

基于改进 CenterNet 的竹条表面缺陷检测方法

高钦泉^{1,2,3}, 黄炳城^{1,2}, 刘文哲³, 童同^{3*}

(1. 福州大学 物理与信息工程学院, 福州 350116; 2. 福建省医疗器械与医药技术重点实验室(福州大学), 福州 350116;

3. 福建帝视信息科技有限公司, 福州 350001)

(* 通信作者电子邮箱 ttravelong@imperial-vision.com)

摘要:在竹条表面缺陷检测中,竹条表面缺陷形状各异,成像环境脏乱,现有基于卷积神经网络(CNN)的目标检测模型面对这样特定的数据时并不能很好地发挥神经网络的优势;而且竹条来源复杂且有其他条件限制,因此没办法采集所有类型的数据,导致竹条表面缺陷数据量少到 CNN 不能充分学习。针对这些问题,提出一种专门针对竹条表面缺陷的检测网络。该网络的基础框架为 CenterNet,而且为提高 CenterNet 在较少的竹条表面缺陷数据中的检测性能,设计了一种基于从零开始训练的辅助检测模块:在网络开始训练时,冻结采用预训练模型的 CenterNet 部分,并针对竹条的缺陷特点从零开始训练辅助检测模块;待辅助检测模块损失趋于稳定时,通过一种注意力机制的连接方式将该模块与采用预训练的主干部分进行融合。将所提检测网络与 CenterNet 以及目前常用于工业检测的 YOLO v3 在相同训练测试集上进行训练和测试。实验结果表明,所提检测网络的平均精度均值(mAP)在竹条表面缺陷检测数据集上比 YOLO v3 和 CenterNet 的 mAP 分别提高了 16.45 和 9.96 个百分点。所提方法能够针对形状各异的竹条表面缺陷进行有效检测,且没有增加过多的耗时,在实际工业运用中具有很好的效果。

关键词:目标检测;缺陷检测;注意力机制;卷积神经网络;深度学习;CenterNet

中图分类号:TP391 **文献标志码:**A

Bamboo strip surface defect detection method based on improved CenterNet

GAO Qinqun^{1,2,3}, HUANG Bingcheng^{1,2}, LIU Wenzhe³, TONG Tong^{3*}

(1. College of Physics and Information Engineering, Fuzhou University, Fuzhou Fujian 350116, China;

2. Key Laboratory of Medical Instrumentation & Pharmaceutical Technology of Fujian Province (Fuzhou University), Fuzhou Fujian 350116, China;

3. Imperial Vision Technology Company Limited, Fuzhou Fujian 350001, China)

Abstract: In bamboo strip surface defect detection, the bamboo strip defects have different shapes and messy imaging environment, and the existing target detection model based on Convolutional Neural Network (CNN) does not take advantage of the neural network when facing such specific data; moreover, the sources of bamboo strips are complicated and there exist other limited conditions, so that it is impossible to collect all types of data, resulting in a small amount of bamboo strip defect data that CNN cannot fully learn. To address these problems, a special detection network aiming at bamboo strip defects was proposed. The basic framework of the proposed network is CenterNet. In order to improve the detection performance of CenterNet in less bamboo strip defect data, an auxiliary detection module based on training from scratch was designed: when the network started training, the CenterNet part that uses the pre-training model was frozen, and the auxiliary detection module was trained from scratch according to the defect characteristics of the bamboo strips; when the loss of the auxiliary detection module stabilized, the module was integrated with the pre-trained main part by a connection method of attention mechanism. The proposed detection network was trained and tested on the same training sets with CenterNet and YOLO v3 which is currently commonly used in industrial detection. Experimental results show that on the bamboo strip defect detection dataset, the mean Average Precision (mAP) of the proposed method is 16.45 and 9.96 percentage points higher than those of YOLO v3 and CenterNet, respectively. The proposed method can effectively detect the different shaped defects of bamboo strips without increasing too much time consumption, and has a good effect in actual industrial applications.

Key words: object detection; defect detection; attention mechanism; Convolutional Neural Network (CNN); deep learning; CenterNet

0 引言

目标检测一直是计算机视觉中的经典问题,包含图像分类和目标定位两个任务,在人脸识别、自动驾驶、工业检测等

领域有着广泛的应用。随着深度学习的发展,许多基于卷积神经网络(Convolutional Neural Network, CNN)的目标检测性能得到进一步的提升。目前基于 CNN 的目标检测主要分为两种:1)二阶段目标检测,利用候选区域生成阶段生成一组候

收稿日期:2020-08-05;修回日期:2020-12-02;录用日期:2020-12-04。

作者简介:高钦泉(1986—),男,福建福清人,副研究员,博士,主要研究方向:人工智能、计算机视觉、医学图像处理与分析、计算机辅助手术导航;黄炳城(1993—),男,福建泉州人,硕士研究生,主要研究方向:人工智能、计算机视觉;刘文哲(1993—),男,福建三明人,硕士,主要研究方向:人工智能、计算机视觉;童同(1986—),男,安徽安庆人,研究员,博士,主要研究方向:人工智能、计算机视觉、医学图像处理与分析。

选目标检测框,然后利用 CNN 对候选区域的特征进行分类和位置回归,典型代表为区域卷积神经网络 (Region-Convolutional Neural Network, R-CNN) 系列^[1-3];2) 单阶段目标检测,它们取消了生成候选区域阶段,直接利用 CNN 来预测边界框,典型代表为 SSD (Single Shot multibox Detector) 系列^[4-5]、YOLO (You Only Look Once) 系列^[6-8]。

本文所依托的骨干网络 CenterNet^[9]将需要成对关键点的问题简化成中心关键点的估计问题,由于 CenterNet 只通过提取目标的中心点且不需要对中心点进行后处理,极大地加快了算法的整体速度。基于关键点目标检测中,大大减少了计算量,降低了基于 anchor-based 所导致的大量正负样本不平衡,减少了大量超参数设计。基于关键点目标检测以其简洁、高效率的优点将一阶段目标检测推向另一高度。一般来说,二阶段检测器具有比单阶段检测器精度更高的优势,而单阶段检测器比二阶段检测器在时间和效率更具优势。

目前,国内已有一些针对竹条表面缺陷检测的研究。贺峰等^[10]提出了一种基于小波多尺度分解的方法实现竹条六种缺陷的检测。该方法首先通过图像共生矩阵的逆差矩差分值得到的最佳小波分解层对竹条纹理进行滤除,然后利用 OTSU 自动阈值检测算法完成缺陷检测。张殿甫等^[11]提出了一种基于机器视觉的竹条表面缺陷检测方法。该方法先将对竹条采集到的彩色图像转为灰度图,之后进行中值滤波去除噪声,平滑后将灰度图转为二值图像;然后用 sobel 算子提取竹条边缘;最后对竹条边缘进行分段处理,结合最小二乘法对每一段进行直线拟合,依据拟合之后直线的斜率大小来判断是否存在缺陷。

近年来,随着深度学习的快速发展,很多传统的工业问题利用深度学习的方法得到了更好的解决。面对传统工业问题,一般是将由 CNN 构成的目标检测网络框架首先使用分类任务数据集(如 ImageNet^[12])进行预训练,使其拥有强大的分类能力,然后利用迁移学习的方法将模型针对对应目标工业数据集进行调优细化,从而实现在工业问题上的检测。但是,采用将一般 CNN 构成的目标检测框架直接迁移应用到工业检测问题,一般存在以下几个影响:

1) 直接采用迁移学习,可能导致模型泛化性能弱。目前现有基于 CNN 的目标检测网络框架主要针对的是一般场景的问题,如 COCO^[13]涵盖 80 类物体,而工业场景数据与这些数据存在较大的差别,所以可能导致预训练模型的特征提取能力在工业场景数据上表现较差。

2) 基于少数数据的模型训练,可能导致检测性能较差。比如竹条来源复杂,没办法采集所有类型的数据,所以竹条表面缺陷数据量少,导致 CNN 不能充分地学习。

3) 基于现实工业竹条表面缺陷检测应用中,其数据不适合用数据增广的方法来提高少量数据缺陷样本。比如采用旋转、翻转等操作,因为在工业应用中,数据的呈现形式是固定的,即不可能出现翻转、旋转后的那种数据形态,所以这些数据增广的方法对提高现实工业竹条表面缺陷检测应用中的性能表现影响不大。

基于以上问题,本文提出一种改进 CenterNet 的基于辅助检测的竹条表面缺陷检测方法,将基于关键点检测的 CenterNet 作为骨干网络,并通过一种注意力机制的连接的方法与辅助学习检测网络结合。实验采用平均精度均值(mean Average Precision, mAP)来评估方法,在竹条表面缺陷检测数据集上使用相同的训练集和测试集,本文方法比 YOLOv3 的 mAP 提高了 16.45 个百分点,比 CenterNet 的 mAP 提高了 9.06 个百分点。

综上所述,本文的工作如下:1) 使用基于深度学习的方法来进行竹条的缺陷检测。利用 CNN 的可学习性,能够同时检出竹黄、竹青、霉斑、虫洞等十种缺陷,减少了传统机器视觉所需的超参数,简化了检测过程,提高了检测效率。2) 提出了一个针对竹条表面缺陷特点的辅助检测网络。该网络模型小,训练速度快,能够得到较精确的目标定位,辅助骨干网络,提高检测性能。3) 使用一个注意力机制的连接方式将辅助检测网络和预训练的骨干网络结合,这一方法能更好地学习不同网络阶段产生的特征图,能够更合理地分配不同重要性程度的特征图的比重,让特征图利用得到最优化,减少冗余特征,提高特征处理速度,使目标检测性能得到优化。4) 针对竹条表面缺陷特点,采用梯度均衡损失函数用于计算对位置坐标回归。相比 L1 损失函数的方法,梯度均衡损失函数在计算位置坐标时在 L1 损失函数的基础上加入了梯度密度的影响,能达到更好的回归效果。

1 网络方案设计

1.1 本文方法分析

本文模型由三个部分组成:骨干网络、辅助检测网络和注意力机制连接。骨干网络采取经过 (ImageNet) 预训练的模型,具备强大的特征提取能力;而辅助检测网络则采取针对竹条表面缺陷数据从零训练。文献[14-15]中表明,当训练数据少时,使用基于预训练的模型能够快速收敛以及得到较好的检测精度;当训练数据足够时,使用预训练模型仅只能比从零开始训练的模型更快地收敛,并不能带来精度的提升;而当模型从零开始训练时,能够得到较精确的目标定位,但相比经过预训练的模型则需要花费更多的时间。本文以 CenterNet-dla34^[9,16]为基础网络,因为它具有高检测精度与高效率的特点。如图 1(a) 所示,将骨干网络最后几层得到的特征图通过注意力机制连接的方式分别与辅助检测网络所产生的特征图结合。为了能够得到更精确的目标信息,在辅助检测网络中采用可变形卷积 (Deformable Convolution, DCN)^[17]替代了普通卷积,因为在可变形卷积中,其卷积核能够根据目标形状的不同而做出相应的改变,能够得到更为精确的竹条表面缺陷信息。

表 1 为竹条表面缺陷数据集的每一类缺陷类型的样本数,表 2 为缺陷的长宽比和面积比详情。可以看出,竹条表面缺陷的样本数存在较大的不平衡,如霉斑、裂边、内碳等具有较多的样本数,而虫洞、竹黄、刀疤等缺陷的样本数则较少。因为所采用的竹条表面缺陷数据极不平衡,所以本文采用 Focal Loss^[18]结合梯度均衡损失函数 (Gradient Harmonizing Mechanism Loss, GHM Loss)^[19]来减小由数据不平衡带来的影响,因为 Focal Loss 及 GHM Loss 均能有效解决数据中的样本不平衡现象。

表 1 竹条表面缺陷数据详情

Tab. 1 Details of bamboo strip surface defects

缺陷	数量	缺陷	数量	缺陷	数量	缺陷	数量
黑节	4 115	虫洞	1 074	霉斑	12 942	竹黄	526
凹黄	6 224	裂边	11 990	竹青	3 643	刀疤	527
裂片	2 385	内碳	12 988				

1.1.1 骨干网络

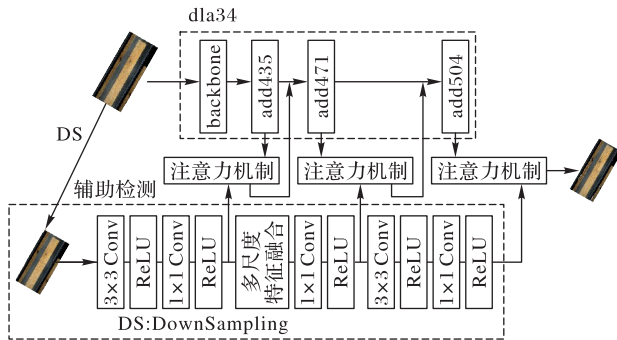
本文以 CenterNet 为基础网络,在 CenterNet 选取 DLA-34 为骨干分类网络,如图 2 所示为本文所采取的 CenterNet-dla34 网络结构。其中 DLA (Deep Layer Aggregation) 为一个带有多级跳跃连接的图像分类网络,由下采样模块、上采样特征融合模块和特征结合传递输出等三个部分组成,该网络使用跳跃

连接(skip connection)的方式将低层卷积层的输出与当前层的卷积层输出通过可变形卷积的方式进行上采样。

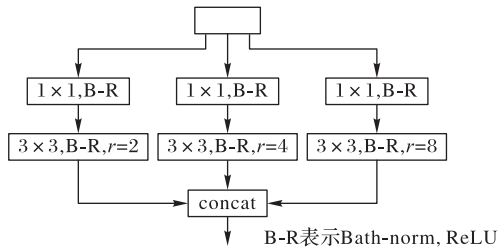
表 2 竹条表面缺陷长宽比与面积比详情

Tab. 2 Details of aspect ratio and area ratio of bamboo strip surface defects

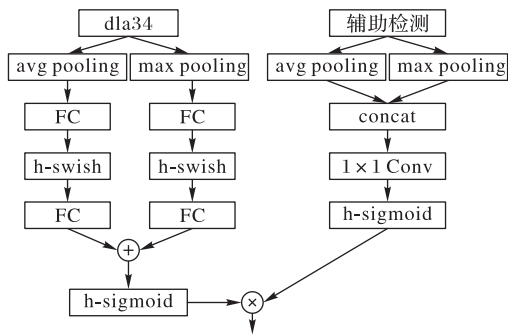
缺陷的长宽比		缺陷的面积比	
数据区间	数量	数据区间	数量
(0, 2.6]	12 830	(0.0, 0.025]	40 856
(2.6, 5.2]	9 530	(0.025, 0.05]	6 763
(5.2, 7.8]	7 328	(0.05, 0.075]	2 663
(7.8, 10.4]	14 707	(0.075, 0.1]	1 310
(10.4, 13.0]	3 114	(0.1, 0.125]	508
(13.0, 15.6]	2 511	(0.125, 0.15]	340
(15.6, 18.2]	1 879	(0.15, 0.175]	382
(18.2, 20.8]	1 573	(0.175, 0.2]	445
(20.8, 23.4]	1 510	(0.2, 0.225]	502
(23.4, 26.0]	1 432	(0.225, 0.25]	2 645



(a) 网络框架



(b) 多尺度特征融合



(c) 注意力机制

图 1 增强 CenterNet 网络框架

Fig. 1 Network framework of enhanced CenterNet

1.1.2 辅助学习检测网络

本文的辅助学习检测网络结构如图 1(a)所示,其中辅助检测网络结构由三个部分组成:

1)降采样特征提取表示,由一个自适应降采样层和卷积层表示。首先输入竹条图像经过一个降采样处理,然后再经过两个连续的卷积层,每个卷积层后都接一个非线性激活函数 ReLU 来增强网络的非线性表达能力。

2)多尺度特征融合,如图 1(b)所示,由多个卷积组成,为了得到多尺度的上下文信息,本文采用了空洞卷积^[20]对其进行处理。首先将降采样特征提取表示部分的输出作为多尺度特征融合的输入,接着由三个卷积核大小为 1×1 的卷积拆分成三个分支,然后采用三个卷积核大小为 3×3 且空洞卷积率 r 分别为 2、4、8 的卷积进行提取特征,接着对提取得到不同感受野的特征进行融合输出到下一模块。在本文方法中,使用空洞卷积有如下几个优点:第一,在不降低图像分辨率且不增加额外计算量的基础上,扩大感受野。经观察统计在竹条表面缺陷中,存在着缺陷目标大小不一且其长宽比等差距较大,如表 1 所示,竹条表面缺陷的长宽比大部分落在 (0, 10.4) 内,且竹条表面缺陷的面积比大部分为 (0, 0.025) 区间,说明本文所采用的竹条数据集中竹条的大部分缺陷为细小且狭长的缺陷。扩大感受野可以检测分割大缺陷目标,且可以精确定位目标。第二,使用不同的空洞卷积率的卷积,可获得不同感受野的特征,即可获得多尺度的缺陷特征。

3)使用卷积层对输出特征信息作进一步处理,使用多个卷积对多尺度特征融合模块的输出信息进行处理,使得到的图像特征更具有代表性。

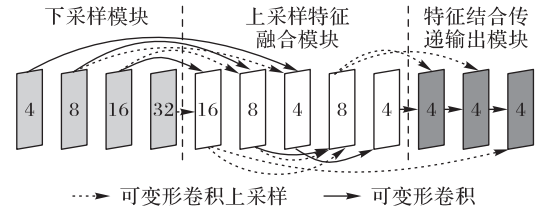


图 2 CenterNet-dla34 网络结构

Fig. 2 CenterNet-dla34 network structure

1.1.3 注意力机制连接

注意力机制连接模块参考了 SENet (Squeeze-and-Excitation Networks)^[21]与 CBAM (Convolutional Block Attention Module)^[22]。文献[23]中表明,h-swish(hard version of swish)和 h-sigmoid 激活函数能比 ReLU 和 Sigmoid 激活函数更有效地提高检测精度,h-swish 和 h-sigmoid 激活函数的计算量相对较小且能有效避免模型在量化计算时的精度损失。受此启发,本文在 CBAM 的基础上分别将 ReLU 和 Sigmoid 换成 h-swish 和 h-sigmoid,在 CBAM 中提到了通道注意力机制主要关注物体的类别信息,而空间注意力机制主要关注的是物体的定位信息,所以在本文中分别利用通道注意力机制和空间注意力机制用于连接骨干分类网络和辅助检测网络,然后将两者的特征进行融合。注意力机制连接模块可以通过学习的方式获得两个网络的特征重要性程度,优化提取的图像特征,避免了大量冗余特征,以获得更好的特征表达,提高网络的计算速度。

1.2 损失函数

Focal Loss 能够解决检测任务中由于大量难易样本数量不平衡的问题,因此被广泛采用,本文也采用 Focal Loss 来计算中心点的损失。假设输入图像 $I \in \mathbb{R}^{W \times H \times 3}$,其中 W 和 H 表示输入图像的宽和高。在预测时,要产生出关键点的热力图(keypoint heatmap): $\hat{Y} \in [0, 1]^{\frac{W}{R} \times \frac{H}{R} \times C}$,其中: R 为输出对应原图的步长,本文中取 $R=4$; C 是在竹条表面缺陷检测中对应检测点的数量,本文中取 $C=10$,表示有 10 个缺陷类别。

$$L_k = \frac{1}{N} \begin{cases} \sum_{ayc} (1 - \hat{Y}_{ayc})^\alpha \log(\hat{Y}_{ayc}), & Y_{ayc} = 1 \\ \sum_{ayc} (1 - Y_{ayc})^\beta (\hat{Y}_{ayc})^\alpha \log(1 - \hat{Y}_{ayc}), & \text{其他} \end{cases} \quad (1)$$

其中: α 和 β 为 Focal Loss 的超参数, 本文中 α 和 β 分别为 2 和 4; N 为图像 I 的关键点数量。假设 p 为图像 I 的中心点, 且 $\tilde{p}$ 为 p/R 的向下取整, 用一个高斯核为:

$$Y_{\text{xyz}} = \exp\left(-\frac{(x - \tilde{p}_x)^2 + (y - \tilde{p}_y)^2}{2\delta_p^2}\right) \quad (2)$$

将关键点分布到特征图上, 其中 δ_p^2 为一个与目标大小相关的标准差。由于对图像进行下采样(下采样因子 $R = 4$), 这样得到的特征图再映射到原始图像上会产生精度误差, 所以针对每个中心点增加一个偏置值, 并采用 L1 Loss 去训练。

$$L_{\text{off}} = \frac{1}{N} \sum_p \left| \hat{O}_{\tilde{p}} - \left(\frac{p}{R} - \tilde{p} \right) \right| \quad (3)$$

其中 $\hat{O} \in \mathbb{R}^{\frac{W}{R} \times \frac{H}{R} \times 2}$ 为一个偏置值。假设 $(x_1^{(k)}, y_1^{(k)}, x_2^{(k)}, y_2^{(k)})$ 为目标 k , 类别为 c 的点, 其中心点为 $p_k = \left(\frac{x_1^{(k)} + x_2^{(k)}}{2}, \frac{y_1^{(k)} + y_2^{(k)}}{2} \right)$, 使用关键点预测去预测所有的中心点, 然后对每个目标 k 的大小进行回归, 回归后得到预测中心点为: $s_k = \left(\frac{x_1^{(k)*} + x_2^{(k)*}}{2}, \frac{y_1^{(k)*} + y_2^{(k)*}}{2} \right)$, 并利用 GHM Loss

来训练 $L_{\text{size}} = \frac{1}{N} \sum_{k=1} \frac{\left(\sqrt{d_i^2 + \mu^2} - \mu \right)}{GD(gr_i)}$, 其中 $d = s_k - p_k$, $gr = \left| \frac{d}{\sqrt{d^2 + \mu^2}} \right|$, 取 μ 为 0.02, 整个物体的损失函数为:

$$L = L_k + \lambda_{\text{off}} L_{\text{off}} + \lambda_{\text{size}} L_{\text{size}} \quad (4)$$

本文参考 CenterNet 参数设置取: $\lambda_{\text{off}} = 1, \lambda_{\text{size}} = 0.1$ 。

2 实验与分析

2.1 实验数据准备和实验设置

本文使用的竹条表面缺陷数据集包含竹青、竹黄、黑节、裂片、裂边、内碳、霉斑、刀疤、虫洞、凹黄等十种竹条表面缺陷, 如图 3 所示。本文采用的竹条表面缺陷数据图片共 37 040 张, 对其中 36 747 张图片采用 LabelImg 进行标注, 而且在标注图片中随机选取 28 747 张图片进行训练, 并按照 8:2 的比例随机选取 22 997 张图片为训练集, 5 750 张图片为验证集, 剩下的 8 000 张标注缺陷数据图与 293 张未标注无缺陷图组成为测试集, 即测试集图片共 8 293 张。

本文使用了相同的设置来初始化和训练模型, 采用 CenterNet-dla34 为本文的骨干网络, 该骨干网络首先在 ImageNet 数据集上预训练, 在实验中, 使用 512×512 的大小输入到骨干网络, 批量大小(batch size)设置为 8, 学习率设定为 0.000 125, 在网络训练过程中首先冻结带有预训练的骨干网络, 只训练本文提出的辅助学习检测网络, 待辅助学习检测网络损失趋于稳定时, 放开冻结, 两者结合一起训练。共训练 140 epoch, 且当在训练到 90 和 120 epoch 的时候, 学习率分别下降到当前值的 1/10, 在训练期间, 所有实验均使用 Adam^[24] 进行优化, 并使用期默认的参数 $\beta_1 = 0.9, \beta_2 = 0.999$ 。本文检测采用的实验环境为 pytorch 框架, 在 Ubuntu16.04 系统下运行, 实验均在 NVIDIA GTX2070 GPU 上进行训练和测试。

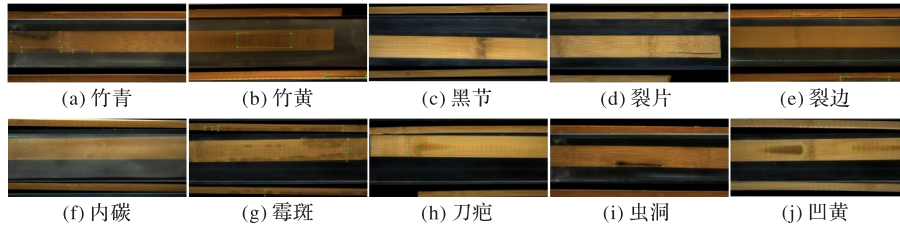


图 3 竹条表面缺陷的类型

Fig. 3 Types of bamboo strip defects

2.2 实验结果与分析

2.2.1 在竹条表面缺陷检测数据集上的结果

1) 数据均衡与数据增广在竹条数据上的表现。

使用竹条表面缺陷数据集来验证本文方法的效果, 并称利用本文方法改进后的模型为增强 CenterNet。表 3 展示了增强 CenterNet 与 YOLOv3、CenterNet 以及针对本文竹条数据采用数据均衡的方法等在相同数据集、相同测试环境下的 mAP 结果对比。其中 CenterNet+均衡数据是指: 将竹条数据中的少量缺陷样本如竹黄、刀疤、虫洞等数据通过平移、旋转、翻转等方法将其增广到与霉斑、裂边、内碳等均衡的数据量并采用 CenterNet 进行学习检测; CenterNet+数据增广则是针对竹条表面缺陷数据中的所有数据进行数据增广到更大的数据量, 采用 CenterNet 进行学习检测。由表 3 可以看出, 采用数据均衡与数据增广的方法在竹条表面缺陷检测上得到的 mAP 分别为 67.47% 和 68.65%, 这与原 CenterNet 方法得到的 66.94% 的 mAP 值相差不大, 因为在竹条表面缺陷中, 其缺陷特征具有固定的表现形式, 与其经过翻转、旋转等得到的数据形式具有一定的差别, 而本文采用针对竹条的缺陷特点通过辅助学习检测网络与骨干网络的结合, 使各类缺陷的检测效

果都有较大的提升。

2) 增强 CenterNet 在竹条数据上的表现。

由表 2 可以看出, 本文使用的竹条表面缺陷数据集的缺陷长宽比、缺陷面积比、缺陷类别样本数存在较大的不平衡, 竹条的缺陷长宽比大部分落在 (0, 10.4), 且竹条的缺陷面积比大部分为 (0, 0.025), 说明本文使用的竹条表面缺陷数据集形状大部分为细小且狭长。由图 4 检测结果可以看出, 在针对狭长的竹黄、内碳、裂边等缺陷的时候, 增强 CenterNet 能够比 YOLOv3 和 CenterNet 进行更有效、更精准的检测。

2.2.2 验证本文模块的有效性

1) 基础网络与本文方法的比较。

为了验证本文方法的有效性, 针对辅助检测网络、多尺度融合模块、注意力机制连接等做了如下实验。在实验过程中, 采用相同的设置来初始化和训练模型。如表 4 所示, 在竹条表面缺陷数据集上用基础网络 CenterNet 得到了 66.94% 的 mAP; 在 CenterNet 的基础上引入辅助学习检测网络, mAP 提高了 5.36 个百分点; 将骨干网络和辅助学习检测网络进行直接结合时, 虽然辅助学习检测网络能帮助模型较精确地定位缺陷位置, 但因为缺乏合理的特征融合方法, 在模型学习得到

的特征中存在大量冗余特征,所以引入了注意力机制的连接方法,这使模型的 mAP 提升了 1.9 个百分点。由表 1 中可知,在竹条表面缺陷数据集中,缺陷的有效样本分布极不均匀,针对辅助学习检测网络中的竹条表面缺陷特点,引入了 GHM

Loss,利用特征梯度分布的方法来引导缺陷定位,使本文方法在竹条表面缺陷数据集上获得了 76.9% 的 mAP。如图 5 所示为基于 CenterNet 上改进方法的损失函数收敛曲线,可看出本文方法在改进 CenterNet 上能够有更优的表现。

表 3 本文方法与其他方法的 AP 与 mAP 比较

Tab. 3 Comparison of AP and mAP between the proposed method and other methods

方法	竹条表面缺陷的 AP/%										mAP/%
	凹黄	黑节	虫洞	裂边	内碳	裂片	竹青	霉斑	竹黄	刀疤	
YOLOv3	85.71	74.44	68.77	63.37	60.39	56.64	55.38	49.42	45.55	44.86	60.45
CenterNet	89.18	79.15	73.23	69.95	80.35	58.12	64.89	52.91	52.21	49.41	66.94
CenterNet+均衡数据	89.17	77.24	75.46	68.46	79.59	59.39	63.42	51.79	57.31	52.82	67.47
CenterNet+数据增广	92.31	80.23	74.93	70.15	81.47	59.88	65.32	52.41	57.55	52.28	68.65
增强 CenterNet	92.58	88.72	84.32	81.20	90.05	70.49	68.87	65.05	70.98	56.65	76.90

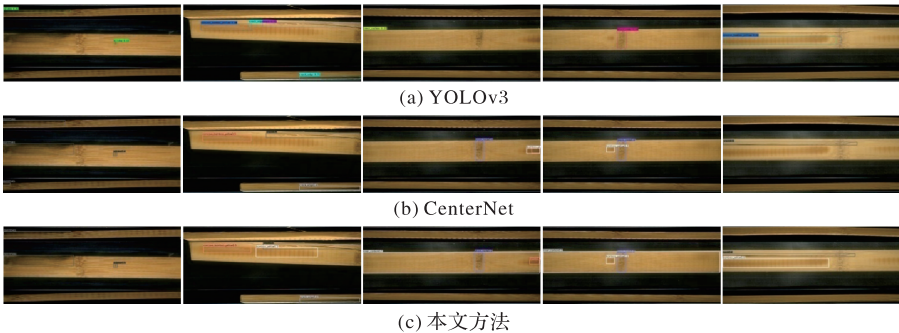


图 4 本文方法与其他方法检测效果比较

Fig. 4 Comparison of detection effect of the proposed method and other methods

表 4 基础网络添加不同模块的 mAP 结果

Tab. 4 mAP results of adding different modules to base network

方法	mAP/%
CenterNet	66.94
CenterNet+辅助检测网络	72.30
CenterNet+辅助检测网络+注意力机制	74.20
CenterNet+辅助检测网络+注意力机制+GHM Loss	76.90

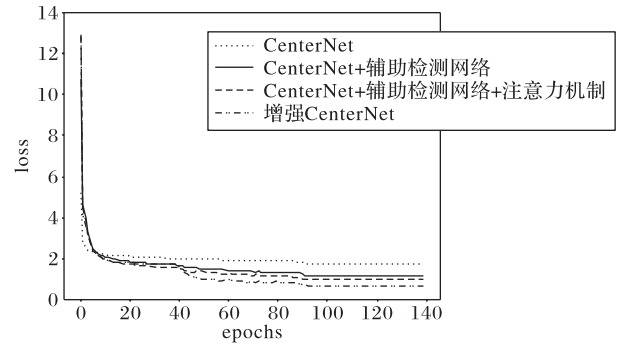


图 5 基础网络添加不同模块的损失曲线

Fig. 5 Loss curves of adding different modules to base network

2)不同注意力机制的连接表示。

在表 5 中,针对注意力机制连接方法采取了五种不同的方法,其中 CT 表示通道注意力,CT⁺表示在通道注意力中采用 h-swish 和 h-sigmoid 激活函数,SPT 表示空间注意力,CBAM⁺表示在 CBAM 方法的基础上将激活函数更改为 h-swish 和 h-sigmoid。采用通道注意力机制的方式作为连接骨干网络与辅助学习检测网络的方法能够比骨干网络和辅助学习检测网络直接结合情况的 mAP 提高 0.5 个百分点,说明引入注意力机制的方式能在一定程度上提高模型的学习能力,但仍需更

有效的方式。当在通道注意力中的 ReLU 和 sigmoid 激活函数替换为 h-swish 和 h-sigmoid 时,获得了 73.7% 的 mAP,说明在针对竹条表面缺陷形状中引入使用分段线性激活函数的通道注意力机制连接的方法能让模型更加准确地关注缺陷特征的学习,有效地提高了检测的精度。但在使用空间注意力机制进行连接时,检测精度比不使用注意力机制连接的 mAP 还下降了 0.7 个百分点。在使用 CBAM 作为骨干网络与辅助学习检测网络的连接方式时,其检测精度比不采用注意力机制连接的 mAP 提高了 1.9 个百分点,且当将 CBAM 的激活函数替换成 h-swish 和 h-sigmoid 时,其检测获得了最好的效果。

表 5 不同注意力机制连接的 mAP 结果

Tab. 5 mAP results of integrating with different attention mechanisms

方法	mAP/%
CenterNet + 辅助检测网络	72.3
CenterNet + 辅助检测网络 + CT	72.8
CenterNet + 辅助检测网络 + CT ⁺	73.7
CenterNet + 辅助检测网络 + SPT	71.6
CenterNet + 辅助检测网络 + CBAM	74.2
CenterNet + 辅助检测网络 + CBAM ⁺	76.9

3)不同空洞卷积率的表示。

尝试三种不同空洞卷积率的空洞卷积进行辅助检测网络中缺陷特征的提取,结果如表 6 所示。在不采取空洞卷积进行特征提取时,mAP 为 69.9%;当采用三个分支分别为 2、4、8 的空洞卷积率进行结合时,获得的 mAP 为 76.9%,模型性能有了一个较好的提升,说明在竹条表面缺陷数据集中采用空洞卷积,通过增大感受野对模型获取上下文的多尺度缺陷特征具有重大意义。

表 6 不同空洞卷积率的 mAP 结果
Tab. 6 mAP results of different dilation rates

$r=2$	$r=4$	$r=8$	mAP/%
✓			69.9
✓	✓		73.5
✓	✓		73.4
✓	✓	✓	76.9

3 结语

本文针对竹条表面缺陷特点利用深度学习的方法来进行竹条表面缺陷检测,能够在较少缺陷数据的情况下获得较好的检测结果。且针对目前现有网络模型的单一训练模式,本文提出一种结合预训练网络迁移学习和从零训练的辅助学习检测网络混合训练的方法,能够有效地提高在仅拥有少量竹条表面缺陷数据情况下的检测性能。在本文所提取的骨干分类网络特征层上具有尺度单一性,下一步研究工作将探寻提取特征更丰富的、结构更小、效率更高的模型,以便能得到更好的检测效果。

参考文献 (References)

- [1] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 580-587.
- [2] GIRSHICK R. Fast R-CNN [C]// Proceedings of the 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2015: 1440-1448.
- [3] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [C]// Proceedings of the 28th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2015: 91-99.
- [4] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]// Proceedings of the 2016 European Conference on Computer Vision, LNCS 9905. Cham: Springer, 2016: 21-37.
- [5] FU C Y, LIU W, RANGA A, et al. DSSD: deconvolutional single shot detector [EB/OL]. [2020-08-03]. <https://arxiv.org/pdf/1701.06659.pdf>.
- [6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 779-788.
- [7] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 7263-7271.
- [8] REDMON J, FARHADI A. YOLOv3: an incremental improvement [EB/OL]. [2020-08-03]. <https://arxiv.org/pdf/1804.02767.pdf>.
- [9] ZHOU X, WANG D, KRÄHENBÜHL P. Objects as points [EB/OL]. [2020-08-03]. <https://arxiv.org/pdf/1904.07850.pdf>.
- [10] 贺峰,秦现生,刘琼,等. 基于小波多尺度分解的竹条表面缺陷检测算法[J]. 机电一体化, 2010, 16(1): 46-49. (HE F, QIN X S, LIU Q, et al. Bamboo defect inspection algorithm based on wavelet multi-scale transform [J]. Mechatronics, 2010, 16(1): 46-49.)
- [11] 滁州市天达汽车部件有限公司. 一种基于机器视觉的竹条表面缺陷检测方法: 中国, 201710416704.0 [P]. 2017-10-17. (Chuzhou Tianda Auto Parts Company Limited. A bamboo strip defect detection method based on machine vision: China, 201710416704.0 [P]. 2017-10-17.)
- [12] DENG J, DONG W, SOCHER R, et al. ImageNet: a large-scale hierarchical image database [C]// Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2009: 248-255.
- [13] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context [C]// Proceedings of the 2014 European Conference on Computer Vision. LNCS 8693. Cham: Springer, 2014: 740-755.
- [14] HE K, GIRSHICK R, DOLLÁR P. Rethinking ImageNet pre-training [C]// Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 4917-4926.
- [15] ZHU R, ZHANG S, WANG X, et al. ScratchDet: training single-shot object detectors from scratch [C]// Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 2263-2272.
- [16] YU F, WANG D, SHELHAMER E, et al. Deep layer aggregation [C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 2403-2412.
- [17] ZHU X, HU H, LIN S, et al. Deformable ConvNets V2: more deformable, better results [C]// Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 9300-9308.
- [18] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]// Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 2999-3007.
- [19] LI B, LIU Y, WANG X. Gradient harmonized single-stage detector [C]// Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2019: 8577-8584.
- [20] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [EB/OL]. [2020-09-22]. <https://arxiv.org/pdf/1511.07122.pdf>.
- [21] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132-7141.
- [22] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]// Proceedings of the 2018 European Conference on Computer Vision. LNCS 11211. Cham: Springer, 2018: 3-19.
- [23] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]// Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 1314-1324.
- [24] KINGMA D P, BA J L. Adam: a method for stochastic optimization [EB/OL]. [2020-09-22]. <https://arxiv.org/pdf/1412.6980.pdf>.

GAO Qinquan, born in 1986, Ph. D., associate research fellow. His research interests include artificial intelligence, computer vision, medical image processing and analysis, computer-aided surgical guidance.

HUANG Bingcheng, born in 1993, M. S. candidate. His research interests include artificial intelligence, computer vision.

LIU Wenzhe, born in 1993, M. S. His research interests include artificial intelligence, computer vision.

TONG Tong, born in 1986, Ph. D., research fellow. His research interests include artificial intelligence, computer vision, medical image processing and analysis.