

基于 SP-DDN 的冲压件表面缺陷检测

张家伟¹, 孙芮², 陈浩¹, 刘银华¹

(1. 上海理工大学 机械工程学院, 上海 200093; 2. 上海交通大学 机械与动力工程学院, 上海 200240)

摘要: 车身冲压件表面缺陷具有缺陷尺寸跨度大、缺陷类别间差异小等特点, 导致传统表面缺陷检测手段的检测准确率低, 难以满足实际工业需求。本文提出一种通道自关联特征金字塔深度学习缺陷检测模型 SP-DDN, 该模型通过对特征金字塔结构进行扩展得到多层扩展特征融合网络 CFPN 结构, 并在该结构中引入具有通道关联分析能力的注意力模块, 以提高模型对多类别缺陷间差异性的特征提取能力, 提升模型的缺陷检测精确度。应用 K-means++ 算法对数据集中缺陷数据开展聚类分析, 以生成特定预选框, 解决默认预选框尺寸与实际缺陷尺寸不匹配问题。最后, 结合自制的冲压件表面缺陷检测数据集 SP-NET 与热轧钢表面缺陷公开数据集, 对提出方法进行应用验证。结果表明, 本文方法相较于基准网络, mAP 分别提升 3.1% 和 4.0%, 召回率分别提升 1.5% 和 5.5%。

关键词: 缺陷检测; 冲压件; K-means++; 自注意力机制

中图分类号: TP 18

文献标志码: A

Defect detection for stamping parts surface based on SP-DDN

ZHANG Jiawei¹, SUN Rui², CHEN Hao¹, LIU Yinhua¹

(1. School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China; 2. School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Traditional defect detection methods for body stampings face challenges due to the wide range of defect sizes and minimal differences between defect classes, resulting in low accuracy and limited industrial applicability. A deep learning defect detection model SP-DDN with channel self-associative feature pyramid was proposed. The model extended the feature pyramid structure to obtain a new feature fusion CFPN structure, and added a channel correlation analysis attention module to the structure to improve the feature extraction capability of the model for the differences between multiple

收稿日期: 2023-08-05

基金项目: 国家自然科学基金面上资助项目(51875362); 上海市自然科学基金资助项目(21ZR1444500); 上海市浦江人才计划(22PJD048)

第一作者: 张家伟(1998-), 男, 硕士研究生。研究方向: 机器视觉与智能检测。E-mail: zjw_yeah98@163.com

通信作者: 刘银华(1983-), 女, 教授。研究方向: 数字化设计与制造。E-mail: liuyinhua@usst.edu.cn

引文格式: 张家伟, 孙芮, 陈浩, 等. 基于 SP-DDN 的冲压件表面缺陷检测[J]. 上海理工大学学报, 2024, 46(6): 647-658.

Citation: ZHANG Jiawei, SUN Rui, CHEN Hao, et al. Defect detection for stamping parts surface based on SP-DDN[J]. Journal of University of Shanghai for Science and Technology, 2024, 46(6): 647-658.

classes of defects and enhanced the defect detection accuracy of the model. K-means++ was applied to cluster the defect data in the dataset to generate specific pre-checked boxes and solved the problem of mismatch between the default pre-checked box size and the actual defect size. Finally, application validation was carried out using a self-made surface defect dataset of stamped parts and a publicly available surface defect dataset of hot rolled steel. The results show that the method in this paper improves the mAP by 3.1% and 4.0% and the recall by 1.5% and 5.5%, respectively, compared with the benchmark network.

Keywords: defect detection; stamped parts; k-means++; self-attention mechanism

随着汽车工业的不断发展，汽车冲压件的表面质量越来越受到关注。由于冲压件表面缺陷尺寸跨度大、缺陷类别间差异小等问题，传统基于油石等辅助方式下的目视检测一直是冲压件表面缺陷检测的最主要手段，但受到检测人员主观性、疲劳等因素影响，检测准确率与检测效率难以满足工业需求。近年来，基于深度学习的目标检测技术逐步成为表面缺陷检测的研究热点，并逐渐应用于缺陷检测领域，相较于传统视觉检测方法^[1-2]，该类方法具有泛化能力强、检测鲁棒性高以及检测速度快等优点。

基于深度学习的缺陷检测框架主要包括两种形式，即一阶段网络和两阶段网络。一阶段网络如 YOLO^[3]、SSD^[4] 和 RetinaNet^[5] 等，通过特征提取即可实现缺陷检测，具有背景误检率低、检测速度快等优点。两阶段网络由于添加候选区域生成模块，网络的目标定位能力以及小目标检测性能较强。常见的两阶段检测网络有：R-CNN^[6]、Fast R-CNN^[7]、Faster R-CNN^[8]、Cascade R-CNN^[9] 等。上述研究为基于深度学习的目标检测方法应用奠定了方法基础。但在具体检测对象的缺陷检测准确性方面仍需进一步提升。为此，学者们针对特定对象的缺陷特点与目标检测框架开展了大量算法改进研究。

在一阶段网络方面，Cheng^[10] 等通过在 RetinaNet 网络中嵌入通道注意力机制，使用具有融合卷积核提取的浅层和深层特征能力的自适应空间特征融合模块，提高网络缺陷特征提取能力。Zheng^[11] 为解决轴承盖表面细微缺陷检测难、检测实时性不足等问题，构建了大规模轴承盖缺陷数据集，并基于瓶颈注意力增强网络提升缺陷特征提取能力。Chang 等^[12] 开发了 TinyDefectNet 缺陷检测系统，使用真实缺陷样本数据进行模型训练，实现了车身漆面微小缺陷的抽样检测。该系

统算法主要是通过丰富缺陷图像数据尺寸以增加特征提取数据形式提高检测准确性。朱文博等^[13] 提出了一种基于改进 YOLOv5 的 O 型圈表面缺陷检测方法，平均精度提高了 4.26%。在两阶段网络方面，Huang 等^[14] 基于 Cascade R-CNN 提出了一种多尺度特征对特征提取方法，在对 RPN 输出的区域进行池化前将特征图扩展 1.5 倍，以增强目标的详细信息并提高检测精度；Ding 等^[15] 针对表面缺陷检测误检率和漏检率高、效率低问题，提出基于 ResNet-101 和特征金字塔网络 (feature pyramid network, FPN) 进行特征提取、融合的微小缺陷检测网络 TDD-net，实验结果表明，该网络具有较强的鲁棒性和泛化性。

综上所述，现有研究多通过改善特征提取方式来提高检测准确率，但现有网络中对缺陷特征变化导致的特征丢失问题、缺陷间的异性识别方面仍存在不足。鉴于车身冲压件表面划痕、压痕等同类缺陷尺寸跨度大，异类缺陷间差异小等问题，如何结合冲压件表面缺陷形态和纹理特征，提取更具区分度的特征，是提升冲压件表面缺陷检测精度的关键。为此，本文提出以 Faster R-CNN 结构为基本骨架，以 ResNet50 为骨干特征提取网络的 SP-DDN 缺陷检测算法，通过引入自注意力机制和可变卷积的通道特征自关联模块与聚类分析，提升模型对多尺寸、多类别缺陷的识别与定位能力，并通过案例分析验证提出方法的有效性。

1 SP-DDN 网络模型

本文提出的冲压件表面缺陷检测模型 SP-DDN 网络结构如图 1 所示。SP-DDN 沿用 Faster R-CNN 网络框架，主要由 5 部分组成：特征提取骨干网络、通道自关联特征金字塔网络 (channel

self-association feature pyramid network, CFPN)、区域生成网络(region proposal network, RPN)、感兴趣区域池化(region of interest pooling, ROI Pooling)和分类回归网络。首先,采用 ResNet50 残差网络作为骨干特征提取网络,对缺陷中的抽象特征进行提取,该过程按照特征图尺寸大小可分为 4 个阶段,每个阶段输出的特征图将作为

CFPN 网络的输入。其次,将经过 CFPN 特征融合后的特征图输入融合自注意力机制和可变形卷积的 SCBAM 模块,以进行通道类间差异性信息自关联分析。然后,将各阶段特征图输入 RPN 生成区域建议预选框,并通过 ROI Pooling 进行区域建议框映射。最后,将特征图送入分类回归网络,进行回归与类别预测。

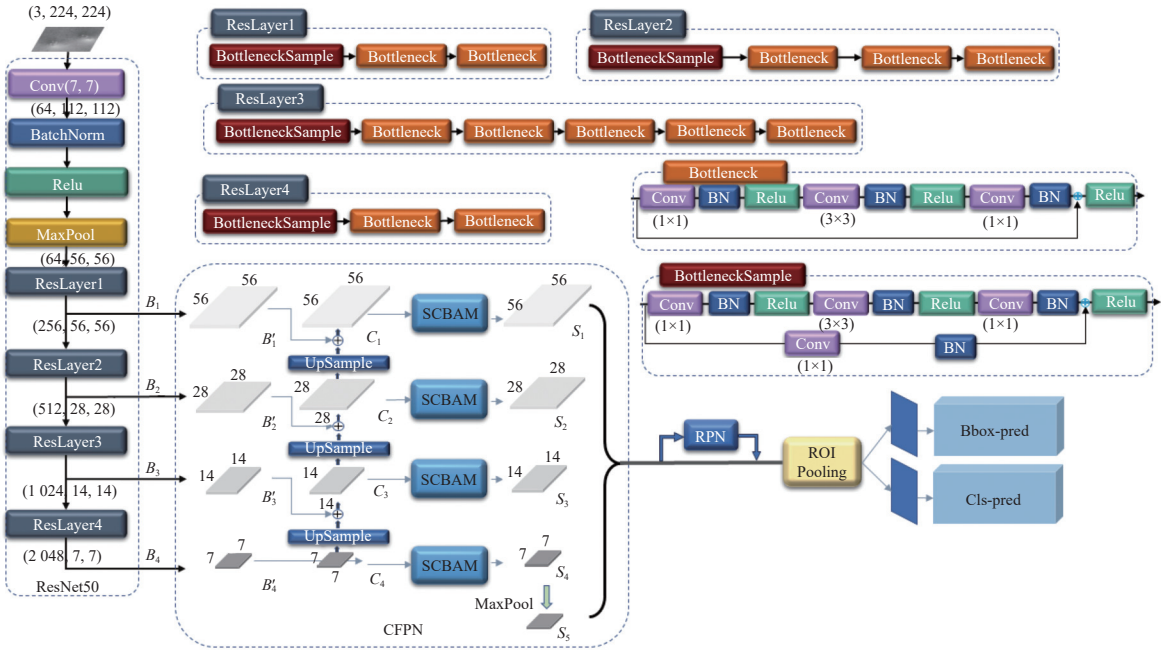


图 1 SP-DDN 网络模型结构

Fig.1 Structure of the SP-DDN network

1.1 通道自关联特征金字塔网络

如图 2 所示,在真实工业现场中,缺陷尺寸跨度较大,尺寸较小的缺陷在图像中仅占十余个像素,而大缺陷则可能横跨整张图像。在特征提取骨干网络中,模型通过一系列卷积操作提取图像的抽象特征信息。然而,随着网络深度的增加,特征图的层级加深,分辨率却逐渐减小。这导致原始图像中不同大小的缺陷所具有的细节特征在分辨率下降过程中丢失。对于较小的缺陷,会出现漏检情况。而对于跨度较大的缺陷,则会出现对缺陷尺寸估计出现偏差,进而影响缺陷的全局定位。为此,本文部署了通道自关联特征金字塔网络(CFPN)。

CFPN 为“4+1”型金字塔结构,该网络通过对较深层中分辨率较低的特征图进行最近邻插值上采样,并与骨干网络中分辨率较高的浅层特征图相加,该过程将不同分辨率的特征进行了融合,实现了不同尺度特征的共享。嵌入 SCBAM

结构的 FPN 加强了不同通道间相关语义信息的应用,为增强模型对多尺度缺陷的全局定位能力,对 SCBAM 输出的最深特征图进行最大池化,得到感受野更大的“+1”特征图,最终形成 CFPN。其中 FPN 结构输出的特征图可表示为

$$C_i = \begin{cases} \text{Conv}(B_i) + U_{2\times}(C_{i+1}), & i = 1, 2, 3 \\ \text{Conv}(B_i), & i = 4 \end{cases} \quad (1)$$

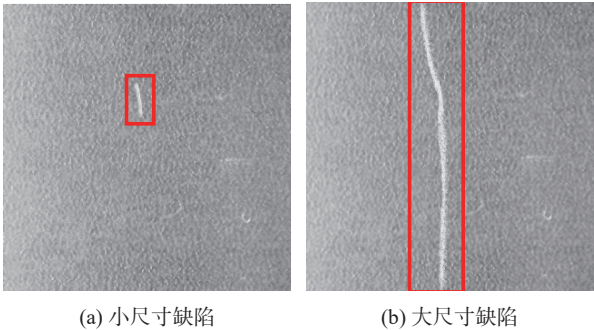


图 2 缺陷尺寸对比图

Fig.2 Defects of different sizes

式中： $\mathbb{C}_i$ 表示CFPN结构中多层特征图融合阶段的第*i*层输出的特征图； B_i 表示由骨干网络(ResNet50)生成的特征图； C_{i+1} 表示待融合的深层特征图； $Conv(\cdot)$ 表示 1×1 卷积操作，其作用为将特征图通道维度整合为256； $U_{2x}(\cdot)$ 代表最近邻插值上采样，上采样倍率取2。 $\mathbb{C}_i$ 作为SCBAM模块的输入，SCBAM的输出结果为 S_i ， $i=\{1,2,3,4\}$ ，对 S_4 进行步长为2的最大池化，得到特征图 S_5 。

1.2 可变卷积注意力模块

特征图中呈现相似外观的不同种类缺陷，不同种类的缺陷共享一些相似的形状、纹理特征。这增加了模型正确识别和区分不同缺陷类别的难

度。为了解决这一问题，受卷积块注意模块(convolutional block attention module, CBAM)^[16]的启发，本文在CBAM的基础上设计了SCBAM模块来增强模型对通道信息的利用率，提升模型的检测精度。SCBAM模块中，多头自注意力机制的通道自注意力模块关注不同通道的重要程度，帮助区分出更具代表性的通道中的语义特征，使得模型可以更好地利用这些语义特征进行计算。可变卷积空间注意力机制通过引入额外的偏移空间采样位置，以适应不规则形状缺陷。

如图3(a)所示，SCBAM整体采用残差结构，主通道采用 3×3 卷积块与改进的CBAM模块并行

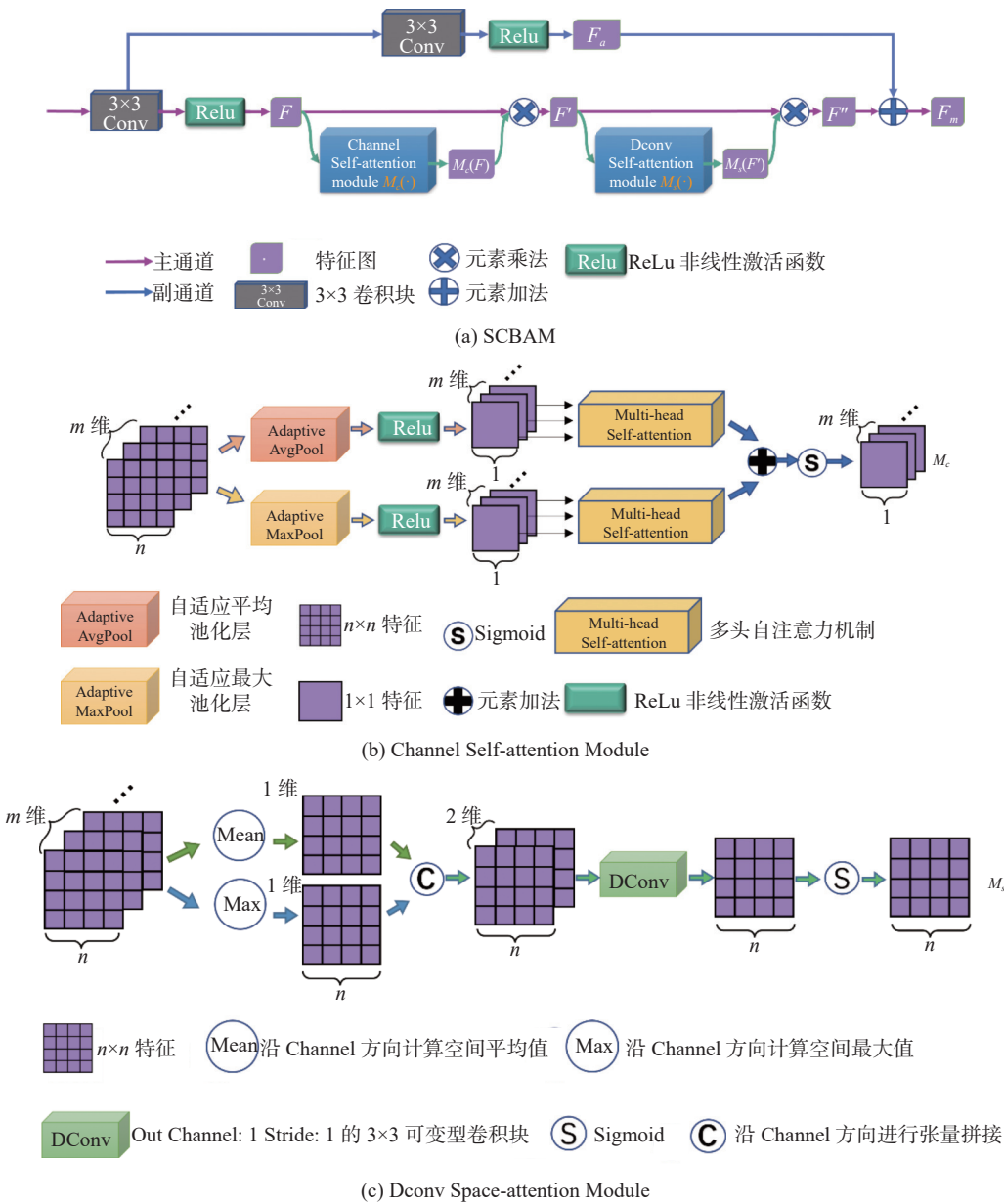


图3 SCBAM 模块结构

Fig.3 Structure of SCBAM module

连接, 副通道并联一个 3×3 卷积块。主通道上的特征信息 $F \in R^{C \times H \times W}$ 先后与通道自注意力模块的输出 $M_c \in R^{C \times 1 \times 1}$ 、可变卷积空间注意模块的输出 $M_s \in R^{1 \times H \times W}$ 进行元素乘法, 整个过程表示为

$$F' = M_c(F) \otimes F \quad (2)$$

$$F'' = M_s(F') \otimes F' \quad (3)$$

式中, $\otimes$ 表示元素乘法, 该模块最终输出 F'' 。图 3(b) 和图 3(c) 分别展示了两个子模块的计算过程。

原有的通道注意力模块分别采用最大池化与平均池化对特征图在空间维度进行压缩, 以表示各通道的聚合空间信息, 并将这些信息输入到多层感知器, 以求得通道注意力图 $M_c \in R^{C \times 1 \times 1}$ 。在此基础上引入多头自注意力机制, 通过计算类间特征的关联, 输出具有通道信息强关联性的注意力图。具体来讲, 沿用原有的通道注意模块的并联结构, 分别采用自适应平均池化与自适应最大池化操作对特征图各通道进行空间信息聚合, 得到通道特征分数 I_{avg} 和 I_{max} , 然后, 将 I_{avg} 和 I_{max} 输入自注意力模块, 得到注意力分数 A_{avg} 和 A_{max} , 模块内部运算可表示为

$$A(I) = W_v(I) S \left(D_P \left((W_k(I))^T, W_q(I) \right) \right) \quad (4)$$

$$W_i(I) = W^i I, \quad i \in v, k, q \quad (5)$$

式中: $W^i, i \in \{v, k, q\}$ 分别表示自注意力机制中的价值集、键值集、查询集参数矩阵; $D_P(\cdot)$ 表示点乘运算; $S(\cdot)$ 表示 SoftMax 函数。

将 A_{avg} 和 A_{max} 进行元素加法后经 Sigmoid 函数处理便得到通道自注意力模块的输出 M_c , 可表示为

$$M_c = \text{Sigmoid}(A_{avg} + A_{max}) \quad (6)$$

原空间注意模块侧重于关注目标的位置信息, 采用 3×3 卷积进行特征提取, 而标准卷积中规则的卷积格点难以适应复杂多样的缺陷形状, 为此本文将标准卷积替换为可变型卷积^[17], 通过额外偏移空间采样位置方式, 提升网络模型对形状复杂特征的建模能力, 最终得到改进的空间注意模块, 其结构如图 3(c) 所示。具体为, 沿通道方向对空间像素点分别取最大值与平均值, 得到聚合通道信息 $P_{max} \in R^{1 \times H \times W}$ 和 $P_{avg} \in R^{1 \times H \times W}$, 并沿通道方向进行特征拼接, 得到二维特征 $P \in R^{2 \times H \times W}$, 同时应用可变形卷积操作并经过 Sigmoid 函数处理后生成空间注意力图 M_s , M_s 的计

算过程为

$$M_s = \text{Sigmoid}(D\text{Conv}(\text{Cat}(\text{Mean}(I), \text{Max}(I)))) \quad (7)$$

式中: $D\text{Conv}(\cdot)$ 表示可变型卷积, 卷积核大小为 3×3 ; $\text{Cat}(\cdot)$ 表示沿通道方向进行元素拼接; $\text{Mean}(\cdot)$ 表示沿通道方向对空间像素取平均值; $\text{Max}(\cdot)$ 表示沿通道方向对空间像素取最大值。

2 基于 K-means++ 的缺陷预选框重优化

SP-DDN 中, 使用了 Faster R-CNN 中的 RPN 结构来生成矩形框。RPN 首先在输入特征图的每个像素位置生成多个初始预选框, 其形状自定义以覆盖不同缺陷形状。这些初始预选框提供模型最初的候选矩形框样本, 然后网络会逐步减少预选框数量, 直至与实际缺陷数量匹配, 并对保留的初始预选框的形状和位置进行微调, 确保最终生成的矩形框能够准确地框选缺陷。初始预选框自定义时具有合理尺寸和形状, 有助于减少模型搜索空间, 提高对缺陷的准确检测和定位能力。

为了让初始预选框自定义生成过程具备更合理的尺寸和形状, 本文采用了 K-means++ 算法对训练集中的标签数据进行聚类分析。K-means++ 是对 K-means 的改进, 主要优化了选择初始聚类中心的过程。传统 K-means 算法随机选择初始化聚类中心, 聚类结果受初始聚类中心影响, 不利于得到全局最优值。而 K-means++ 通过引入概率分布函数来选择初始聚类中心, 使得选择下一个聚类中心时更有针对性, 从而克服初始聚类中心选择不佳的问题。首先, 设置 K 为聚类簇的数量, 并随机选择一个数据样本作为第一个聚类中心。然后, 分别计算每个样本到已有聚类中心的最短欧氏距离, 计算过程如式 8 所示。然后, 根据距离计算结果, 选择距离较远的样本为新的聚类中心, 以保证各中心间的距离相对较大。重复上述过程, 直到得到 K 个初始聚类中心。之后, 根据 K 个初始聚类中心循环聚类、划分、更新簇直至分成 K 个簇并得到聚类中心。

$$\rho = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \quad (8)$$

式中: ρ 代表两个样本间的欧氏距离; x_i, y_i 分别代表每个样本框的高度与宽度; x_c 和 y_c 分别代表聚类中心的高度与宽度。图 4 中列出了 K-means++ 应

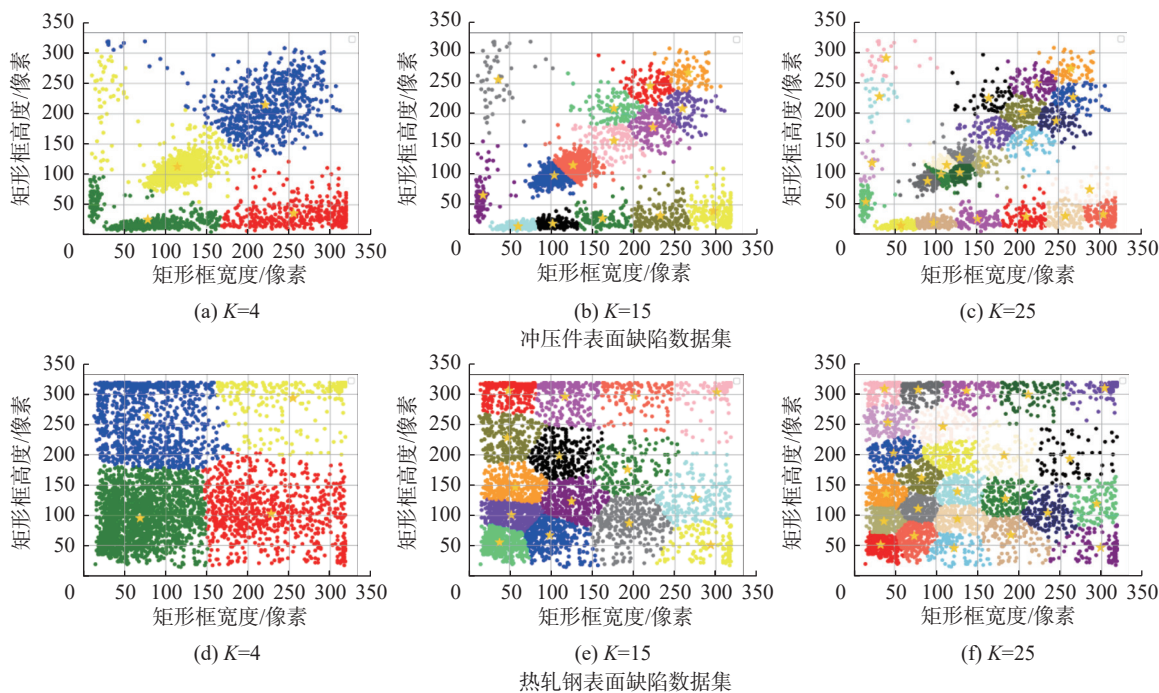


图 4 K-means++在两个表面缺陷数据集上对预选框的聚类结果

Fig.4 K-means++ clustering results for pre-selected boxes on two surface defect datasets

用于两个不同表面缺陷数据集的聚类结果。在各图中，每个圆点表示训练集中的一个缺陷样本的人工标注矩形框。圆点对应的横坐标为矩形框的宽度，纵坐标为矩形框的高度。K-means++算法根据自定义的 K 值，根据人工标注矩形框的宽度信息与高度信息对训练集中所有的人工标注矩形框进行聚类，在图中不同聚类簇使用不同颜色进行区分。聚类中心用金色星号表示，其到簇内样本的平均距离最小，因此聚类中心可以反应整个簇的代表性。本文将聚类中心所对应的宽度、高度作为初始预选框的宽度、高度。将聚类簇的种类数 K 作为初始预选框的种类数。

3 冲压件表面缺陷数据集

由于实际工业生产过程中良品率高，冲压件表面缺陷数据的样本不足。为此，本文通过自主构建冲压件表面缺陷数据集的方式，为开展研究提供数据基础。数据集的构建过程包括图像采集、标注及数据预处理。为确保数据集的有效性和可靠性，数据集制作过程的关键步骤包括：

a. 鉴于光照条件对图像质量和缺陷检测准确性的显著影响，本文在数据采集时采用图 5 所示的侧面白色光源、明场照明方式。

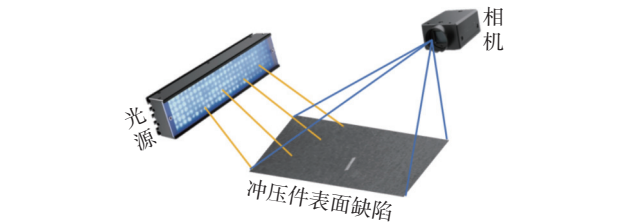


图 5 冲压件表面图像采集示意图

Fig.5 Camera and light source placement during defective image acquisition

b. 为获取高质量图像，本文采用 CMOS 传感器相机进行图像采集。

c. 标注是保证数据集质量的关键步骤，使用 Labeling 图像标注软件进行缺陷标注。

d. 为增强数据集的语义性和适用性，选取凹坑、凸起、划痕、压痕等常见缺陷类型，并且每张图像包含多个缺陷实例。

e. 为增强缺陷特征的可检测性，本文对图像进行了灰度化和二维自适应滤波去噪处理，以减少噪声和增强缺陷特征可检测性。

基于上述流程，构建了如图 6 所示的冲压件表面缺陷数据集 SP-NET，该数据集包含 2028 张缺陷图像，最小的缺陷直径小于 1 mm，缺陷类型分别为：凹陷、划痕、凸起、压痕。各缺陷类型数目如表 1 所示，以此为冲压件表面缺陷检测和

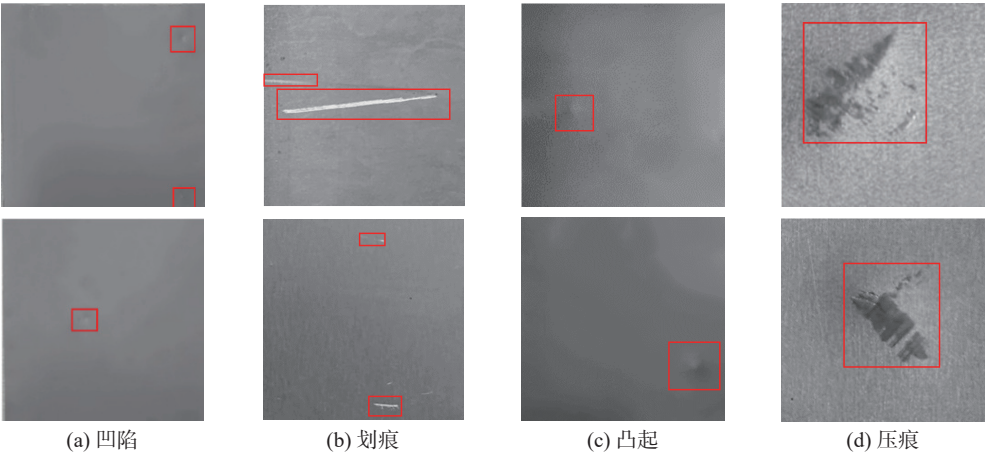


图 6 冲压件表面缺陷数据集图片样例

Fig.6 Sample image of defect data set on the stamped part surface

表 1 SP-NET 数据集图像数据分布情况表

Tab.1 Distribution of image data in SP-NET dataset

缺陷类型	训练集	测试集	总和
凹痕	284	76	360
划痕	493	113	606
凸起	325	93	418
压痕	520	124	644

识别算法验证提供数据基础。

4 实验及结果分析

4.1 实验数据集

本文分别结合自制冲压件表面缺陷数据集 SP-

NET 和热轧钢带的开源数据集 NEU-DET^[18] 对上述方法开展验证。NEU-DET 数据集在缺陷检测领域被广泛使用，它包含 6 种缺陷类型，分别为：氧化（rolled-in scale）、斑块（patches）、裂纹（crazing）、麻点（pitted surface）、杂质（inclusion）、划痕（scratches），共包含 1800 张图片。表面缺陷样本如图 7 所示，实验过程中数据集的具体分配如表 2 所示。

4.2 算法评价标准与实验设置

4.2.1 算法评价标准

为评价 SP-DDN 网络的针对多类别缺陷检测的能力，采用平均精度均值（mean average precision, mAP）作为衡量目标检测算法的性能指标，其值越高说明模型在所有类别上的性能越

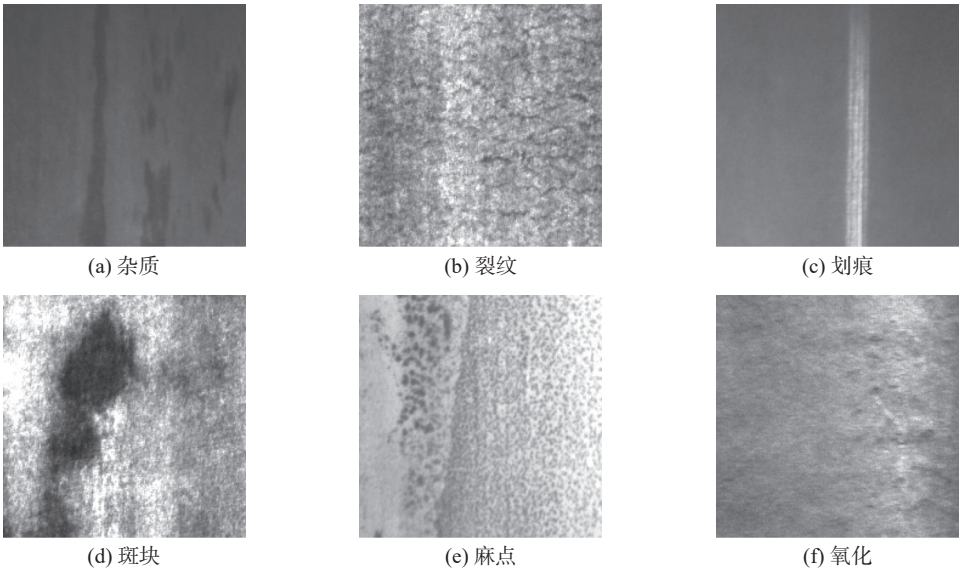


图 7 NEU-DET 数据集缺陷样本

Fig.7 Samples of defects in the NEU-DET dataset

表 2 NEU-DET 数据集图像数据分布情况
Tab.2 Distribution of image data in the NEU-DET dataset

缺陷类型	训练集	测试集	总和
氧化	241	59	300
斑块	234	66	300
裂纹	239	61	300
麻点	242	58	300
杂质	249	51	300
划痕	235	65	300

好，其表达式如下：

$$\text{mAP} = \frac{\sum AP}{C} \tag{9}$$

式中： C 表示检测缺陷类别数； AP 为平均精确率（average precision），它是对每个类别的精度均值的平均值，其数值越高，说明模型在该类别上的性能越好。平均精确率 AP 、精确率 P 与召回率 R 定义为

$$AP = \int_0^1 P(R) dR \tag{10}$$

$$P = \frac{TP}{TP + FP} \tag{11}$$

$$R = \frac{TP}{TP + FN} \tag{12}$$

式中： TP (true positive) 表示正样本被正确识别为正样本，本文将锚框与真实标签之间重叠率大于 0.5 的识别结果视为 TP ； FP 表示负样本被错误识别为正样本； FN (false negatives) 表示正样本被错误识别为负样本。同时，漏检率是工业目标检测领域中一项重要指标，由于召回率 R 表示检测正确的缺陷占全部缺陷的比例，召回率越高，缺陷漏检率越低，模型在识别目标方面的性能越好，因此本文将召回率作为模型评价的另一指标。此外，采用图像推理速度 (frames per second, FPS) 作为衡量模型检测速率的评价指标，并计算各网络模型的参数量 (Param)。

4.2.2 实验设置

实验过程中采用的硬件配置参数如下，CPU: AMD Ryzen 7 5800H CPU@3.20GHZ，GPU: NVIDIA GeForce RTX 3060 6G 显存，操作系统: Windows 11。采用基于 Python3.8，pytorch1.8 的 MMDetection 深度学习目标检测框架，并通过 CUDA11.1，CuDNN8.0.5 进行 GPU 加速。

根据硬件性能，epoch 设置为 20，batch size 设置为 4，采用经过 ImageNet 训练得到的 resnet50 模型权重作为本文模型的预训练权重，参数优化器采用随机梯度下降 (stochastic gradient descent, SGD)，学习率为 0.02。冲压件表面缺陷数据集 SP-NET 采用 320×320 图像大小进行网络训练，而开源数据集 NEU-DET 采用 224×224 图片大小进行训练。

4.3 模型验证

图 8 中为 SP-DDN 应用于两个数据集的检测效果示例，其中第 1、2 行为自制冲压件表面缺陷数据集的检测结果，第 3、4、5 行为热轧钢带表面缺陷开源数据集的检测结果。

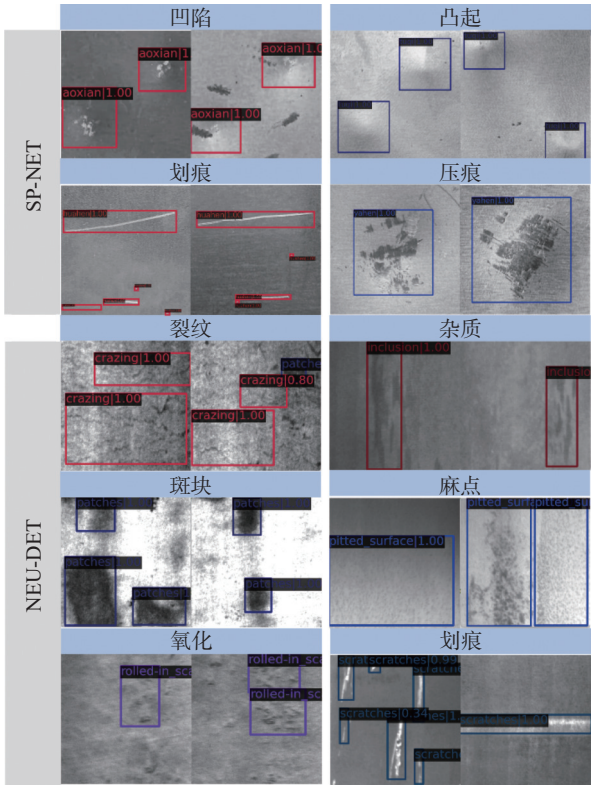


图 8 缺陷检测结果
Fig.8 Defect detection results

本文采用消融实验验证 SP-DDN 的性能和各组件对模型性能的贡献程度。所有实验基于两个数据集开展，结果如表 3 所示。

对上述缺陷检测的结果分析如下：

a. 用于对比的基准网络采用特征金子塔层数为 4 层的 Faster R-CNN FPN 网络。从表 3 中可以看出，在两个数据集上，CFPN 结构的应用使得模型的 mAP 分别提升 0.5%，0.5%，模型召回率 R 分别提升 0.5%，1.6%，由此可见“+1”层金字塔结

表 3 消融实验结果

Tab.3 Results of ablation experiments

CFPN	SCBAM	K-means++	SP-NET		NEU-DET	
			mAP / %	R / %	mAP / %	R / %
			90.3	97.4	71.1	82.0
√			90.8	97.9	71.6	83.6
√	√		93.0	98.4	74.7	84.8
√	√	√	93.4	98.9	75.1	87.5

构提升了模型的检测准确率和缺陷查全率。

b. 在 Faster R-CNN CFPN 中, 模型通过特征金字塔结构将骨干网络不同阶段多尺度特征信息进行融合, 导致大量的特征信息丢失。试验嵌入 SCBAM 模块, 通过自注意力机制分析通道间特征的关联, 以提高模型的检测准确性。如表 3 所示, 在两个数据集上, 模型因 SCBAM 的嵌入, mAP 分别提升了 2.2%、3.1%, 同时召回率分别提升了 0.5%, 1.2%。

c. SCBAM 模块中的 Multi-head Self-attention 部分为多头自注意力机制模块, 负责对不同通道特征信息进行关联性分析, 该模块有两个主要的自定义参数分别为注意力头个数与隐藏层权重个数, 其中注意力头的数量决定了模块可以同时关注的不同信息源的数量, 而隐藏层权重数量代表注意力权重向量的维度, 决定了模块对信息关联性的表达能力。本试验在缺陷类型更丰富的 NEU-DET 数据集上分析了模块安装位置及自身参数设置对于模型性能的影响, 实验结果如表 4 所示, 其中 C4/150 具体含义如下: C4 代表将 SCBAM 模块置于图 1 的 C4 线路上, 150 代表 Multi-head Self-attention 中注意力头个数×隐藏层权重数量。通过

表 4 SCBAM 模块位置与参数分析实验结果

Tab.4 Experimental results of SCBAM module location and parameter analysis

实验序号	组合形式	mAP / %	R / %
①	基准网络	71.6	83.6
②	C4/150	71.6	83.5
③	C4/300	72.4	84.2
④	C1/150	71.3	82.8
⑤	C1/200	72.0	82.9
⑥	C3&C4/300	72.6	83.9
⑦	C2&C3&C4/300	73.4	83.7
⑧	C1&C2&C3&C4/300	75.1	87.5
⑨	C1&C2&C3&C4/30	74.7	84.8

②③和④⑤可以得出当模块安装位置相同时, 模块参数量越多, 模型的检测精度与召回率越高。且模块安置于感受野范围更大的 C4 路线时, 对模块产生更加积极的影响。通过③⑥⑦⑧可以得出当模块数量逐渐增加时, 模型的检测精度与召回率逐渐提升。同时, 分析⑧⑨可以得出模块数量相同时, 参数量越多模型性能越好的规律。综合考虑模型检测精度、召回率以及模型参数的影响, 选择⑧中所示的模块数量及参数设置进行应用。

d. 预选框聚类提升模型性能: 采用 K-means++算法对两个数据集中缺陷样本的标定框聚类分析, 自制冲压件表面缺陷数据集的聚类结果如图 9(a)所示, 实验过程中将 K 设为 25。热轧钢表面缺陷数据集的聚类结果如图 9(b)所示, 试验过程中将 K 设为 25。最终模型在两个数据集上的 mAP 分别提升 0.4%, 0.4%, 召回率分别提升 0.5%, 2.7%。

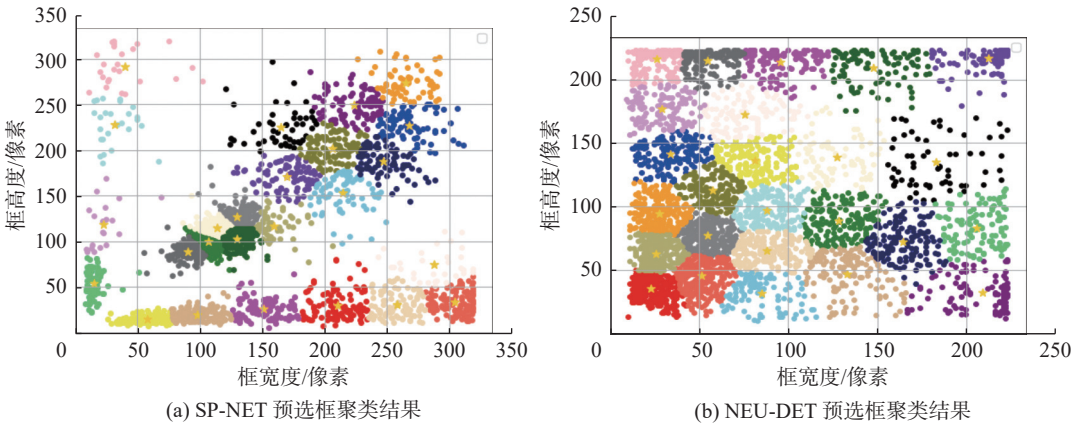


图 9 聚类结果

Fig.9 Results of clustering

4.4 对比分析

为了验证本文提出网络的性能，文章分别在两个数据集上与其他网络进行比较，训练过程中各网络采用的损失函数如表 5 所示。

图 10 为 SP-DDN 检测模型在 SP-NET 数据集训练过程中的回归损失 (BboxLoss) 和类别损失 (ClassLoss)，横轴为模型训练迭代次数，纵轴为损失值，在训练初始阶段模型较快收敛，最终两

者分别于 0.06，0.03 附近趋于稳定。表 6 为 SP-DDN 与其他目标检测模型在冲压件表面缺陷数据集上的性能比较。可见本文方法在凹陷类中的 AP 比基准网络 Faster R-CNN FPN 高出 9.4%，最终的 mAP 为 93.4%，模型的类别召回率为 98.9%。

表 7 列出了本文方法与其它目标检测方法在开源数据集 NEU-DET 上的检测结果。NEU-DET 数据集中缺陷种类更多，对于缺陷检测器的类型

表 5 网络损失函数示意图
Tab.5 Loss function for network

模型	分类损失	边界框回归损失	置信度损失
Faster R-CNN FPN	Cross Entropy Loss	L1 Loss	
Cascade R-CNN	Cross Entropy Loss	Smooth L1 Loss	
YOLO v3	Cross Entropy Loss	MSE Loss,Cross Entropy Loss	Cross Entropy Loss
Retina Net	Focal Loss	L1 Loss	
YOLO X ^[19]	Focal Loss	GIoU Loss,DIoU Loss	
EFPN ^[20]	Cross Entropy Loss	L1 Loss	
本文	Cross Entropy Loss	L1 Loss	

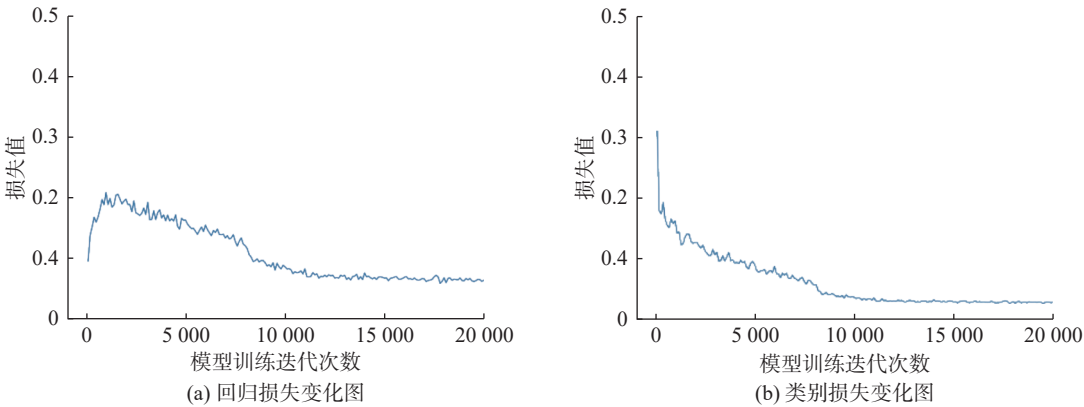


图 10 SP-NET 训练损失
Fig.10 SP-NET training loss

表 6 不同检测算法在自制 SP-NET 数据集的性能表现

Tab.6 Performance of different detection algorithms on a dataset of surface defects on stamped parts

检测算法	AP / %				mAP / %	R / %	FPS / (task·s ⁻¹)	Param / M
	凹陷	划痕	凸起	压痕				
Faster R-CNN	90.4	90.4	90.2	90.1	90.3	97.4	17.8	41.15
Cascade R-CNN	99.5	90.8	90.7	90.5	92.9	97.9	14.5	68.94
YOLO v3	87.2	74.5	90.6	90.0	95.3	95.3	23.3	3.67
Retina Net	90.9	88.1	90.8	90.6	90.1	97.7	19.0	35.62
YOLO X	90.8	90.2	99.8	89.8	92.7	99.35	19.1	8.94
EFPN	99.8	70.6	90.8	90.9	88.0	94.15	16.6	43.27
本文	99.8	91.7	91.2	90.8	93.4	98.9	16.7	41.15

表 7 不同检测算法在 NEU-DET 公开数据集上的表现

Tab.7 Performance of different detection algorithms on hot rolled steel surface defect dataset

检测算法	AP / %						mAP / %	R / %	FPS/(task·s ⁻¹)	Param / M
	裂纹	杂质	斑点	麻点	氧化	划痕				
Faster R-CNN FPN	34.0	74.8	77.1	86.6	64.6	89.6	71.1	82.0	18.1	41.15
Cascade R-CNN	34.9	66.8	85.1	87.7	62.7	89.5	71.2	80.9	14.5	68.94
YOLO v3	16.4	62.7	73.7	74.2	41.0	76.2	57.4	67.8	23.3	3.67
Retina Net	45.6	70.5	84.6	86.5	71.1	81.8	73.3	95.7	19.0	35.62
YOLO X	31.4	72.3	86.0	83.2	60.4	88.3	70.3	96.35	19.1	8.94
EFPN	31.6	44.5	74.3	85.0	61.0	61.4	59.6	72.85	16.4	43.27
本文	45.9	75.3	86.7	87.3	66.9	88.6	75.1	87.5	16.7	41.15

辨别能力有着较高要求。从表 7 中可以看出, 本文所提算法最终的 mAP 结果为 75.1%, 优于其他检测算法。

5 结 论

本文针对车身冲压件表面缺陷检测难题, 提出 SP-DDN 模型实现冲压件表面缺陷位置与类型识别。首先基于 ResNet50 作为模型的骨干网络以进行特征提取, 同时为满足缺陷多尺寸检测要求对特征金字塔网络进行扩展得到 CFPN 结构, 通过嵌入融合自注意力机制和可变卷积的 SCBAM 模块加强模型提升通道类间信息的利用率。本文提出方法分别在自制数据集以及开源数据集上进行验证。结果表明, 相比于基准网络 Faster R-CNN FPN 网络, 本文提出的 SP-DDN 在 mAP 和召回率上均有效提升。然而, 该模型计算量的增加导致了模型检测速度的下降, 后续将研究模型检测速度提升方法并减少模型参数量, 以适应实时性检测场景的要求。

参考文献:

[1] ARNAL L, SOLANES J E, MOLINA J, et al. Detecting dings and dents on specular car body surfaces based on optical flow[J]. *Journal of Manufacturing Systems*, 2017, 45: 306–321.

[2] MOLINA J, SOLANES J E, ARNAL L, et al. On the detection of defects on specular car body surfaces[J]. *Robotics and Computer-Integrated Manufacturing*, 2017, 48: 263–278.

[3] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, real-time object detection[C]//2016

IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 779–788.

[4] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//Proceedings of the 14th European Conference on Computer Vision. Amsterdam: Springer, 2016: 21–37.

[5] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 318–327.

[6] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 580–587.

[7] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE, 2015: 1440–1448.

[8] REN S Q, HE K M, 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.

[9] CAI Z W, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 6154–6162.

[10] CHENG X, YU J B. RetinaNet with difference channel attention and adaptively spatial feature fusion for steel surface defect detection[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 2503911.

[11] ZHENG Z H, ZHAO J, LI Y. Research on detecting bearing-cover defects based on improved YOLOv3[J]. *IEEE Access*, 2021, 9: 10304–10315.

[12] CHANG F, LIU M, DONG M Y, et al. A mobile vision inspection system for tiny defect detection on smooth car-body surfaces based on deep ensemble learning[J].

Measurement Science and Technology, 2019, 30(12): 125905.

[13] 朱文博, 夏林聪, 陈龙, 等. 基于改进 YOLOv5 的 O 型密封圈缺陷检测方法 [J]. 上海理工大学学报, 2022, 44(5): 440–448.

[14] HUANG Z H, XIAO H, ZHANG R Y, et al. Multi-scale feature pair based R-CNN method for defect detection[C]//2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). Atlanta: IEEE, 2019: 46–51.

[15] DING R W, DAI L H, LI G P, et al. TDD-net: A tiny defect detection network for printed circuit boards[J]. CAAI Transactions on Intelligence Technology, 2019, 4(2): 110–116.

[16] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the 15th European Conference on Computer Vision. Munich:

Springer, 2018: 3–19.

[17] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017: 764–773.

[18] HE Y, SONG K C, MENG Q G, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(4): 1493–1504.

[19] GE Z, LIU S T, WANG F, et al. YOLOx: exceeding YOLO series in 2021[DB/OL]. [2022-04-05]. <http://arxiv.org/abs/2107.08430>.

[20] DENG C F, WANG M M, LIU L, et al. Extended feature pyramid network for small object detection[J]. IEEE Transactions on Multimedia, 2022, 24: 1968–1979.

(编辑: 黄 娟)

(上接第 637 页)

[8] 蔡飞龙, 许思传, 常国峰. 纯电动汽车用锂离子电池热管理综述 [J]. 电源技术, 2012, 36(9): 1410–1413.

[9] HUO Y T, RAO Z H, LIU X J, et al. Investigation of power battery thermal management by using mini-channel cold plate[J]. Energy Conversion and Management, 2015, 89: 387–395.

[10] TANG A K, LI J M, LOU L S, et al. Optimization design and numerical study on water cooling structure for power lithium battery pack[J]. Applied Thermal Engineering, 2019, 159: 113760.

[11] SHANG Z Z, QI H Z, LIU X T, et al. Structural optimization of lithium-ion battery for improving thermal performance based on a liquid cooling system[J]. International Journal of Heat and Mass Transfer, 2019, 130: 33–41.

[12] NIETO N, DÍAZ L, GASTELURRUTIA J, et al. Novel thermal management system design methodology for power lithium-ion battery[J]. Journal of Power Sources, 2014, 272: 291–302.

[13] 杨莹莹, 魏学哲, 刘耀锋, 等. 车用锂离子电池交流加热研究 [J]. 汽车工程, 2016, 38(7): 901–908.

[14] SHENG L, SU L, ZHANG H Y, et al. An improved calorimetric method for characterizations of the specific heat and the heat generation rate in a prismatic lithium ion battery cell[J]. Energy Conversion and Management, 2019, 180: 724–732.

[15] SAW L H, SOMASUNDARAM K, YE Y, et al. Electro-thermal analysis of lithium iron phosphate battery for electric vehicles[J]. Journal of Power Sources, 2014, 249: 231–238.

[16] BERNARDI D, PAWLIKOWSKI E, NEWMAN J. A general energy balance for battery systems[J]. Journal of the Electrochemical Society, 1985, 132(1): 5–12.

[17] CHEN S C, WAN C C, WANG Y Y. Thermal analysis of lithium-ion batteries[J]. Journal of Power Sources, 2005, 140(1): 111–124.

[18] 杨锋. 电动汽车用锂离子电池液冷散热仿真分析 [D]. 南昌: 南昌大学, 2020.

(编辑: 黄 娟)