

基于改进型YOLOv5算法的偏振片缺陷识别研究

贾晓斌*, 罗柏文, 金祝红

湖南科技大学机电工程学院, 湖南 湘潭

收稿日期: 2025年3月4日; 录用日期: 2025年4月2日; 发布日期: 2025年4月10日

摘要

偏振片作为TFT-LCD的关键部件之一, 其表面缺陷严重影响液晶显示器的成像质量。为了实现偏振片缺陷智能化在线检测从而替代目前因人眼检测导致效率低的问题, 提出了一种改进型YOLOv5检测算法, 即在Backbone层增加了CBAM注意力机制; 在Prediction层增加了一个新的输出层; 将传统边框回归损失函数改为CIOU_Loss。通过缺陷样本测试实验表明, 改进型算法尽管增加了一个输出层导致参数增加, 且FPS略微降低, 但mAP却提升了4个百分点, 并且检测的最高置信度达到了0.93。故改进型YOLOv5算法增强了缺陷目标识别精度和准确度。

关键词

缺陷检测, YOLOv5算法, CBAM注意力机制, 四尺度输出层

Research on Polarizer Defect Recognition Based on Improved YOLOv5 Algorithm

Xiaobin Jia*, Bowen Luo, Zhuhong Jin

School of Mechanical Engineering, Hunan University of Science and Technology, Xiangtan Hunan

Received: Mar. 4th, 2025; accepted: Apr. 2nd, 2025; published: Apr. 10th, 2025

Abstract

Polarizer is one of the key components of TFT-LCD. Its surface defects seriously affect the imaging quality of LCD. In order to realize the intelligent on-line detection of polarizer defects and replace the current problem of low efficiency caused by human eye detection, an improved YOLOv5 detection algorithm is proposed, that is, CBAM attention mechanism is added in the backbone layer; A new output layer is added in the prediction layer; The traditional border regression loss function is changed to

*通讯作者。

CIOU_Loss. The defect sample test experiment shows that although the improved algorithm adds an output layer, resulting in an increase in parameters and a slight decrease in FPS, the map increases by 4 percentage points, and the highest confidence of detection reaches 0.93. Therefore, the improved YOLOv5 algorithm enhances the accuracy and accuracy of defect target recognition.

Keywords

Defect Detection, YOLOv5 Algorithm, CBAM Attention Mechanism, Four Scale Output Layer

Copyright © 2025 by author(s) and Hans Publishers Inc.

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

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



Open Access

1. 引言

近年来,以 TFT-LCD(薄膜晶体管液晶显示器件)为代表的液晶显示器在显微镜、液晶电视、手机和平板电脑等领域获得了十分广泛的应用[1]。偏振片是 TFT-LCD 的关键部件之一,直接影响到显示器的成像效果。质量正常的偏振片其表面必须保证干净整洁,不允许有明显的脏污、折痕、划痕、气泡、凹凸点等缺陷。但偏振片在生产、运输和存储等环节中不可避免地会产生这些缺陷,因此对偏振片进行缺陷检测是不可缺少的工序。

目前国内,包括大型厂家,对于偏振片的缺陷检测依然停留在人工检测层面,除了检测效率低,长时间对人的眼睛和身心都会带来不小的伤害。采用机器视觉技术,通过相机代替人眼捕捉偏振片缺陷形成图像,再交由计算机进行一系列图像处理,可达到自动、快速和准确检测偏振片缺陷的目的。然而目前国内、外针对偏振片缺陷检测采用机器视觉技术的研究比较少,其方法也主要停留在传统算法上。例如,曾小星[2]采用黑白间隔的条纹光做光源,提高缺陷的对比度,然后使用 RPCA(鲁棒性主成分分析),将缺陷从条纹光提取出来。赖文威[3]采用主动光扫描法,通过主动光四步扫描将缺陷捕捉,然后使用 Rangefilt 滤波原理实现了缺陷的提取。许少鹏[4]采用结构光饱和成像方法继续增大缺陷对比度,然后采用基于 RPCA 的图像检测方法将缺陷提取了出来。贺健[5]提出了一种差分法也成功提取出了缺陷。但是这些方法虽然提取出了缺陷,而并未很好地解决缺陷分类和缺陷位置的提取问题,不利于现场在线检测。而随着机器深度学习的发展,各类目标缺陷检测算法[6]日益增多,例如,赵睿[7]等人提出一种改进的 YOLOv5s 检测安全帽的算法,采用 DenseBlock 模块代替主干网络中的切片(Focus)结构,提高了网络提取特征的能力;王静[8]等人采用 YOLOv5 检测白细胞,在主干网络中添加坐标注意力机制 CA [9],改良了网络的性能;李春霖[10]等采用 YOLOv3-Tiny 网络来提取偏振片缺陷,使用 Dense Block 模块和 SPP-Net 模块来优化其特征提取网络,取得了良好的效果。

本文根据深圳某偏振片生产厂家提出的人眼可见的缺陷为研究目标,例如:气泡、凹凸点等圆状物缺陷半径不小于 0.1 mm;折痕、划痕等线状物缺陷长度不小于 2 mm,宽度不小于 0.05 mm。以 YOLOv5 算法为基础,但针对偏振片缺陷数据集的特点又对 YOLOv5 算法进行改进,使之更适合偏振片自动化检测;并通过在原来网络输出三种特征图的基础上再增加一层特征图的输出,增加注意力机制和修改损失函数三个方面对原有网络进行优化,使之模型的泛化能力更强,从而增强了对目标不同尺度大小下的检测能力[11]-[17]。

2. 改进型 YOLOv5 算法

本文是在 YOLOv5 网络的基础上进行改进,与经典的 YOLOv4 [18]算法相比, YOLOv5 总体结构也

由四个部分组成:输入端(Input),主干网络(Backbone),多尺度特征融合模块(Neck)以及输出层(Prediction)。YOLOv5 算法拥有 4 个版本的网络模型,包括:YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x,本文是以 YOLOv5l 为基础的研究模型。改进型算法主要体现在主干网络和特征融合模块。把原来的 YOLOv5l 的网络模型 Backbone 中的 Conv 全部改进为了 Conv_CBAM;这样可以提高目标识别的精度。损失函数采用了 CIOU_LOSS 函数,从中心距离、重复面积、两框的长宽比 3 个方面综合考虑 CIOU_LOSS 能够显著提升目标检测的回归精度。并且增加了输出层,是针对小型气泡检测效果不佳的。进而提升整个改进算法的精确度。下面分别具体介绍注意机制添加、损失函数的选择、输出层的增加。

2.1. 添加注意力机制

通常情况下目标检测任务都是目标占比小而背景占比大,导致许多检测算法只关注背景,不关注缺陷;导致检测缺陷的精确度不佳。引入了注意力机制(Attention Mechanism)帮助算法忽略不重要的信息[19],把注意力集中在重要的信息上。本文引入注意力机制旨在提高检测算法对缺陷的敏感性。目前注意力机制种类较多,常见的有通道注意力机制(SENet) [20]和卷积模块的注意力机制(CBAM) [21]两种。

SE (Squeeze-and-Excitation)是一个模块,将现有的网络嵌入 SE 模块的话,那么该网络就是 SENet。SENet 认为一张特征图上并不是所有的通道都同等重要,也就是说每个通道其实还应该有一个重要性权值才行,然后每个通道的重要性权值乘以每个通道原来的值,就是我们求的真正特征图(feature map)。SENet 的结构如图 1 所示。

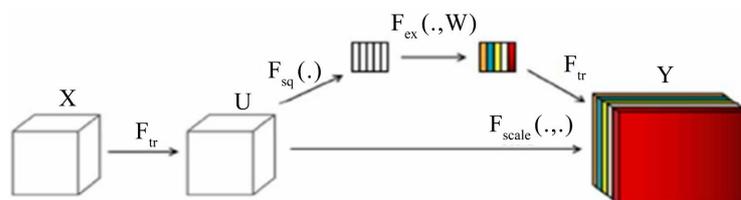


Figure 1. SENet structure diagram
图 1. SENet 结构图

SENet 解决了卷积池化过程中不同通道重要性不同带来的损失问题。进一步考虑到通道相同但位置不同的像素的重要性也可能不同,所以出现了 CBAM。CBAM 不仅考虑了不同通道像素的重要性,还考虑了同一通道不同位置的像素的重要性。CBAM 它比 SENet 多了个全局池化操作,这意味着它提取了更多的高级信息。鉴于此,本文采用 CBAM。

CBAM (图 2)是将通道注意机制(图 3)和空间注意机制(图 4)结合在一起的技术。对于一个特征图, CBAM 模块会从通道和空间两个维度推断注意力图,然后将注意力图与输入特征图相乘进行优化。本文结合原算法中的标准卷积模块 Conv 和 CBAM 模块设计 Conv_CBAM 模块,将主干网络中的 Conv 模块替换为 Conv_CBAM 模块。如图 5 所示。

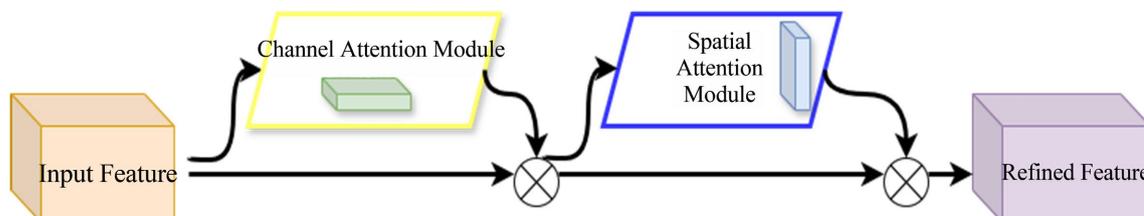


Figure 2. CBAM structure diagram
图 2. CBAM 结构图

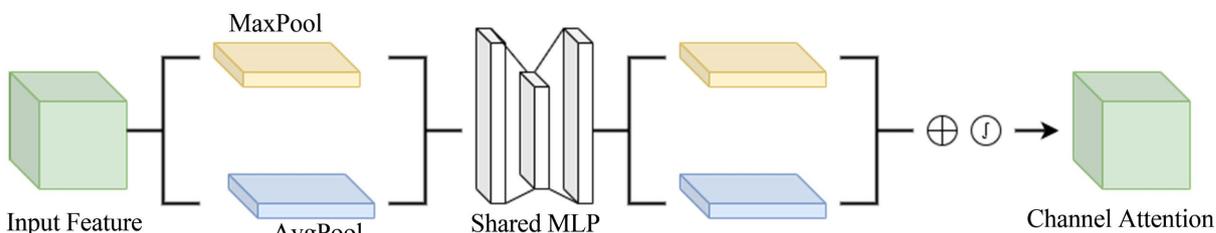


Figure 3. Channel attention module
图 3. 通道注意力机制

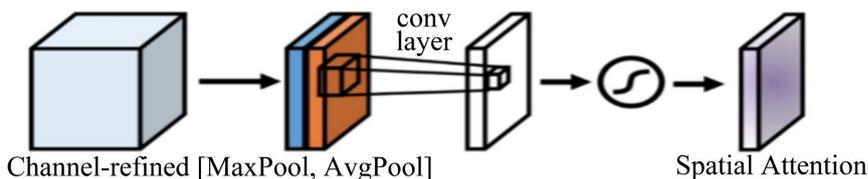


Figure 4. Spatial attention module
图 4. 空间注意力机制



Figure 5. Conv_CBAM module
图 5. Conv_CBAM 结构

2.2. 采用 CIOU_Loss 损失函数

YOLOv5 的损失函数包括三个方面：置信度损失函数(L_{obj})、分类损失函数(L_{cls})和边框回归损失函数(L_{box})。总的损失函数计算公式如(1)所示：

$$Loss = L_{obj} + L_{cls} + L_{box} \tag{1}$$

其中边框回归损失函数是一个不断演变的过程，最开始使用的是 IOU_Loss。IOU_Loss 是在 2016 年被提出来的，其计算公式如(2)所示：

$$IOU_Loss = 1 - IOU \tag{2}$$

IOU_Loss 很好地解决了预测框冗余的问题，但也存在局限：当预测框和真实框不相交，或者两框相交但交集大小相同时，这两种情况都会导致 IOU_Loss 相同，这样就无法做正确的回归。为此，2019 年又提出了 GIOU_Loss 来进行改进。

GIOU_Loss 的演示图如图 6 所示。其计算公式如(3)所示：

$$GIOU_Loss = 1 - GIOU = 1 - \left(IOU - \frac{L}{C} \right) \tag{3}$$

传统 YOLOv5 算法针对边框回归多采用 GIOU_Loss [22]损失函数，该函数虽然考虑了目标框和预测框的最小外接矩形，但忽略了预测框在目标框内部且不同预测框大小一致的问题如图 7 所示，这样导致 GIOU_Loss 退化成普通的 IOU_Loss 函数。

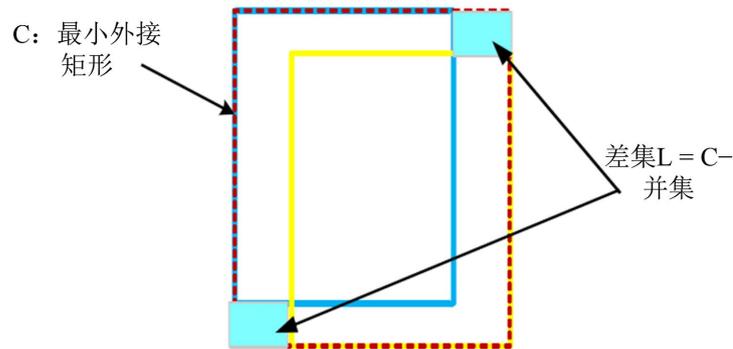


Figure 6. GIOU_Loss demo image
图 6. GIOU_Loss 演示图

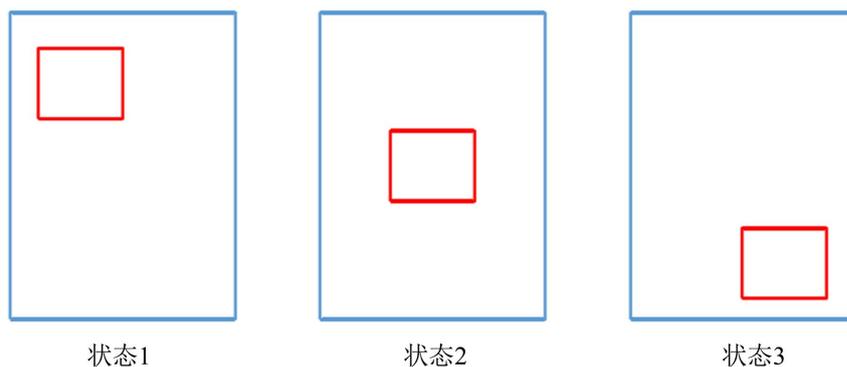


Figure 7. The prediction box is inside the target box
图 7. 预测框在目标框内部

为了解决这种问题相继出现了 DIOU_Loss 函数和 CIOU_Loss 函数。DIOU_Loss 函数考虑到了重复面积和两框的中心点距离，而 CIOU_Loss 函数既考虑了中心点距离，又考虑到了两框的长宽比，所以能够提高边框回归的精度和速度。本文采用 CIOU_Loss 函数作为边框回归的损失函数。CIOU_Loss 的计算公式见(4)：

$$CIOU_LOSS = 1 - CIOU = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \tag{4}$$

其中 b 是预测框中心点， b^{gt} 是目标框中心点， ρ 表示两框中心点距离， c 是两框的最小外接矩形的对角线长度， α 是权重， v 是表示长宽比的参数。公式如(5)、(6)所示：

$$\alpha = \frac{v}{(1 - IOU) + v} \tag{5}$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \tag{6}$$

2.3. 增加输出层

最初的网络有三个输出，当输入图片的尺寸为 608×608 像素时，经过主干网络的 Focus 和 4 个步长为 2 的卷积之后，特征图大小分别变为 304×304 ， 152×152 ， 76×76 ， 38×38 和 19×19 ，随后网络选取最后三种尺寸大小的特征图进行输出， 76×76 的感受野最小，用于检测小目标， 38×38 的感受野中

等，用于检测中等大小目标， 19×19 的感受野最大，用于检测大目标。经实验，以上特征检测面积大的目标效果较好，但针对偏振片上一些较小的缺陷，如小型气泡则检测效果不佳。为此，本文在原来的基础上再增加了一层 152×152 的输出，拓展为四个尺度。最后共得到四个尺度不同的输出结果。相比于原来的输出，新的输出层多了个更大的预测特征图，该特征图的感受野最大，能够更好地检测小目标，并且中间两种尺度的特征图分别融合了多个其它尺度的浅层信息：如第 12 层融合了 6 层的信息；第 16 层融合了 4 层的信息；第 20 层融合了 2 层的信息；第 23 层融合了 18 层的信息；第 26 层分别融合了 14 层的信息，第 29 层分别融合了 10 层的信息这种多融合的方式减小了过程中特征信息的丢失。改进后的最终网络结构如图所示图 8 所示：

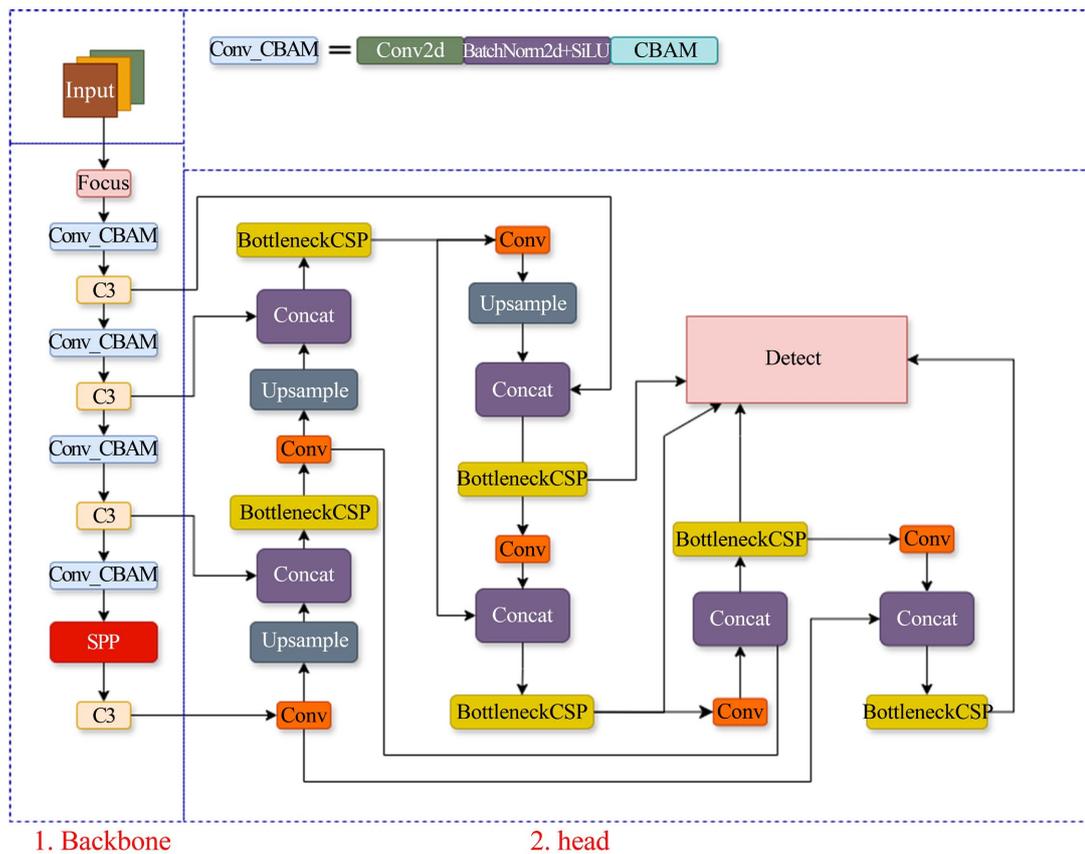


Figure 8. Improved network structure
图 8. 改进后的网络结构

3. 实验与分析

3.1. 实验环境

实验相关设施如表 1 所示，偏振片来源于深圳某厂家生产线上的不合格样品，其主要缺陷为划痕和气泡，样品大小为 $150 \text{ mm} \times 150 \text{ mm}$ 。为了扩大样本图数量通过人工做标记方式，并利用旋转，缩放和模糊等手段扩充到 1000 张，典型的样本图如图 9 所示。得到缺陷样本图后，再使用 labelimg 软件进行标注，采集到的缺陷样本中，划痕数量为 3097 个，气泡数量为 5589 个。多个缺陷，所以缺陷数量的总和远大于缺陷图的数量。由于一张缺陷图上可能存在每标注完成一张图后会在指定文件夹下生成一个对应的 xml 数据集文件，该文件里保存了标注的信息，如类别，边框的坐标等。标注界面如图 10 所示。

Table 1. Experimental configuration
表 1. 实验配置

名称	配置
CPU	INTER COREI7
GPU	NVIDIA GEFORCE GTX 1080Ti
操作系统	Ubuntu 16.0406 LTS
开发环境	Python3.7 Pycharm2019
深度学习框架	Pytorch1.6
工业相机	MV-GEA1201M-T
镜头	AFT-2514MP
偏振片样本	150 mm × 100 mm

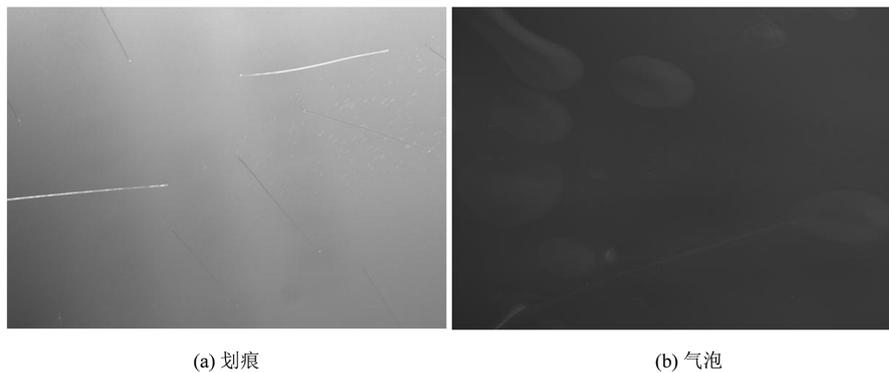


Figure 9. Defect sample chart
图 9. 缺陷样本图

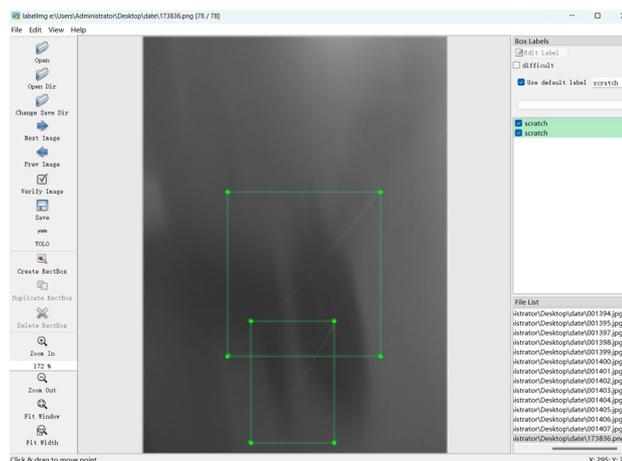


Figure 10. Defect labeling
图 10. 缺陷标注

3.2. 评价指标及结果分析

本文采用的算法评价指标有精确率 P (Precision)、召回率 R (Recall)、和平均精度均值 mAP (Mean Average Precision)和帧率 FPS 共四个指标衡量算法性能。指标计算公式如式(7)~(9)所示:

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

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

$$mAP = \frac{1}{c} \sum_{i=1}^c AP_i \tag{9}$$

将缺陷样本数据集按照 1:8 的比例分为测试集和训练集，使用预训练权重 YOLOv5l.pt 进行初始化，并将训练得到的权重文件用作测试的权重文件，图片的输入大小设为 416×416 ，迭代的 epochs 设为 400 次，batch_size 设为 4，学习率设为 0.01，本次实验训练了 2 种模型，分别是 YOLOv5 和本文的改进型 YOLOv5 算法。实验流程如图 11 所示。

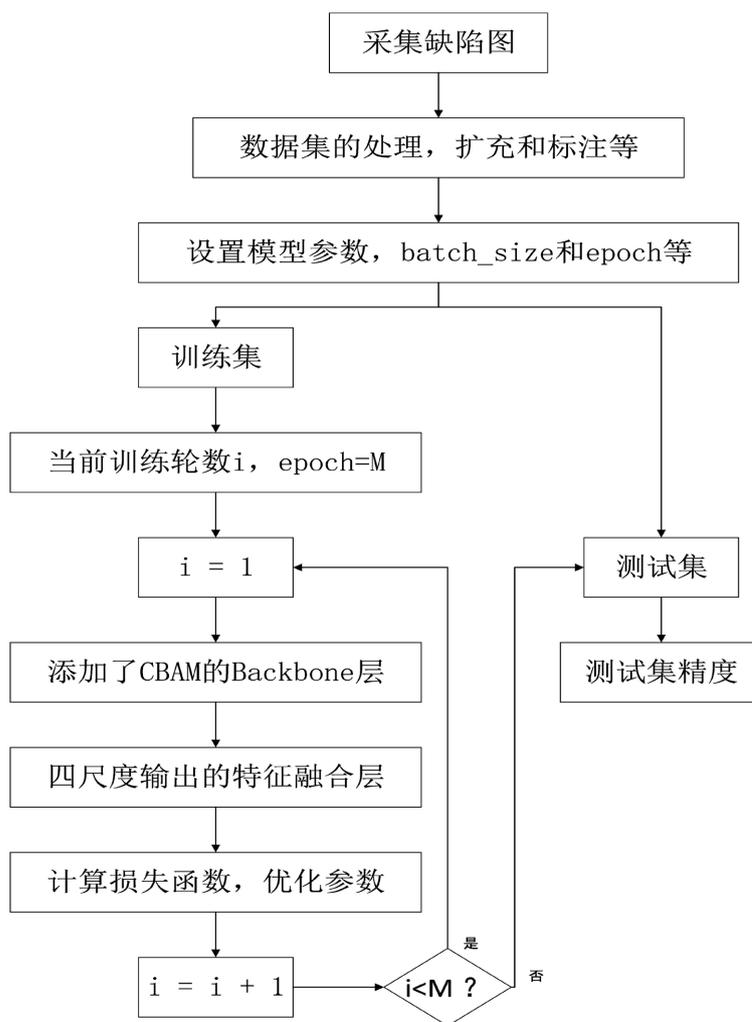


Figure 11. Experimental process
图 11. 实验流程

如表 2 所示，为了验证改进型算法的性能，将两种算法进行了两组对比实验。通过一步步在网络中添加注意力机制，更改损失函数和增加输出层之后，YOLOv5 的 mAP 值由最初的 0.8162 增长到了 0.8598；推理速度方面原算法为 52.6 FPS，改进型算法为 51.2 FPS，由此可见 mAP 提升了 4%，而推理速度却并

未下降太多,充分体现了改进型算法的优良性。

Table 2. Comparison of algorithm parameters

表 2. 算法参数对比

模型	P	R	mAP@0.5	FPS
YOLOv5	81%	90%	0.8162	52
YOLOv5-new	85%	92%	0.8598	51

最后得到的损失函数结果如图 12 所示。其中(1)、(2)、(3)分别为改进前算法 YOLOv5-ori 和改进后算法 YOLOv5-new 的定位损失、置信度损失和分类损失函数的对比图,横坐标是训练轮数,纵坐标是损失值。如图 12 所示,训练到 350 个 epochs 基本上开始收敛,而且改进后的损失函数曲线比改进前下降得更多,更加直观地表明了改进型算法比原算法更优良。

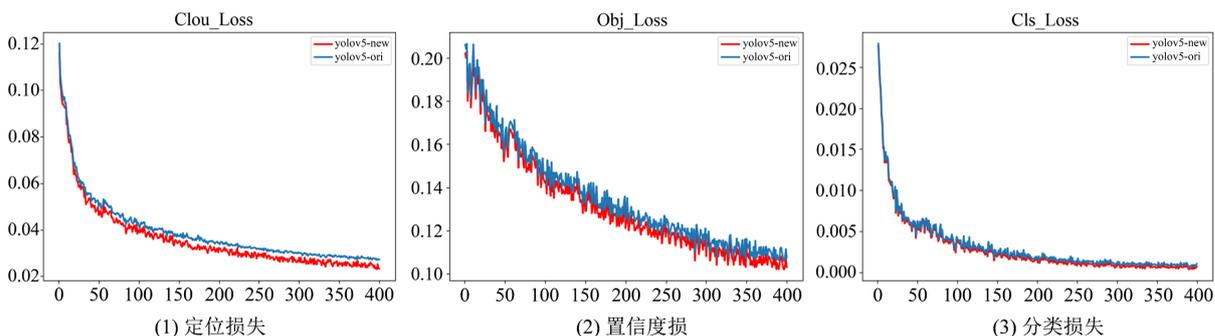


Figure 12. Loss function contrast curve

图 12. 损失函数对比曲线



Figure 13. Improved pre- and post-test results

图 13. 改进前后检测结果

现随机选取了几张缺陷样本图在两种算法下展示测试效果(放大图),如图 13 所示,左列为原算法结果,右列为改良算法结果。图中可见,改进型算法在检测精度上有不少提升,并且原算法检测不出来(较淡)的缺陷都能由改进型算法检测出来,并由此验证了算法改进的正确性。

由图可见,相对于改良之前,改良之后的算法在检测精度上有不少提升,并且原算法检测不出来的较淡的缺陷本文算法都能检测出来,这也验证了本文所作改进的正确性。

4. 结论

为了智能化在线检测偏振片缺陷,本文提出采用 YOLOv5 深度学习方法,并对 YOLOv5 算法进行改进:

- 1) 在 Backbone 层增加 CBAM 注意力机制,使得算法更加注重目标;
- 2) 在 Prediction 层增加了一个新的输出层,以此增强检测小目标的能力;
- 3) 修改了边框回归损失函数为 CIOU_Loss,以此提高边框回归的精度。

通过缺陷样本实验测试表明,在检测效果方面,改进型算法尽管增加了一个输出层导致参数增加,且 FPS 略微降低,但 mAP 却提升了 4 个百分点,并且检测的精准度达到了 85%。

参考文献

- [1] Sohn, S., Lee, D., Choi, H., Suh, J. and Bae, H. (2007) Detection of Various Defects in TFT-LCD Polarizing Film. In: Beliczynski, B., Dzielinski, A., Iwanowski, M. and Ribeiro, B., Eds., *Adaptive and Natural Computing Algorithms. ICANNGA 2007. Lecture Notes in Computer Science*, Springer, 534-543. https://doi.org/10.1007/978-3-540-71629-7_60
- [2] 曾小星. 基于结构光的偏光片外观缺陷检测技术研究[D]: [硕士学位论文]. 深圳: 深圳大学, 2016.
- [3] 赖文威. 偏光片外观缺陷成像机理与检测技术研究[D]: [硕士学位论文]. 深圳: 深圳大学, 2017.
- [4] 许少鹏. 基于机器视觉的偏光片缺陷检测技术研究[D]: [硕士学位论文]. 深圳: 深圳大学, 2018.
- [5] 贺健. 偏光片外观缺陷成像仿真与检测[D]: [硕士学位论文]. 深圳: 深圳大学, 2016.
- [6] 李柯泉, 陈燕, 刘佳晨, 牟向伟. 基于深度学习的目标检测算法综述[J]. 计算机工程, 2022, 48(7): 1-17.
- [7] 赵睿, 刘辉, 刘沛霖, 雷音, 李达. 基于改进 YOLOv5s 的安全帽检测算法[J]. 北京航空航天大学学报, 2023, 49(8): 2050-2061.
- [8] 王静, 孙紫雲, 郭苹, 张龙妹. 改进 YOLOv5 的白细胞检测算法[J]. 计算机工程与应用, 2022, 58(4): 134-142.
- [9] Hou, Q., Zhou, D. and Feng, J. (2021) Coordinate Attention for Efficient Mobile Network Design. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 13708-13717. <https://doi.org/10.1109/cvpr46437.2021.01350>
- [10] 李春霖, 谢刚, 王银, 谢新林, 刘瑞珍. 基于 YOLOv3-Tiny-D 算法的偏光片缺陷检测[J]. 计算机集成制造系统, 2022, 28(3): 787-797.
- [11] Pan, X., Ge, C., Lu, R., Song, S., Chen, G., Huang, Z., et al. (2022) On the Integration of Self-Attention and Convolution. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 805-815. <https://doi.org/10.1109/cvpr52688.2022.00089>
- [12] Chen, C., Liu, M., Tuzel, O. and Xiao, J. (2017) R-CNN for Small Object Detection. In: Lai, S.H., Lepetit, V., Nishino, K. and Sato, Y., Eds., *Computer Vision—ACCV 2016. Lecture Notes in Computer Science*, Springer International Publishing, 214-230. https://doi.org/10.1007/978-3-319-54193-8_14
- [13] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90. <https://doi.org/10.1145/3065386>
- [14] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., et al. (1989) Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, **1**, 541-551. <https://doi.org/10.1162/neco.1989.1.4.541>
- [15] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I. (2017) Attention Is All You Need. In: Guyon, I., Von Luxburg, U., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S. and Garnett, R., Eds., *Advances in Neural Information Processing Systems*. <https://doi.org/10.48550/arXiv.1706.03762>

-
- [16] Niu, Z., Zhong, G. and Yu, H. (2021) A Review on the Attention Mechanism of Deep Learning. *Neurocomputing*, **452**, 48-62. <https://doi.org/10.1016/j.neucom.2021.03.091>
- [17] Cordonnier, J.B., Loukas, A. and Jaggi, M. (2019) On the Relationship between Self-Attention and Convolutional Layers.
- [18] Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M. (2020) YOLOv4: Optimal Speed and Accuracy of Object Detection.
- [19] Wang, C., Mark Liao, H., Wu, Y., Chen, P., Hsieh, J. and Yeh, I. (2020) CSPNet: A New Backbone That Can Enhance Learning Capability of CNN. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, 14-19 June 2020, 1571-1580. <https://doi.org/10.1109/cvprw50498.2020.00203>
- [20] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/cvpr.2018.00745>
- [21] Woo, S., Park, J., Lee, J. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018. Lecture Notes in Computer Science*, Springer International Publishing, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1
- [22] Zheng, Z., Wang, P., Liu, W., *et al.* (2019) Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. *Proceedings of the AAAI Conference on Artificial Intelligence*, **34**, 12993-13000.