

基于改进 YOLOv5 的新材料地板表面缺陷检测研究

张忠^a, 魏国亮^b, 张之江^a, 蔡贤杰^a, 王耀磊^c

(上海理工大学 a.光电信息与计算机工程学院 b.管理学院 c.理学院, 上海 200093)

摘要: **目的** 提升质检过程中新材料地板的表面缺陷检测精度。**方法** 通过翻转、水平迁移和垂直迁移对采集到的缺陷图像进行扩充, 构建新材料地板缺陷数据集。基于 YOLOv5 算法, 增加一个预测头, 使算法对微小缺陷更加敏感; 其次在网络的特征融合层应用 Swin Transformer 模块, 形成注意力机制预测头, 提高网络特征提取效率; 然后在网络主干末端加入 SE 模块, 使网络提取有用的信息特征, 提高模型精度。**结果** 实验结果表明, 提出的新材料地板表面缺陷检测方法能够准确判别地板好坏, 并能够识别出白色杂质、黑斑、边损、气泡胶等 4 类表面缺陷, 各缺陷类型的平均精确均值为 82.30%, 比 YOLOv5 Baseline 提高了 6.58%, 相比其他典型目标检测算法也能够更准确和快速地识别地板表面缺陷。**结论** 通过改进的 YOLOv5 算法可以更准确地对地板表面缺陷进行分类与定位, 从而大大提高工业质检效率。

关键词: 新材料地板; 缺陷检测; YOLOv5; 预测头; 注意力机制

中图分类号: TP273⁺.2 文献标识码: A 文章编号: 1001-3563(2023)07-0196-08

DOI: 10.19554/j.cnki.1001-3563.2023.07.022

Surface Defect Detection of New Material Floor Based on Improved YOLOv5

ZHANG Zhong^a, WEI Guo-liang^b, ZHANG Zhi-jiang^a, CAI Xian-jie^a, WANG Yao-lei^c

(a. School of Optical Electrical and Computer Engineering b. Business School c. College of Science,
University of Shanghai for Science and Technology, Shanghai 200093, China)

ABSTRACT: The work aims to improve the surface defect detection accuracy of new material floor during quality inspection. The defect images collected were expanded by flipping, horizontal migration and vertical migration, and the defect data set of new material floor was constructed. Based on YOLOv5, a prediction head was added to make the algorithm more sensitive to tiny defects. Next, the Swin Transformer module was applied in the feature fusion layer of the network to form the attentional mechanism prediction head and improve the efficiency of network feature extraction. Then, the SE module was added at the end of the backbone to enable the network to extract useful feature information and improve the model accuracy. The experimental results showed that the proposed method could accurately distinguish the quality of floor and identify four kinds of surface defects including white impurity, black spot, edge damage and bubble gum. The mean average precision for each defect type was 82.30%, which was 6.58% higher than that of YOLOv5 Baseline. Compared with other typical target detection algorithms, it could identify floor surface defects more accurately and quickly. The improved YOLOv5 algorithm can classify and locate the surface defects of the floor more accurately, thus greatly improving the efficiency of industrial quality inspection.

KEY WORDS: new material floor; defect detection; YOLOv5; prediction head; attentional mechanism

收稿日期: 2022-05-01

基金项目: 上海市“科技创新行动计划”国内科技合作项目 (20015801100)

作者简介: 张忠 (1997—), 男, 硕士生, 主要研究方向为机器视觉。

通信作者: 魏国亮 (1973—), 男, 博士, 教授, 主要研究方向为非线性系统、多智能体协同控制。

2020 年 5 月 7 日,《全球森林资源评估》报告中指出,当前森林面积为 40.6 亿公顷,相当于人均 0.52 公顷。自 1990 年以来,全球因毁林而损失了约 4.2 亿公顷森林,虽然近几年森林流失的速度已大幅下降,但是全球森林砍伐仍在继续。由于我国人口基数大、资源的人均占有率低^[1],木材工业面临着资源供应不足的风险,这严重限制了实木地板的发展,与此同时,为了保护环境,急需一种新型环保材料来代替实木地板,于是新材料地板产品^[2]开始盛行起来。

随着工业自动化技术的快速发展,流水线生产效率的不断提高,单纯依靠人工来判别地板质量优劣的方式,很容易出现漏检、误检,而且效率极低,已经不再满足日益高效的工业生产要求。目前,已经有研究者提出一些传统图像处理的检测算法:张怡卓等^[3]提出了基于 R 分量的形态学分割方法来确定缺陷目标,进行特征提取后,用 SOM 网络对缺陷进行分类;许超^[4]提出了基于特征融合与压缩感知的实木地板的缺陷检测方法;郭慧等^[5]提出了人造板表面缺陷检测图像的自适应快速阈值分割算法。这些检测方法虽然能够实现地板缺陷的自动检测,但是对一些小的缺陷以及在具有复杂背景属性的检测场景下的检测存在明显不足。

近些年来,深度学习在缺陷检测方面的逐渐应用广泛,它在缺陷特征提取和目标定位上取得了非常好的效果。基于 CNN (Convolution Neural Network) 的目标检测模型分为多种类型:单阶段检测器,如 YOLOX^[6]、FCOS^[7]、DETR^[8]、ScaledYOLOv4^[9]、EfficientDet^[10]等;两阶段检测器,如 VFNet^[11]、CPNNet^[12]等;基于锚框的检测器,如 ScaledYOLOv4^[9]、YOLOv5^[13]等;无锚检测器,如 CornerNet^[14]、YOLOX^[6]、RepPoints^[15]等。

目前,基于深度学习的缺陷检测已有如下工作:Wei 等^[16]以 SSD 为基础网络,结合轻量级神经网络 MobileNet,提高了网络推理速度,并引入 Inception 网络,构建了 SSD-MobileNet 算法模型提高了识别人造板不同缺陷的能力;Zhao 等^[17]通过重构 Faster Rcn 的网络结构,设计可变形卷积网络替换常规卷积网络,提高了神经网络对不规则缺陷的识别能力,并采用多尺度融合训练网络来检测钢材表面的小面积缺陷;Anwar 等^[18]提出了一种基于 YOLOv4 改进的深度学习算法,并使用空间注意力机制 SAM (Spatial Attention Mechanisms) 来提高所提出方法的性能,从而更准确地对轨面缺陷进行检测与分类。

相较于其他经典目标检测算法,YOLOv5 算法表现优异,检测目标准确率与速度都较高^[13]。YOLOv5 算法是由 Ultralytics 在 2020 年提出的,其性能与 YOLOv4 不相伯仲,是目前最先进的目标检测算法之一,具有较快的推理速度。该网络模型主干部分采用 CSPDarknet53^[19]的架构,以 PANet^[20]为

颈部,并采用 3 个 YOLO 检测头,但是原 YOLOv5 算法对新材料地板的微小缺陷检测效果较差,而且存在检测精度不足,因此,本文提出一种基于改进的 YOLOv5 网络模型的新材料地板表面缺陷分类检测方法。

1 改进的 YOLOv5 新材料地板检测方法

改进的 YOLOv5 的框架如图 1 所示。主干网络 (Backbone): 在 CSPDarknet53^[19]主干网络末端加入 SE 模块。颈部 (Neck): 使用类似 PANet^[20]的结构。头部 (Head): 探索 Swin Transformer 的预测潜力,使用该注意力机制预测头^[21]代替原来的预测头。通过修改原始的 YOLOv5 模型,使其更好地应用于新材料地板数据集。

1.1 预测微小缺陷

通过观察新材料地板数据集,发现有的缺陷非常小,因此增加一个预测头进行微小缺陷的检测。根据国际组织 SPIE 的定义,对于相对尺寸角度,小于 256×256 的 0.12% 就为小目标,在文中,小于 640×640 的 0.12% 即为微小缺陷。结合其他 3 个预测头 (head2、head3 和 head4),可以检测不同尺度的缺陷。如图 1 所示,所添加的预测头 (head1) 是由浅层和高分辨率的特征图生成的。在第 18 层后,为了使得特征图进一步扩大,采用上采样 (Upsample) 等方式对特征图进行处理,同时,在第 22 层时,通过 Concat 操作将骨干网络 (Backbone) 中第 2 层特征图与第 20 层得到的大小为 160×160 的特征图进行通道维度上的拼接,以此获取更大、包含更多信息的特征图,使得它对微小缺陷更加敏感。增加检测头后,虽然增加了计算量和内存开销,但对微小缺陷的检测性能有所提高。

1.2 Swin Transformer 模块

基于 YOLOv5,对图 1 网络结构中 C3 模块进行改进,将其中的 Bottleneck 模块替换为 Swin Transformer^[22]模块,构成了图 1 中的 STR 模块。STR 模块处在网络颈部 (Neck) 的输出预测层 (图 1 中第 22、25、28、33 层),形成了注意力机制预测头^[18]。

Swin Transformer^[22]模块是基于原 Transformer Layer^[23]的标准多头自注意力机制提出的,主要区别在于加入了滑动窗口和局部注意的机制。如图 2 所示,给定一个尺寸为 $H \times W \times C$ 的图片输入,Swin Transformer 首先将输入分割成不重叠的 $M \times M$ 的局部窗口,再将其重塑为一个 $(HW/M^2) \times M^2 \times C$ 的特征图,其中 HW/M^2 为窗口总数。然后,分别计算每个窗口的标准自我注意力 (局部注意)。对于局部窗口特征

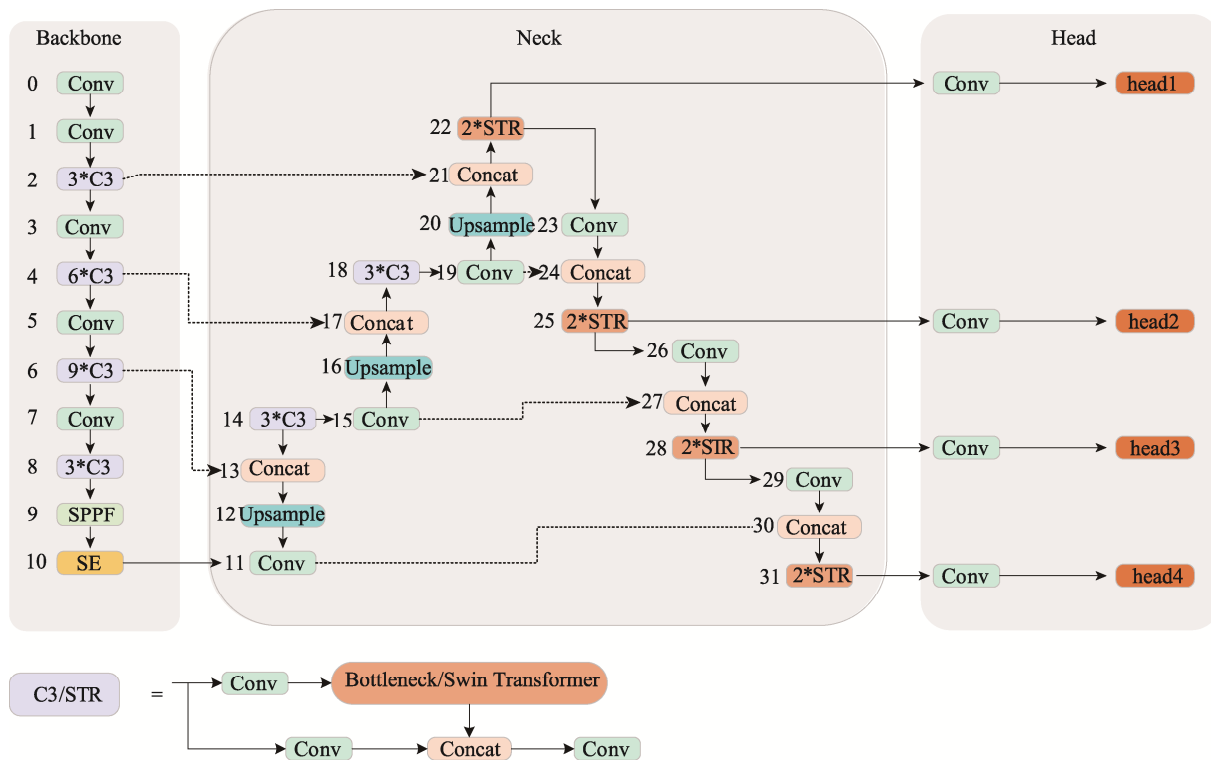


图1 改进 YOLOv5 的模型结构
Fig.1 Model structure based on improved YOLOv5

$X \in \mathbb{R}^{M^2 \times C}$, 查询矩阵 Q 、键矩阵 K 和值矩阵 V 计算见式 (1)。

$$Q = XP_Q, K = XP_K, V = XP_V \quad (1)$$

式中: P_Q 、 P_K 和 P_V 为在不同窗口中共享的投影矩阵。一般来说, $Q, K, V \in \mathbb{R}^{M^2 \times d}$, 因此, 注意力矩阵由局部窗口中的自我注意力机制计算:

$$\text{Attention}(Q, K, V) = \text{SoftMax}(QK^T / \sqrt{d} + B)V \quad (2)$$

式中: B 为可学习的相对位置编码。在实践中, 并行执行了 h 次的注意力机制函数, 并将结果连接起来用于多头自我注意力 (MSA)。

接下来, 使用多层感知器 (MLP) 进行特征变换, 多层感知器 (MLP) 具有 2 个全连接层, 全连接层之间具有 GELU 非线性。在 MSA 和 MLP 之前添加 LayerNorm (LN) 层, 并且 2 个模块都采用残差连接。整个过程表述见式 (3) — (4)。

$$X = \text{MSA}(\text{LN}(X)) + X \quad (3)$$

$$X = \text{MLP}(\text{LN}(X)) + X \quad (4)$$

当不同层的分区被固定时, 就不存在跨局部窗口的连接, 因此, 为了可以跨窗口连接^[22], 常规和滑动窗口分区被交替使用, 其中滑动窗口分区意味着在分区前将特征偏移 $(\lfloor M/2 \rfloor, \lfloor M/2 \rfloor)$ 像素。通过采用滑动窗口的机制实现跨窗口连接, 使模型可以关注相邻窗口的其他特征信息, 进行跨窗口的特征交互, 从而在一定程度上扩大了感受野, 提高了网络特征提取效率。

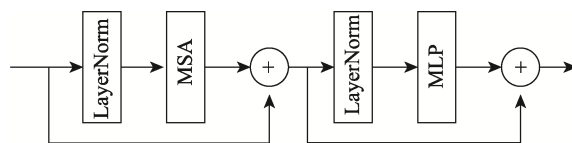


图2 Swin Transformer 模块组成
Fig.2 Swin Transformer module composition

1.3 SE 模块

SE 模块是一种计算单元, 通过引入压缩和激励^[24]显式地建立特征通道之间的相互依赖关系来改进网络。如图 3 所示, 将 F_{tr} 看作一个标准的卷积运算符, 通过一系列卷积操作, F_{tr} 将输入转换成输出 $U = [u_1, u_2, \dots, u_c]$ 。 u_c 的表达式如式 (5) 所示。

$$u_c = v_c * X = \sum_{s=1}^C v_c^s * x^s \quad (5)$$

式中: $*$ 为卷积; $v_c = [v_c^1, v_c^2, \dots, v_c^C]$; $X = [x_1, x_2, \dots, x^C]$; v_c^s 为一个二维卷积核, 表示作用于对应通道 X 上的单个 v_c 通道。

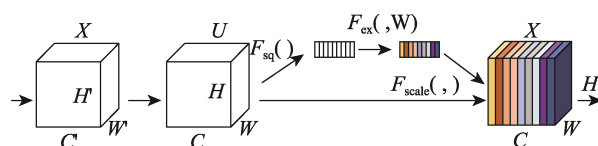


图3 SE 模块结构
Fig.3 Squeeze-and-Excitation module structure

SE 模块中的压缩操作是将上面获取到的 U 通过全局平均池化进行压缩, 变成一维的实数 $z \in \mathbb{R}^C$, 其中 z 的第 c 个元素的表达式如式 (6) 所示。

$$z_c = F_{sq}(u_c) = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H u_c(i, j) \quad (6)$$

通过激励操作来全面捕获通道依赖性, 学习通道之间的非线性关系。利用 Sigmoid 形式的门控机制, 得到信息 s 的表达式见式 (7)。

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (7)$$

为了防止模型复杂化, 在这里采用的参数分别为 W_1 和 W_2 的全连接层, 同时二者之间使用激活函数 δ (ReLU) 来引入非线性, 最后通过归一化函数 σ (Sigmoid) 来输出权重向量。SE 模块的最终输出 $\tilde{X} = [\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_c]$, 是通过激励部分输出与重新调整变换后的输出 U 相乘得到的。 $\tilde{x}_c$ 的表达式如式 (8) 所示。

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c \cdot u_c \quad (8)$$

为了减少网络参数增加, 只在主干末端 (第 10 层) 应用 SE 模块, 将前面卷积层的融合特征作为 SE 模块的输入, 通过 SE 模块上分支结构的全连接层与 ReLU、Sigmoid 激活函数, 实现利用 ReLU 激活函数增加网络的线性复杂度, 同时也能利用 Sigmoid 激活函数将输出缩小到 0~1, 从而使得有效信息特征的关注度得到增强, 无关信息特征的影响力被削弱^[25]。

2 材料与方法

2.1 数据样本结构

本文的新材料地板缺陷图片由浙江晶通新材料有限公司所提供, 这些图片均由黑白相机所拍摄, 共 1 700 张。本文实验选取 4 种常见的缺陷 (白色杂质、黑斑、边损、气泡胶), 如图 4 所示。

将采集的缺陷图片进行数据扩充, 主要的数据扩充操作为翻转、水平迁移、垂直迁移这 3 种方式, 如表 1 所示。将扩充得到的 1 860 张图片的 80% 作为训练集、10% 作为验证集、10% 作为测试集, 然后利用标注软件 Colabeler 对新材料地板表面缺陷进行标注。

2.2 实验平台

本文所有实验均在开源操作系统 Ubuntu 上进行的, 版本号为 20.04。该工作站的 CPU 型号为 Intel® Xeon (R) Gold 5220R, 具有 24 个内核和 48 个线程, 基本时钟速度可以达到 2.20 GHz; GPU 型号为 GeForce RTX3060, 它具有 64 GB 的 DDR4 内存, 专为 AI 和机器学习任务所设计。软硬件平台编程语言主要是 Python, 选用 Pytorch 作为搭建 YOLOv5 算法的基础框架, 并配置版本型号为 11.3 的 CUDA 对 GPU 运算进行加速^[16]。网络通过多次训练表明, 当动量 (Momentum) 参数赋值为 0.937、迭代批量 (Batch Size) 参数赋值为 8、权重衰减 (Weight Decay) 参数

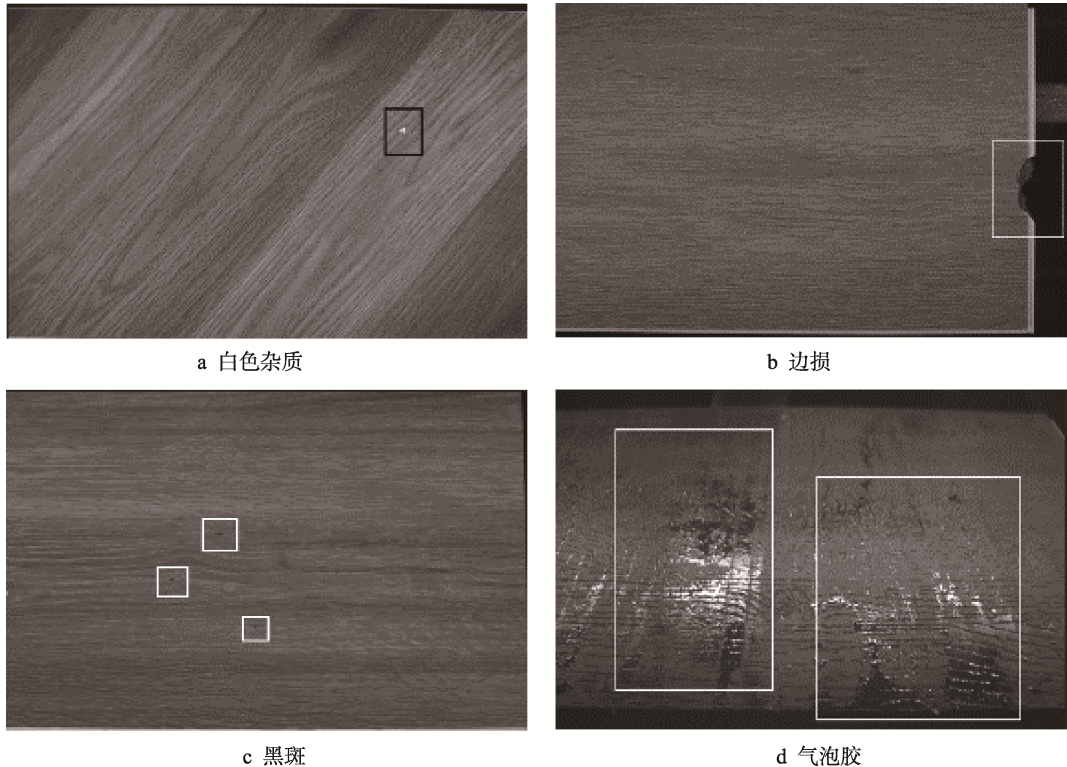


图 4 部分采集图像
Fig.4 Partially acquired image

表 1 数据扩充方法
Tab.1 Data augment method

方法	改变
翻转	将缺陷图像从 0° 旋转到 180°
水平迁移	横向偏移图像, 偏移值为图像长度的 10%
垂直迁移	垂直偏移图像, 偏移值为图像长度的 10%

赋值为 0.000 5、初始学习率 (Learning Rate) 参数赋值为 0.01、图片训练尺寸设置为 640×640 时, 训练速度最快。

3 实验结果与分析

3.1 算法性能分析

3.1.1 综合性能评估

评价算法的性能指标主要有各缺陷类别的平均精度 (Average Precision, δ_{AP}), 均值平均精度 (Mean Average Precision, δ_{mAP}), 检测每张图片平均所用时间 t_r , 准确率 P , 召回率 R 。具体计算式见式 (9) — (12)。

$$P = \frac{T_p}{T_p + F_p} \quad (9)$$

$$R = \frac{T_p}{T_p + F_N} \quad (10)$$

$$\delta_{AP} = \int_0^1 P(r) dr \quad (11)$$

$$\delta_{mAP} = \frac{1}{C} \sum_{k=1}^C \delta_{AP}(k) \quad (12)$$

式中: T_p 表示预测为正的正样本; F_p 表示预测为正的负样本; F_N 表示预测为负的正样本; δ_{mAP} 表示所有类别 δ_{AP} 的平均值, 即表示该模型在所有类别上的检测性能^[26]。

将测试集中的 186 张图片放入改进后的网络模型中, 具体检测结果如表 2 所示。从表 2 可以看出, 与 YOLOv5 基础模型相比, 虽然平均检测时间 t_r 有所增加, 但是改进后的 YOLOv5 模型的平均准确率均值提升了 6.58%, 说明本文提出的算法具有更好的检测性能。

表 2 改进网络模型检测结果对比
Tab.2 Comparison of detection results by improved network model

模型	$\delta_{AP}/\%$				$\delta_{mAP}/\%$	t_r/s
	白色杂质	黑斑	边损	气泡胶		
YOLOv5	78.11	84.23	58.33	82.20	75.72	0.020
Improved YOLOv5	78.20	90.28	67.49	93.24	82.30	0.033

3.1.2 数据集样本检测结果分析

本次实验任选其中 2 种缺陷 (以白色杂质和黑斑为例), 利用改进前后的 YOLOv5 网络模型对新材料地板数据集中的缺陷进行检测, 检测结果如图 5 所示。图 5a 和 b 为 YOLOv5 基础网络模型的检测结果, 图 5c 和 d 为改进 YOLOv5 网络模型 (本文) 的检测结果。将图 5a 和图 5c 进行对比, 可以看出改进后的置信度更高, 检测效果更好; 将图 5b 和 5d 进行对比, 可以看出虽然改进后的交并比有所下降, 但是改进后的算法可以检测到原网络检测不到的微小缺陷。

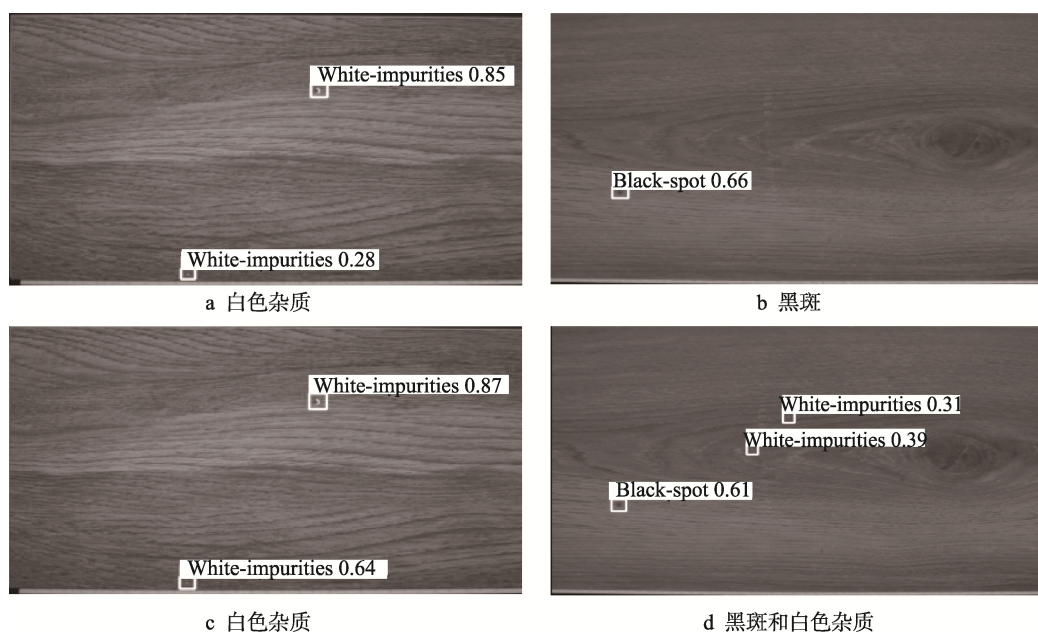


图 5 改进前后检测结果对比

Fig.5 Comparison of detection results before and after improvement

3.1.3 不同网络模型的性能对比

为了更全面地评测改进后 YOLOv5 的缺陷检测性能, 使用 SSD^[27]、Faster Rcn^[28]、YOLOX^[6]这 3 个网络模型进行对比实验, 并以 δ_{mAP} 和 t_r 作为评价算法检测性能的指标, 实验结果如表 3 所示。

从表 3 可以看出, 在 SSD、Faster Rcn、YOLOX 和 YOLOv5 这 4 种网络模型中, 均值平均精度最高的是 Faster Rcn, 检测时间最短的是 YOLOv5。与 Faster Rcn 相比, 本文方法的均值平均精度提高了 3.97%, 检测时间也缩短了约 44.1%; 与 YOLOv5 基础模型相比, 检测时间虽然有所增加, 但是均值平均精度大幅提高, 约为 6.58%。由此可见, 改进后的 YOLOv5 算法降低一点检测速度, 但大幅提高了检测缺陷的准确性, 能够更好地完成地板表面缺陷检测任务。

表 3 5 种网络模型检测效果对比
Tab.3 Comparison of detection effects of five kinds of network models

网络模型	$\delta_{mAP}/\%$	t_r/s
SSD	67.80	0.038
YOLOX	66.49	0.049
Faster Rcn	78.33	0.059
YOLOv5	75.72	0.020
Improved YOLOv5	82.30	0.033

3.2 不同比例模型的效果对比

S、M、L 和 X 是 YOLOv5 提供的 4 个不同比例的模型, 分别代表着小型、中型、大型、超大型。每种比例都对模型的深度和宽度乘以不同的系数, 这就意味着模型整体结构是保持不变的, 每个模型的大小和复杂度是按比例的。网络模型越复杂, 并不代表着检测准确性越好, 但网络的推理时间更长。经过多次实验, 发现使用 L 这种比例的模型效果最好, 如表 4 所示。

表 4 改进 YOLOv5 不同比例模型的效果对比
Tab.4 Comparison of effects of improved YOLOv5 models in different scales

比例	$\delta_{mAP}/\%$
S	79.96
M	81.25
L(本文)	82.30
X	80.00

3.3 消融实验

在新材料地板缺陷数据集上, 融入了不同组件进

行检测效果对比。通过组合不同组件, 在该数据集上进行检测效果对比的结果见表 5。

1) 额外预测头的效果。增加了一个针对小目标的预测头, 虽然使网络层数从 367 层增加到了 434 层, 但是检测指标的均值平均精度比于原网络的大幅提高了约 2.63%。

2) Swin Transformer 模块的效果。在使用 Swin Transformer 模块之后, 均值平均精度大幅提高了约 2.76%。

3) SE 模块的效果。加入 SE 模块使得网络提取有用的特征信息, 抑制无关的背景干扰, 从而提高网络模型的精度, 均值平均精度也有相应的提升。

表 5 融入不同组件检测效果对比
Tab.5 Comparison of detection effect after integration of different components

方法	$\delta_{mAP}/\%$
YOLOv5	75.72
YOLOv5+head	78.35(↑2.63)
YOLOv5+SE	76.76(↑1.04)
YOLOv5+STR	78.48(↑2.76)
YOLOv5+head+STR	81.00(↑5.28)
YOLOv5+head+SE	80.75(↑5.03)
Improved YOLOv5(本文)	82.30(↑6.58)

4 结语

为了能够更加准确地识别地板表面缺陷, 本文提出一种改进 YOLOv5 网络模型的方法。该改进主要是在 YOLOv5 主干网络的末端加入注意力机制, 使得网络关注有用的信息特征, 并增加一个小目标预测头, 在输出预测层融入注意力机制, 提高模型对微小缺陷的检测能力以及对信息特征的提取效率。实验结果表明, 相较于原 YOLOv5 网络模型, 改进后的 YOLOv5 算法的检测准确率得到了明显的提升, 而且能够检测到微小缺陷, 可以更好地实现新材料地板不同类别缺陷的分类检测。通过与其他 3 种网络模型对比可知, 改进后的 YOLOv5 算法检测精度更高, 但是在检测速度方面存在不足, 如何在保证检测性能的情况下提升检测速度是未来需要攻克的难题。

参考文献:

[1] 谢政峰. 基于深度学习的新材料地板缺陷检测技术研究[D]. 成都: 四川大学, 2021: 13-14.
XIE Zheng-feng. Research on Defect Detection Tech-

- nology of New Material Flooring Based on Deep Learning[D]. Sichuan: Sichuan University, 2021: 13-14.
- [2] 刘华江, 朱小斌. 设计师的材料清单[M]. 上海: 同济大学出版社. 2017: 25-26.
LIU Hua-jiang, ZHU Xiao-bin. Designer's Bill of Materials[M]. Shanghai: Tongji University Press. 2017: 25-26.
- [3] 张怡卓, 曹军, 许雷, 等. 实木地板缺陷形态学分割与 SOM 识别[J]. 电机与控制学报, 2013, 17(4): 116-120.
ZHANG Yi-zhuo, CAO Jun, XU Lei, et al. Wood Floor Defects Segmentation and Recognition Based on Morphological and SOM[J]. Electric Machines and Control, 2013, 17(4): 116-120.
- [4] 许超. 基于特征融合与压缩感知的实木地板缺陷检测方法研究[D]. 黑龙江: 东北林业大学, 2015: 5-10.
XU Chao. Research on Feature Fusion and Solid Wood Floor Defects Detection Method Based on Compressed Sensing[D]. Heilongjiang: Northeast Forestry University, 2015: 5-10.
- [5] 郭慧, 王霄, 刘传泽, 等. 人造板表面缺陷检测图像自适应快速阈值分割算法[J]. 林业科学, 2018, 54(11): 134-142.
GUO Hui, WANG Xiao, LIU Chuan-ze, et al. Research on Adaptive Fast Threshold Segmentation Algorithm for Surface Defect Detection of Wood-Based Panel[J]. Scientia Silvae Sinicae, 2018, 54(11): 134-142.
- [6] LIU B, HUANG J, LIN S, et al. Improved YOLOX-S Abnormal Condition Detection for Power Transmission Line Corridors[C]// 2021 IEEE 3rd International Conference on Power Data Science (ICPDS), IEEE, 2021: 13-16.
- [7] TIAN Z, SHEN C H, CHEN H, et al. Fcos: Fully Convolutional One-Stage Object Detection[C]// In Proceedings of the IEEE/CVF international conference on computer vision. 2019: 9627-9636.
- [8] DAI X, CHEN Y, YANG J, et al. Dynamic detr: End-to-End Object Detection with Dynamic Attention[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 2988-2997.
- [9] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. Scaled-YOLOv4: Scaling Cross Stage Partial Network[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13029-13038.
- [10] TAN M, PANG R, LE Q V, et al. Efficientdet: Scalable and Efficient Object Detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790.
- [11] ZHANG H, WANG Y, DAYOUB F, et al. Varifocalnet: An Iou-Aware Dense Object Detector[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 8514-8523.
- [12] DUAN KAIWEN, XIE LINGXI, QI HONGGANG, et al. Corner Proposal Network for Anchor-Free, Two-Stage Object Detection[C]// 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA), IEEE, 2021: 6-11.
- [13] ZHOU F, ZHAO H, NIE Z. Safety helmet detection based on YOLOv5[C]// 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA), IEEE, 2021: 6-11.
- [14] LAW H, DENG J. Cornernet: Detecting objects as Paired Keypoints[C]// Proceedings of the European Conference On Computer Vision (ECCV), 2018: 734-750.
- [15] YANG Z, LIU S, HU H, et al. Reppoints: Point set representation for object detection[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 9657-9666.
- [16] 魏智锋, 肖书浩, 蒋国璋, 等. 基于深度学习的人造板表面缺陷检测研究[J]. 林产工业, 2021, 58(2): 21-26.
WEI Zhi-feng, XIAO Shu-hao, JIANG Guo-zhang, et al. Research on Surface Defect Detection of Wood-Based Panels Based on Deep Learning[J]. China Forest Products Industry, 2021, 58(2): 21-26.
- [17] ZHAO Wei-dong, CHEN Feng, HUANG Han-cheng, et al. A New Steel Defect Detection Algorithm Based on Deep Learning[J]. Computational Intelligence and Neuroscience, 2021, 2021: 5592878.
- [18] ANWAR N, SHEN Z, WEI Q, et al. YOLOv4 Based Deep Learning Algorithm for Defects Detection and Classification of Rail Surfaces[C]// 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), IEEE, 2021: 1616-1620.
- [19] PATTANSHETTI S S, NIVADE S I. Real-Time Object Detection with Pre-eminent Speed and Precision using YOLOv4[J]. International Journal of Research in Engineering, Science and Management, 2021, 4(7): 26-31.
- [20] LIU S, QI L, QIN H, et al. Path Aggregation Network for Instance Segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [21] ZHU X K, LYU S H, WANG X, et al. TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-Captured Scenarios[C]//

- Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 2778-2788.
- [22] LIU Z, LIN Y, CAO Y, et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 10012-10022.
- [23] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all You Need[C]// Proceedings of the IEEE Conference On Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [24] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [25] 倪黎, 邹卫军. 基于 SE 模块改进 Xception 的动物种类识别[J]. 导航与控制, 2020, 19(2): 106-111.
- NI Li, ZOU Wei-jun. Recognition of Animal Species Based on Improved Xception by SE Module[J]. Navigation and Control, 2020, 19(2): 106-111.
- [26] 鲍峰, 王俊红, 张锋, 等. 基于 YOLO V3 的管道环焊缝缺陷检测[J]. 焊接, 2021(8): 56-61.
- BAO Feng, WANG Jun-hong, ZHANG Feng, et al. Inspection of Pipeline Girth Weld Defects Based on YOLO V3[J]. Welding & Joining, 2021(8): 56-61.
- [27] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single Shot Multibox Detector[C]// European Conference on Computer Vision, Springer, Cham, 2016: 21-37.
- [28] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

责任编辑: 曾钰婵