化关键特征表达。骨干层

需要大量的卷积来提取足够的图像特征，标准卷积通道之间需要高度相关[15]。BiFPN 引入了权重信息进行特征融合，这些权重是网络学习的参数。计算公式如所示。

$$P_{6td} = \text{Conv} \left(\frac{w_1 \cdot P_{6in} + w_2 \cdot \text{resize}(P_{7in})}{w_1 + w_2 + \varepsilon} \right) \quad (1)$$

式中： P 为特征金字塔中的特征图；下标 6 为金字塔的第 6 层；下标 td 为 top-down 的缩写； P_{6td} 为自顶向下路径中生成的中间特征图； w_1 、 w_2 为代表融合时各自输入所对应的可学习权重； P_{6in} 为第 6 层的输入特征图(in 表示 input)； P_{7in} 为第 7 层的输入特征图； $\text{resize}(P_{7in})$ 为尺寸调整操作； ε 为一个极小的正数，用来保证分母不为零，从而保持数值稳定性。

$$P_{6out} = \text{Conv} \left(\frac{w_{1'} \cdot P_{6in} + w_{2'} \cdot P_{6td} + w_{3'} \cdot \text{resize}(P_{5out})}{w_{1'} + w_{2'} + w_{3'} + \varepsilon} \right) \quad (2)$$

式中： P_{6out} 为第 6 级特征图经过融合和卷积后的最终输出特征； $w_{1'}$ 、 $w_{2'}$ 、 $w_{3'}$ 为可学习的标量权重； P_{6td} 为第 6 级的自顶向下路径特征； P_{6in} 为第 6 级的原始输入特征； $\text{resize}(P_{5out})$ 为来自第 5 级的输出特征，经过尺寸调整后，与第 6 级特征图的分辨率对齐。

此外，BiFPN 通过为每个输入特征图分配可学习的权重来实现加权特征融合。在 YOLOv11-B 实验中，BiFPN 模块的堆叠次数设置为 1，即仅执行一轮自顶向下和自底向上的双向特征融合，随后直接送入检测头。旨在验证加权特征融合机制的有效性，而非增加网络深度。

得益于上述改进，模型在复杂背景和尺度变化显著的缺陷检测任务中，能够有效降低误检与漏检率显著提升检测精度与鲁棒性。

3.4. EfficientNetV2 网络模块

EfficientNetV2 在整体结构上延续了卷积神经网络的分阶段设计思想，但在基本模块与训练策略方面进行了系统优化。在网络单元设计上，引入 Fused-MBConv 模块替代部分传统 MBConv 结构。该模块将逐点卷积与深度可分离卷积进行融合，有效减少内存访问开销，从而显著提升训练效率。Fused-MBConv 普通卷积融合和输出公式为：

$$X_f = \text{Conv}_{3 \times 3}(X) \quad (3)$$

式中： X 为输入特征图(input feature map)； X_f 为输出特征图(output feature map)； $\text{Conv}_{3 \times 3}$ 为卷积核大小为 3×3 的标准卷积操作。

$$Y = \text{Conv}_{1 \times 1}(X_f) \quad (4)$$

式中： X_f 为输入特征图(input feature map)； $\text{Conv}_{1 \times 1}$ 为卷积核大小为 1×1 的标准卷积操作； Y 为输出特征图(output feature map)。

在浅层网络中采用 Fused-MBConv，以提高计算效率；在深层网络中保留 MBConv 结构，以增强特征表达能力。这种“浅层高效、深层精细”的设计实现了速度与精度之间的良好平衡。Depthwise 卷积公式：

$$X_d = \text{DWConv}(X_e) \quad (5)$$

式中： X_e 为输入特征图；DWConv 为深度卷积操作(Depthwise Convolution)； X_d 为输出特征图；是深度卷积处理后的结果。

而在深层网络中，仍然使用 MBConv 结构以保持较高的表达能力。这种“浅层高效、深层精细”的

设计,使模型在速度与精度之间取得更优平衡。MBConv 核心公式为:

$$X_e = \text{Conv}_{1 \times 1}(X) \quad (6)$$

式中: X 为输入特征; X_e 为扩张后特征。

在 YOLOv11-E 实验中采用 EfficientNetV2-S 作为骨干网络,该网络共包含 8 个阶段(Stage 0~7),各阶段的具体配置: Stage 0 为 $\text{Conv}_{3 \times 3}$ (输出通道 24, 步长 2, 1 层); Stage 1 为 Fused-MBConv1 (输出通道 24, 步长 1, 2 层); Stage 2 为 Fused-MBConv4 (输出通道 48, 步长 2, 4 层); Stage 3 为 Fused-MBConv4 (输出通道 64, 步长 2, 4 层); Stage 4 为 MBConv4 with SE0.25 (输出通道 128, 步长 2, 6 层); Stage 5 为 MBConv6 with SE0.25 (输出通道 160, 步长 1, 9 层); Stage 6 为 MBConv6 with SE0.25 (输出通道 256, 步长 2, 15 层); Stage 7 由 $\text{Conv}_{1 \times 1}$ 、全局池化和全连接层组成(输出通道 1280)。

在模型缩放策略方面, EfficientNetV2 对传统复合缩放方法进行了改进,采用分阶段缩放策略,使网络各阶段能够根据任务需求灵活调整,从而提高计算资源利用率。

此外,该模型引入渐进式学习(Progressive Learning)策略,在训练初期使用低分辨率与弱增强以加速收敛,随后逐步提高输入分辨率与增强强度,从而在保证训练效率的同时提升模型的泛化能力。

4. 实验及结果分析

4.1. 实验参数设置

计算机硬件方面,采用配备 NVIDIA RTX5060 (16 GB)显卡和 AMD R9-7845HX 处理器的设备进行模型训练,以满足深度学习任务对算力和显存的需求。

软件环境方面,实验基于 Python3.8.10 搭建开发环境,并使用 PyTorch2.8.0 + cu128 作为深度学习框架,以支持 GPU 加速计算。

模型训练配置方面,输入图像尺寸统一设定为 640×640 像素,批量大小(batchsize)为 8,数据加载线程数设置为 2。初始学习率设定为 0.01,动量因子设定为 0.937,权重衰减系数为 0.0005,用于正则化防止过拟合,优化器选择 SGD (随机梯度下降)。模型总训练轮次为 150 轮(epochs),并在训练的最后 10 轮关闭马赛克(Mosaic)数据增强,以提升模型在后期训练阶段的稳定性和收敛效果。

4.2. 评价指标

在目标检测领域中,混淆矩阵是一种常用的分类性能评估工具。它以矩阵形式对比模型预测结果与真实标签之间的关系,能够直观反映模型在不同类别上的判别情况及错误类型。对于二分类任务,样本通常被划分为正类(Positive)和负类(Negative)。根据预测结果与真实标签的一致性,可分为四种情况:预测正确的称为“真”(True),预测错误的称为“假”(False)。由此可得到真正例(True Positive, TP)、假正例(False Positive, FP)、真负例(True Negative, TN)以及假负例(False Negative, FN)四类结果,这四种情况共同构成一个 2×2 的混淆矩阵,如表 1 所示。

Table 1. Confusion matrix

表 1. 混淆矩阵表

	预测为正例	预测为负例
实例为正例	TP (真正例)	FN (假负例)
实例为负例	FP (假正例)	TN (真负例)

根据混淆矩阵中的四项基础统计量——真正例(TP)、假负例(FN)、假正例(FP)及真负例(TN),可进一

步推导出若干关键评估指标，用以量化模型的实测表现。

其中精确度(P)、召回率(R)、平均精度(AP)、平均精度均值(mAP)的计算公式为：

$$P = \frac{TP}{TP + FP} \tag{7}$$

式中：TP 实际是正例，模型也预测为正例的样本数量；FP 实际是负例，模型却错误预测为正例的样本数量；TP + FP 是所有被模型预测为正例的样本总数。

$$R = \frac{TP}{TP + FN} \tag{8}$$

式中：FN 实际是正例；但模型错误预测为负例的样本数量；TP + FN 是所有真实为正例的样本总数。

$$AP = \int_0^1 P(R) dR \tag{9}$$

式中： $P(R)$ 是一个以召回率 R 为自变量、精确度 P 为因变量的函数。

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \tag{10}$$

式中： N 代表缺陷类别总数

4.3. 对比试验

为了保证本研究结果的有效性和参考性，实验采取 FasterR-CNN 和两款主流 YOLOv8、YOLO5 作为对比模型。如表 2 所示。

Table 2. Comparative experiments
表 2. 对比试验

检测模型	$P/\%$	$R/\%$	$mAP50/\%$
FasterR-CNN	79.6	85.3	69.4
YOLOv5	74.3	75.8	79.6
YOLOv8	79.8	83.7	86.3
YOLOv11	85.6	76.5	83.2
YOLOv11-EB	90.9	88.4	90.5

试验结果表明，FasterR-CNN 作为两阶段目标检测的经典模型，在检测任务中展现出了较高的召回率(85.3%)，基础目标发现能力较强，但其精确度仅为 79.6%，mAP50 低至 69.4%，可见误报率较高、综合检测性能明显不足的短板。在 YOLO 系列模型中，YOLOv5 的精确度(74.3%)和召回率(75.8%)均处于较低水平，mAP50 仅为 79.6%，整体检测能力较弱；YOLOv8 召回率(83.7%)较高，且高于 YOLOv11，这主要是由于 YOLOv11 在召回率上表现偏弱，但 YOLOv8 精确度却只有 79.8%；YOLOv11 精确度高达 85.6%，但其召回率 76.5%，出现了漏检问题，导致 mAP50 至 83.2%。综合来看，YOLOv8 在 mAP50 指标上表现更优，而 YOLOv11 在精确度方面具有一定优势，但由于召回率偏低，整体性能仍存在提升空间。通过改进模型 YOLOv11-EB 通过多层次特征融合与骨干网络的协同优化，实现了检测性能的全面跃升，其精确度、召回率与 mAP50 分别达到 90.9%、88.4%与 90.5%，在保证极低漏检与误检的同时，具有综合优势，充分验证了该改进策略的有效性与优越性。

4.4. 骨干网络对比

为全面评估不同骨干网络对 YOLOv11 芯片表面缺陷检测准确度性能的影响，本实验从检测精度、

召回率、Map50 三个维度，选取 EfficientNetV2、ShuffleNetV2 及 MobileNetV4 三种典型骨干网络进行对比实验。如表 3 所示。

Table 3. Backbone network comparison

表 3. 骨干网络对比

模型	<i>P</i> /%	<i>R</i> /%	mAP50/%
ShuffleNetV2	80.2	75.8	79.8
EfficientNetV2	88.5	77.2	84.8
MobileNetV4	76.5	75.9	79.4

实验结果表明，ShuffleNetV2 的精确度为 80.2%，mAP50 为 79.8，但召回率仅 75.8%，整体表现不够突出，存在漏检。MobileNetV4 的精确度仅 76.5%，mAP50 为 79.4%，均为最低，召回率 75.9%也处于较低水平，检测性能明显偏弱，难以达到对精度有较高要求的检测要求。EfficientNetV2 凭借最高的精确度(88.5%)与最优的 mAP50 (84.8%)，召回率也达到 77.2%，三者均居首位，综合检测性能显著领先，具有较强的目标识别与检测能力。

4.5. 消融实验

为了验证本文所提改进方案对提升 YOLO11 模型性能的有效性，系统性地设计并开展了消融实验。如表 4 所示。

Table 4. Ablation experiments

表 4. 消融实验

检测模型	<i>P</i> /%	<i>R</i> /%	mAP50/%
Yolo11	85.6	76.5	83.2
Yolo11-E	88.5	77.2	84.8
Yolo11-B	88.2	83.3	88.6
Yolo11-EB	90.9	88.4	90.5

消融实验结果表明，各改进模块对模型性能均具有显著提升作用。引入 EfficientNetV2 后，模型精确度明显提高，但召回率提升有限，说明其主要增强了特征判别能力。引入 BiFPN 后，召回率大幅提升，表明其在多尺度特征融合方面具有显著优势。

当两者结合构成 YOLOv11-EB 模型时，各项指标均达到最优，体现出良好的协同效应，使模型在精确度与召回率之间实现了更优平衡。

5. 结论

本文围绕芯片表面缺陷检测中存在的缺陷特征相似、小目标占比高及尺度差异显著等关键问题，提出了一种改进的 YOLOv11 模型(YOLOv11-EB)，并通过实验验证了其有效性与优越性。主要结论如下：

- 1) 针对传统模型在复杂场景下特征表达能力不足的问题，本文引入 EfficientNetV2 作为骨干网络，有效提升了模型对细粒度特征的提取能力，使其在相似缺陷区分方面表现更加稳定。
- 2) 通过构建加权双向特征金字塔网络(BiFPN)替代原有颈部结构，显著增强了多尺度特征融合能力，提高了模型对不同尺寸缺陷目标的检测性能，尤其在小目标检测方面效果明显改善。

3) 实验结果表明,改进模型在精确度、召回率及 mAP50 指标上分别达到 90.9%、88.4%和 90.5%,相较于原模型均有显著提升,验证了各改进策略的协同优化效果。

4) 所提出的方法在复杂背景条件下能够有效降低误检与漏检率,提升检测稳定性和鲁棒性,具备良好的工业应用潜力。

5) 本研究成果对于提升芯片制造过程中的质量检测效率具有重要意义,不仅有助于提高产品良率、降低生产成本,同时也为后续高精度缺陷检测模型的研究提供了参考思路。

参考文献

- [1] 杨璐霞, 崔硕, 张红瑞, 等. 基于 HRM-YOLO 的交通标志检测算法[J]. 现代电子技术, 2026, 49(7): 190-198.
- [2] Liao, J.G. and Tian, H.N. (2024) MBB-YOLO: A Comprehensively Improved Lightweight Aigorithm for Crowded Object Detection. *Concurrency and Computation: Practice and Experience*, **36**, e8219. <https://doi.org/10.1002/cpe.8219>
- [3] 王翔, 黄娟, 顾寄南, 等. 基于方向解耦与多尺度增强的芯片表面缺陷检测模型[J]. 半导体技术, 2026, 51(1): 68-76.
- [4] 刘琛, 袁小芳, 田争鸣, 等. 基于点模式匹配的芯片缺陷检测算法研究[J]. 激光杂志, 2020, 41(1): 39-44.
- [5] 陈文凤, 张淼, 欧幸福, 等. 基于嵌入式的 IC 芯片外观质量检测系统设计[J]. 电子制作, 2016(9): 43-44.
- [6] 李本红, 张淼, 欧幸福. 基于机器视觉的 SOP 芯片引脚缺陷检测系统设计[J]. 电子器件, 2017, 40(1): 171-178.
- [7] 李绎铃. 基于机器视觉的 SOP 芯片引脚缺陷检测研究[J]. 科技资讯, 2018, 16(10): 2-5.
- [8] 金贺楠. 基于机器视觉的 SOP 芯片引脚缺陷检测的研究与应用[D]: [硕士学位论文]. 沈阳: 东北大学, 2019.
- [9] Li, K., Xu, L., Su, L., Gu, J., Ji, Y., Wang, G., *et al.* (2024) X-Ray Detection of Ceramic Packaging Chip Solder Defects Based on Improved YOLOv5. *NDT & E International*, **143**, Article 103048. <https://doi.org/10.1016/j.ndteint.2024.103048>.
- [10] Liu, Y., Liu, Y., Guo, X., Ling, X. and Geng, Q. (2025) Metal Surface Defect Detection Using SLF-YOLO Enhanced YOLOv8 Model. *Scientific Reports*, **15**, Article No. 11105. <https://doi.org/10.1038/s41598-025-94936-9>
- [11] Zhang, T., Zhang, J., Pan, P. and Zhang, X. (2024) YOLO-RRL: A Lightweight Algorithm for PCB Surface Defect Detection. *Applied Sciences*, **14**, Article 7460. <https://doi.org/10.3390/app14177460>
- [12] 贺智轩, 陈里里, 王翔, 等. DMF-YOLOv11: 基于改进 YOLOv11n 的无人机航拍图像目标检测算法[J]. 计算机工程与应用, 2025, 61(14): 88-100.
- [13] 孙佩月, 黄娟, 顾寄南, 等. 基于 YOLO11n 的轻量化芯片表面缺陷检测方法[J]. 半导体技术, 2025, 50(7): 707-713.
- [14] Khanam, R. and Hussain, M. (2024) YOLOv11: An Overview of the Key Architectural Enhancements. <https://arxiv.org/abs/2410.17725v1>
- [15] 李慧琴, 宋赵铭, 肖亚涛, 等. 基于改进 YOLOv8n 的番茄果实检测模型[J]. 河南农业大学学报, 2026, 60(2): 347-355.