

基于多维度动态衰减 Transformer 的轮胎检测算法应用^{*}

朱焕宇 王明泉 贾虎 商奥雪 谢绍鹏

(中北大学信息与通信工程学院 太原 030051)

摘要: 针对当前国内子午线轮胎缺陷分割困难、成本昂贵的问题,本文提出了如下解决方案,提出一种基于 Swin Transformer 和注意力特征金字塔的子午线轮胎缺陷分割算法 Swin DAA,其中主要使用 Swin Transformer 作为主干特征提取网络,经过动态衰减注意力特征金字塔增强语义表达能力,搭建由 Python 语言编写的软件平台,同时级联 X 射线载重轮胎检测系统采集图像,并且使用 TCP 协议与上位机通信、传输图像数据,最终将缺陷分割软件系统与 MES 工控系统连接,完成无人监守的自动化子午线轮胎缺陷分割。实验对比数据显示,本文中提出的 Swin DAA 网络的精准度达到了 82.87%,召回率达到了 85.22%,每秒传输帧数达到了 11,所集成的软件能良好的完成子午线轮胎的实际监测要求。

关键词: 动态衰减;编解码器;深度学习;缺陷分割;子午线轮胎

中图分类号: TP391; TN919.8 文献标识码: A 国家标准学科分类代码: 520.6

Application of tire detection algorithm based on multi-dimensional dynamic attenuation Transformer

Zhu Huanyu Wang Mingquan Jia Hu Shang Aoxue Xie Shaopeng

(Information and Communication Engineering, North University of China, Taiyuan 030051, China)

Abstract: In response to the current difficulties and high costs in segmenting radial tire defects in China, this paper proposes the following solution: a radial tire defect segmentation algorithm called Swin DAA based on Swin Transformer and attention feature pyramid. Swin Transformer is mainly used as the backbone feature extraction network, and the semantic expression ability is enhanced through the Dynamic Attenuation Attention feature pyramid, Build a software platform written in Python language, cascade the X-ray heavy-duty tire detection system to collect images, and use TCP protocol to communicate with the upper computer and transmit image data. Finally, connect the defect segmentation software system with the MES industrial control system to complete unmanned automated radial tire defect segmentation. The experimental comparison data shows that the Swin DAA network proposed in this article has an accuracy of 82.87%, a recall rate of 85.22%, and a transmission frame rate of 11 per second. The integrated software can effectively meet the actual monitoring requirements of radial tires.

Keywords: dynamic attenuation attention; transformer; deep learning; defect segmentation; radial tire

0 引言

近些年来,国内人均消费水平逐渐提升,现代化的轿车逐渐走进人民的生活中,轮胎作为汽车工业的生命线,在工业生产中的重要性逐渐提升,轮胎的安全性也随之得到了重点关注。子午线轮胎通常分为胎冠、胎侧和子口部分,胎

冠区域为胎面行驶面范围,通常以第一根防擦线作为标准,胎侧区域通常为第一根防擦线到胎体反包断电或钢包反包端点范围,胎体反包端点或钢包反包端点到胎圈部位以内的一部分为子口区域。内部通常由钢丝帘线组成的帘布层构成,靠近胎冠中心位置通常由多层反复堆叠的带束层构成。常见的子午线轮胎缺陷分为钢丝帘线开口、钢丝帘线弯曲、

钢丝帘线稀疏、异物、气孔等缺陷^[1]。

当前国内在子午线轮胎^[2]钢丝帘线缺陷检测领域的研究甚少,通常采用的人工检测、传统图像检测法和神经网络法,人工检测法相对落后,通常指检测人员通过肉眼观察子午线轮胎 X 光成像图,精度和效率均由检测人员自身水平控制。传统图像检测法通常指利用图像处理的方法,例如图像滤波、去噪、边缘检测等方法明显化缺陷,便于检测人员查找。神经网络法是当前国内的主流方法,但检测的精度和效率与神经网络的选取有较强关联,且目前在多分类问题环境下,常用的神经网络表现较差,难以满足工业生产需求。

针对上述所提出国内当前在子午线轮胎缺陷检测困难的问题,本文提出由以下几点改进方案所组成的系统设计。首先利用全自动 X 射线载重轮胎检测系统(XTI15)采集图像,通过自主研发的通信模块完成上位机、下位机之间的通信,以及在上位机平台中完成缺陷的自动化检测,将缺陷判级的结果上传至 MES^[3]工控系统存储,完成子午线轮胎缺陷检测自动化的功能。其次提出 Swin DAA 网络,其中主干特征提取网络为 Swin Transformer,同时提出动态衰减注意力(dynamic attenuation attention, DAA)特征金字塔,

相比常规的特征金字塔具有更强的语义信息和空间信息的表达能力。本文所提出的 Swin DAA 神经网络与 DETR^[4]、YOLOv8、YOLOX、和 Swin Transformer^[5]相比,在子午线轮胎的工业生产环境下有更高的检测精度。

1 神经网络模块研发

1.1 Swin Transformer 模块

Swin Transformer 是由微软研究院于 2021 年公开发表的网络,由 4 个 Patch Merging 和 Swin Transformer 模块级联构成,Patch Merging 模块会将输入图像划分成 4 个 Patch,之后将每个 Patch 中相同位置的像素点拼接成特征图,并且将得到的 4 个特征图在纵向进行 Concat 堆叠,通过 LayerNorm 层归一化后,最后通过全连接层在纵向上做线性变化,降低特征图的深度。后续经过 Swin Transformer Block 完成特征提取,如图 1 中 Swin Transformer Block 所示,由 4 个全连接层、两个多层感知机(MLP)以及两个注意力机制模块构成,其中两个注意力机制模块分别是窗口多头注意力机制模块(W-MSA)以及移窗多头注意力机制(SW-MSA)^[6]。

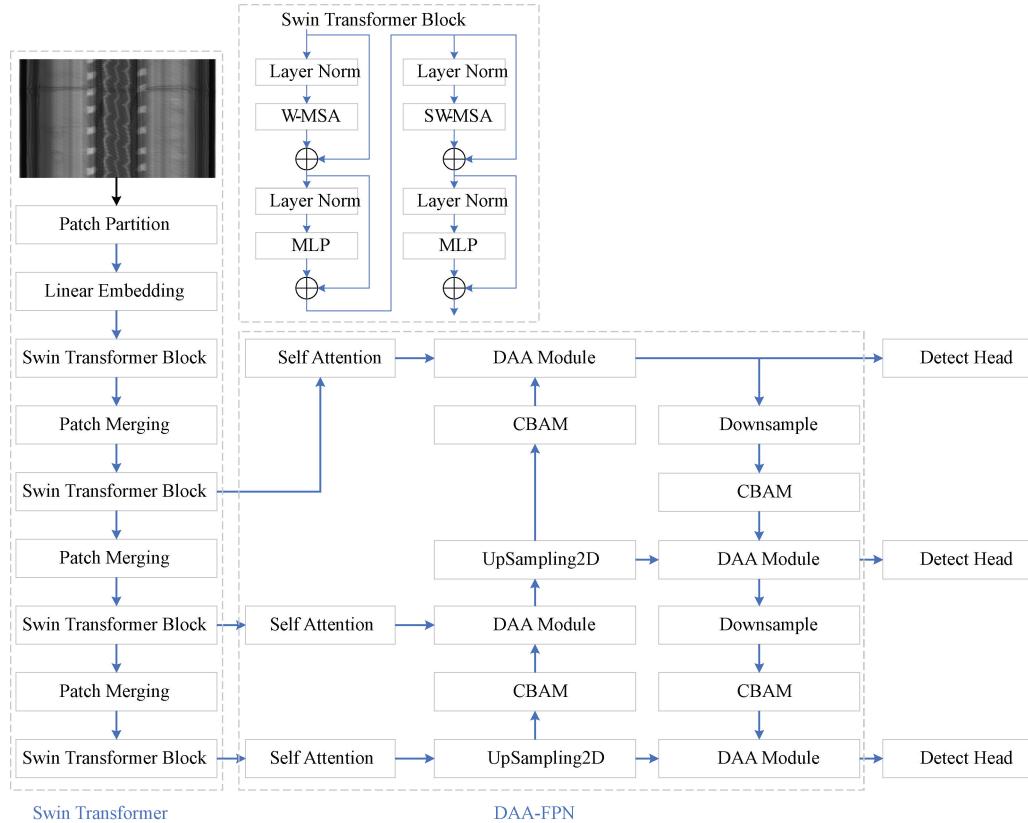


图 1 网络总体结构图

W-MSA 模块的作用是降低总计算量,通过将特征图划分为 4 个窗口,然后单独对 4 个窗口的内部进行自注意力运算,易知 W-MSA 模块会缩小感受野,不同窗口之间几乎没有信息交互,故在 W-MSA 模块后增添了 SW-MSA 模块,用

于完成窗口间信息交互,具体操作如图 2 所示,将原图向右下角移动一定像素,分割窗口位置不变,将原有的 4 个窗口分割成 9 个窗口,并且按照一定的规律进行重排序,分别将 FD、HB、ACIG 合并成一个窗口,最终可见与窗口多头注意

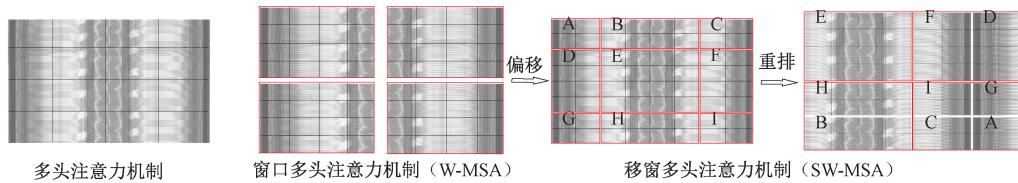


图 2 W-MSA 与 SW-MSA 机制

力机制一样均为 4×4 的窗口, 可见计算量未发生变化, 但在一定程度上完成相邻的窗口间完成信息交互^[7]。

1.2 动态衰减注意力金字塔模块概述

根据上述的 Swin Transformer 模块, 不难发现该模块具有感受野小, 同时特征提取能力要弱于卷积神经网络, 为解决上述问题, 本文提出如图 3 所示的动态衰减注意力特征金字塔模块, 首先将 Swin Transformer 主干特征提取网络中获取的 3 个大小不一的特征图经过自注意力机制扩充感受

野大小, 获取更多的特征依赖关系, 但该注意力机制获取有效信息的能力相对较小, 对原图的依赖性不强, 无法充分运用图像本身的尺度和平移不变性, 故该模块再融合交叉注意力机制, 更新特征图在通道维度上和空间维度上的权重, 如图 4 所示, 交叉注意力机制获取的特征图之与自注意力机制获取的特征图进行 L2 范数叠加, 并追加反向传播更新网络^[8], 最后经过 CSPLayer^[9] 残差网络增强特征提取能力, 增强特征表达能力和改善特征的语义信息^[10]。

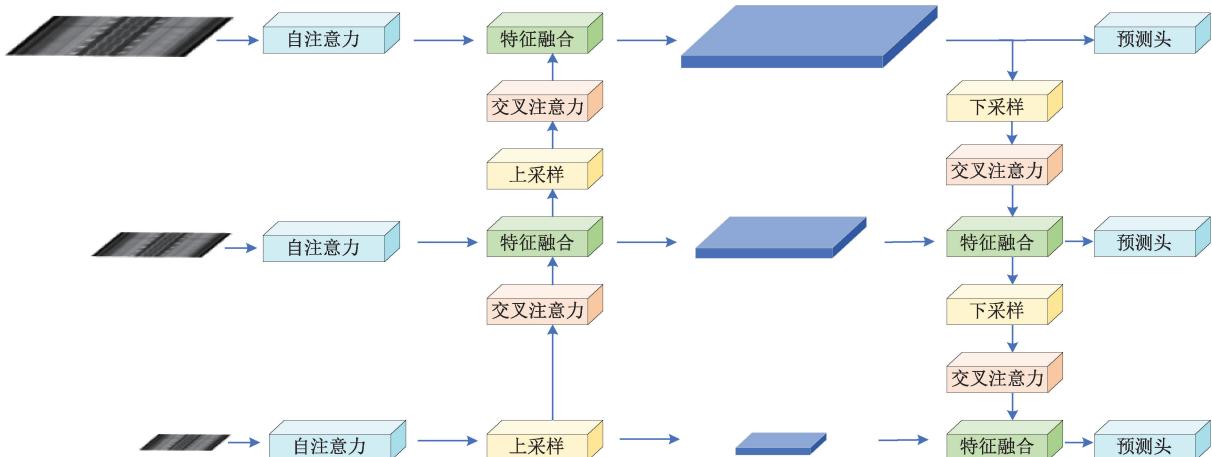


图 3 DAA 注意力特征金字塔模块

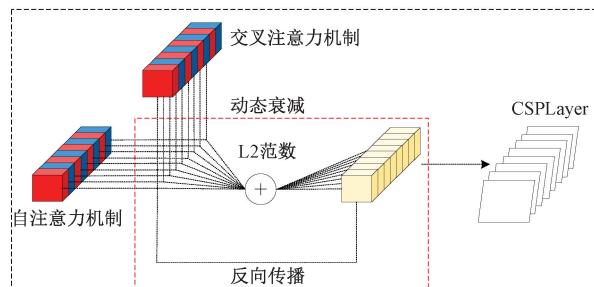


图 4 DAA 动态衰减注意力机制

通常将由多个样本成本函数写为:

$$J\theta = \frac{1}{m} \sum_{i=1}^m L(\hat{y}[i], y^{[i]}) \quad (1)$$

将其引入 L2 范数正则化^[11]后:

$$Jw, b = \frac{1}{m} \sum_{i=1}^m L(\hat{y}[i], y^{[i]}) + \frac{\lambda}{2m} \|W\|_F^2 \quad (2)$$

易知其在一定程度上抑制了过拟合^[12]现象, 通过对成本函数中斜率较高者给予惩罚, 斜率低者给予鼓励。而在

本文的 DAA 模块中, 输入端存在两个特征矩阵, 为兼顾两者的特征向量, 单纯逐像素相加效果并不理想, 故本文引入 L2 范数正则化过程, 设 x, y 分别为两个不同的特征向量:

$$\|x, y\|_2 = \sqrt{\sum_i^n (x_i - y_i)^2} \quad (3)$$

随后将式(2)带入上述式中, 可知:

$$\min[J(w^1, b^1, w^2, b^2)] = \frac{1}{m} \sum_{i=1}^m \mathcal{L}(\hat{y}(i), y^i) + \frac{\lambda}{2m} \sum_{l=1}^L \|x, y\|_F^2 \quad (4)$$

随后在等式两边对 w 求微分, 利用反向传播进行参数更新:

$$\frac{J}{dw} = dz \times a + \frac{\lambda}{m} w \quad (5)$$

$$w := w - a \times dw \quad (6)$$

将式(6)代入式(5)中得:

$$\omega := \omega - \frac{a\lambda}{m} \omega^T - a \left(\frac{\partial \frac{1}{m} \sum_{i=1}^m \mathcal{L}(\hat{y}^{(i)}, y^{(i)})}{\partial \omega} \right) \quad (7)$$

从等式可知, ω 权重无论如何均会减小, 其中 λ 通常称之为正则化参数, 使用验证集或者交叉验证来配置此参数, 故需要引入反向传播更新其权重。

通常情况下深层的特征层具有较强的语义信息和较弱的定位信息, 而浅层的特征层具有较强的定位信息和较弱的语义信息, 本文通过构建 DAA 动态衰减注意力机制, 保留了深层特征层的语义信息和浅层特征层的定位信息, 同时在另一方面增添交叉注意力机制(CBAM), 用于优化在上采样中被忽略了的不同特征通道之间的权重, 优化算法在空间维度上的权重, 提高网络的特征表达能力, 改善特征的语义信息。

1.3 网络结构与网络预设参数概述

本文中提出的神经网络结构总参数量为 139.989 M, 总计算量为 301.488 GFLOPS, 左侧为 Swin Transformer 主干特征提取网络, 负责提取出 3 个大小分别为 $80 \times 80 \times 256$ 、 $40 \times 40 \times 512$ 、 $20 \times 20 \times 1024$ 的特征图。右侧为本文提出的 DAA 注意力特征金字塔, 随后输出 3 个大小不同的 DetectHead, 用于作为检测时的预测头。本文中所有的开发环境基于 Python3.7、TensorFlow2.4.0、Pytorch1.10.1, 硬件环境中 GPU 为 RTX GeForce 2080Ti、CPU 为 Intel(R) Core(TM)i9 9900 k, 数据集来源于丹东市仪器仪表研究

所、中策橡胶集团有限公司、山东玲珑轮胎股份有限公司现场实际采取, 经过数据筛选增强处理, 共 500 张左右含缺陷的图像。

本文中所有的网络预设参数均相同, 采用 Swin Transformer Base 版本的预权重, 冻结阶段训练次数 Freeze epoch=100, Freeze batch size=8, 解冻阶段训练参数 UnFreeze epoch=500, UnFreeze batch size=4, 优化器使用的 Adam, 学习率的下降方式使用的 cos。

2 硬件模块研发与判级系统设计

2.1 通信调试模块开发

硬件模块在本文中的主要作用为传输经过 X 光成像的子午线轮胎图像, 并且完成系统间的信息交互, 通过匀速旋转子午线轮胎, 最终获取长条型的子午线轮胎 X 光成像图, 在通过预设的 TCP^[13] 报文协议, 报文格式如表 1 所示, 最终获取子午线轮胎灰度图像。通信调试模块的成像系统需要实时将接收到的固定行曝光数据转换成 8 位灰度图数组图像, 同时完成对缺陷的自动判级等功能。完成自动判级后需要与由西门子 S7-1515-2pn 设备进行信息交互, 获取子午线轮胎的重量、型号以及后续生成的报表信息等, 并且返回当前检测的子午线轮胎是否合格的信息。最终将人工判级结果与神经网络的判级结果上传至 MES 工控系统中, 同时生成缓存, 以便后续在对判级结果产生争执的情况下进行复检。

表 1 TCP 数据报文格式

包头	数据报文长度	图像偏移高度	图像宽度	图像高度	灰度值	包尾
AABBCCDDEEFF	20 00 28 00	00 04	00 14	00 02	AABBCCDDEEFF

2.2 判级标准及其系统设计

本文中的判级标准将全钢子午线轮胎分为以下 4 种类型, 合格品、整理胎、C 级品、D 级品, 合格品指不存在缺陷或存在微小缺陷的子午线轮胎, 整理胎指经过维修后可重新使用的子午线轮胎, C 级品通常不进行维修, 而是作为

原材料重新生产, D 级品指废品, 通常直接进行销毁操作。常见的子午线轮胎缺陷分为钢丝帘线开口、钢丝帘线弯曲、钢丝帘线稀疏、异物、气孔等缺陷^[14], 如图 5 所示, 其中各缺陷对应的判级标准如表 2 所示, 大多数都是通过对缺陷的数量或面积大小作为判级标准。



图 5 常见全钢子午线轮胎缺陷

表 2 缺陷判级标准

等级	帘线开口	帘线弯曲	帘线稀疏	异物	气孔
合格品	≤ 3 处	≤ 2.5 处	≤ 3 处	直径 ≤ 6 mm	—
整理胎	—	—	—	$6 \text{ mm} < \text{杂物直径} \leq 15 \text{ mm}$	直径 $\leq 100 \text{ mm}, \leq 3$ 处
C 级品	> 3 且 ≤ 5 处	> 2.5 且 ≤ 4 处	> 3 处	直径 $> 15 \text{ mm}$ (不在钢丝部件之间)	直径 $> 100 \text{ mm}$
D 级品	> 5	> 4	—	直径 $> 15 \text{ mm}$ (在钢丝部件之间)	> 8 处

判级模块主要体现在软件设计上,通过设计一个判级算法,即可接收神经网络的检测结果,对检测后的建议框进行统计数量、面积等操作,并通过一个逻辑判断,最终返回当前检测轮胎是否为合格品,并且上传至云端。根据上述方案,搭建了如图 6 所示的子午线轮胎缺陷自动判级系统。首先通过采集设备和 MES 工控系统获取当前检测的子午线轮胎条形码,随后在铅房中通过 X 光对子午线轮胎成像,并将图像数据传输至数据传输机中,之后通过 TCP 协议传输至核心判级软件,使用本文中所提出的 Swin DAA 神经网络进行缺陷分割,最终将数据与判定结果接

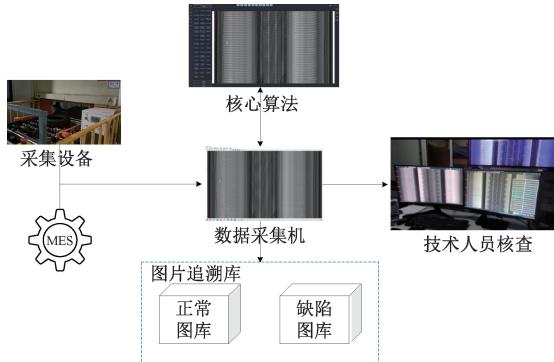


图 6 子午线轮胎缺陷自动判级系统

由技术人员判级。神经网络在本系统中用于提示操作人员指定位置可能有缺陷,提高操作人员的工作效率。

3 实验与分析

3.1 对比实验

在实验部分,本文引入未经优化的 DETR、YOLOX^[15]、YOLOv8、Swin YOLOX^[16]和 CD-YOLO^[17]作为对比实验组,开发环境为 Python3.7、TensorFlow2.4.0、Pytorch1.10.1,硬件环境中 GPU 为 RTX GeForce 2080Ti、CPU 为 Intel(R)Core(TM)i9 9900 k,实验结果如表 3 所示。根据表格中的数据易知,在异物和气孔缺陷上,受限于 YOLO 的锚框^[18],对于微小的缺陷如异物、气孔等检测的难度有一定的挑战,但随着总参数量、计算量和模型大小的增加,气孔的检测得到了一定的改善。而帘线缺陷通常尺寸较大,故上述网络均能在一定程度上精确定位帘线缺陷。

通过对上述网络性能参数数据的对比,本文提出的 Swin DAA 神经网络在识别子午线轮胎常见缺陷时,平均精准度达到了 82.87%,F1 值达到了 0.76,每秒可处理图片帧数为 11,综合性能超过了常规的传统图像检测法,且与上述的几种神经网络相比,具有更好的综合性能。

表 3 各模型对比实验结果

模型	精准度(Precision)/%				MAP	F1	计算量/GFLOPS	参数量/MB	FPS/(帧·s ⁻¹)
	帘线开口	帘线稀疏	异物	气孔					
DETR	90.37	71.15	17.16	8.37	46.76	0.49	208.949	55.702	10
YOLOv8	93.99	61.47	76.75	22.22	63.61	0.64	258.547	68.230	29
YOLOX	96.78	91.44	56.15	16.67	65.26	0.70	26.927	8.968	40
CD-YOLO	87.48	95.76	55.43	50.59	72.29	0.68	16.633	17.351	67
Swin YOLOX	90.64	95.00	69.15	66.67	80.37	0.74	273.556	92.574	14
Swin DAA	95.75	96.78	72.26	66.67	82.87	0.76	301.488	139.989	11

3.2 消融实验

为证明本文提出的设计方案对 Swin DAA 的网络性能评估系数有影响,故引入消融实验,将 Swin Transformer 主干特征提取网络和 DAA 模块分别替换 YOLOX 神经网络的 Backbone 和 FPN 模块,同时将未经优化的 DETR 和 YOLOX 作为对照组,开发环境为 Python3.7、TensorFlow2.4.0、Pytorch1.10.1,硬件环境中 GPU 为 RTX GeForce 2080Ti、CPU 为 Intel(R)Core(TM)i9 9900 k,做出以下 4 组对照组实验,实验结果如表 4 所示。

根据表格中 MAP 值易知,将 YOLOX 的主干特征提取网络替换成 Swin Transformer 模块,因其增添了大量的网络参数,会显著提升对微小缺陷的检测效果,而 DAA 模块对于网络的整体均有优化,各缺陷的精准度均有不小提升,在一定程度上提高了 MAP 值。

表 4 消融实验结果

模型	精准度(Precision)/%				mAP
	帘线开口	帘线稀疏	异物	气孔	
YOLOX	96.78	91.44	56.15	16.67	65.26
YOLOX DAA	97.09	96.51	68.94	11.11	68.41
Swin YOLOX	90.64	95.00	69.15	66.67	80.37
Swin DAA	95.75	96.78	72.26	66.67	82.87

3.3 泛化性能测试实验

在工厂实际检测过程中,通常有不同类型不同型号的子午线轮胎参与检测,且环境因素也会影响图像的成像效果,对网络的泛化性能要求较高,故引入泛化性能对比实验。该实验的测试集来源于后续实际采集所得,共计 1 000 张含缺陷的图像,统计各缺陷分布的结果作为样本

总数,通过观察召回率的大小来评价各方法在实际生产中的可靠性。实验结果如表 5 所示。

表 5 泛化性能实验

类型	帘线 开口	帘线 稀疏	异物	气孔	耗时/ h	召回率/ %
总数	1 068	993	715	816		
YOLOX	951	928	571	550	0.28	82.43
Swin DAA	964	942	587	576	0.31	85.22
人工检测	961	893	643	734	16.7	90.00

上述数据中人工检测的采样为中策轮胎公司、玲珑轮胎公司工业需求,仅作为参考数据,且本实验中涉及的数据来源于不同设备,不同规格的射线源采集获得。从上述数据不难得出,对于帘线开口、帘线稀疏的检测效果,大多都能达到精度要求,而在异物、气孔上,上述网络效果并不理想,距离工业生产需求仍有差距,但检测耗时远超人工检测效率,故本文提出的 Swin DAA 在一定程度上解决了子午线轮胎生产过程中缺陷难以检测、检测效率低下等问题。如图 7 所示,为泛化性能测试实验的实际效果图,可证明本文提出的 Swin DAA 具有一定的泛化性能。

4 结 论

综上所述,本文首次 Swin Transformer 作为子午线轮

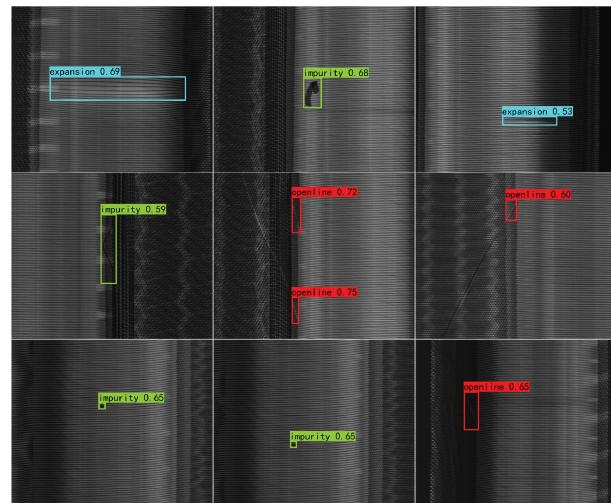


图 7 部分检测结果展示

胎缺陷目标检测的主干神经网络,并且提出一种全新的动态衰减注意力金字塔模型,增强特征提取能力和泛化能力,更适合用于解决当前国内大环境下子午线轮胎缺陷检测的难题,如图 8 为交工后的软件实际检测效果图。根据第四章对比实验、消融实验证明,本文中提出的 Swin DAA 神经网络的性能参数更佳,在精准度和召回率上均有不小提升。本文中提出的全钢子午线轮胎的自动判级系统已投入工厂实际生产,目前处于试用期。文中出现的所有网络、数据均可开源。



图 8 软件实际检测结果展示

参 考 文 献

- [1] 宋典武,王瑞均. 一种抗湿滑子午线轮胎胎面结构: CN202211678823.0[P]. CN116039296A[2024-04-19].
- [2] CAMPOLINA M, DIAS C, JR L J, et al. Static analysis of the carcass geometry influence on an automotive radial tyre using finite element model[J]. International Journal of Vehicle Systems Modelling and Testing, 2022, 16(4): 13-327.
- [3] 黄华星,韦海燕. MES 系统功能构架建设中的 IE 应用实践[J]. 装备制造技术,2023(7):213-215,223.
- [4] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]. Computer Vision-ECCV 2020: 16th European Conference, Glasgow, 2020.
- [5] 冷冰,冷敏,常智敏,等. 基于 Transformer 结构的深度学习模型用于外周血白细胞检测[J]. 仪器仪表学报, 2023,(5):113-120.
- [6] 王经纬,高艳鹏,宋瀚兴,等. 基于 Transformer 的飞机状态预测[J]. 计算机工程与设计, 2024, 45 (2): 477-483.
- [7] 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.
- [8] YUAN J L, ZHOU F, GUO Z T, et al. HCformer: Hybrid CNN-transformer for LDCT image denoising[J]. Journal of Digital Imaging, 2023, 36(5):2290-2305.
- [9] 陈星,欧鸥,杨长志,等. 基于改进 YOLOX 的落石检测方法[J/OL]. 计算机测量与控制:1-8[2023-09-15]. <http://kns.cnki.net/kcms/detail/11.4762.tp.20230713.1120.026.html>.
- [10] 吕振虎. 轻量化卷积神经网络的结构优化方法研

- 究[D]. 徐州:中国矿业大学,2022.
- [11] LIU Q X, DUAN G T, MATSUNAGA T, et al. A free-surface particle regularization scheme based on numerical integration for particle methods [J]. Engineering Analysis with Boundary Elements, 2023, 156:251-274.
- [12] RUI Q W, FEI Z, XU Y Z, et al. Training with scaled logits to alleviate class-level over-fitting in few-shot learning[J]. Neurocomputing, 2023, 522:142-151.
- [13] 冯学伟,徐恪,李琦,等.轻量级链式验证的网络传输层安全性增强方法[J/OL].软件学报:1-19[2023-09-16]. DOI:10.13328/j.cnki.jos.006941.
- [14] 赵鹏鹏,陈仁龙,李金屏.基于穿线法的胎侧散线缺陷检测[J].济南大学学报(自然科学版),2018,32(3):218-222.
- [15] FENG J T, WANG Z P, WANG S, et al. MSDD-YOLOX: An enhanced YOLOX for real-time surface defect detection of oranges by type [J]. European Journal of Agronomy, 2023. Doi: 10.1016/J.EJA. 2023.126918.
- [16] ZHENG X, QIAN S, WEI S, et al. The combination of transformer and you only look once for automatic concrete pavement crack detection [J]. Applied Sciences, 2023. Doi:10.3390/APP13169211.
- [17] 朱焕宇,王明泉,李磊磊,等.胎体钢丝帘线缺陷的CD-YOLO分割算法应用[J].国外电子测量技术,2023,42(7):175-180.
- [18] REDMON J, DIVVALA K S, GIRSHICK B R, et al. You Only Look Once: unified, real-time object detection[C]. 29th IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788.

作者简介

朱焕宇,硕士研究生,主要研究方向为图像识别与处理。

E-mail: zhuhuanyu6649@163.com

王明泉(通信作者),教授,主要研究方向为数字图像处理、分析和重建理论和应用研究。

E-mail: wangmq@nuc.edu.cn

贾虎,硕士研究生,主要研究方向为图像重建、人工智能。

E-mail: 1169214436@qq.com

商奥雪,硕士研究生,主要研究方向为图像重建方向。

E-mail: 1689669859@qq.com

谢绍鹏,硕士研究生,主要研究方向为图像识别与处理。

E-mail: 2090117685@qq.com