

多尺度特征自适应融合的轻量化织物瑕疵检测

杨毅, 桑庆兵

(江南大学 人工智能与计算机学院, 江苏 无锡 214122)

摘要: 织物瑕疵检测是纺织行业保证产品质量的重要环节, 针对织物瑕疵检测中存在小目标瑕疵检测困难、不同种类瑕疵长宽比差异大、对实时性要求高等问题, 提出一种新的轻量化织物瑕疵检测算法。以YOLOv4网络为基础, 使用轻量化网络MobileNetv2为主干网络, 有效减少模型参数总量与运算量, 以满足实时性需求。在MobileNetv2的逆残差结构中加入CoordAttention注意力模块, 将空间精确位置信息嵌入到通道注意力中, 增强网络聚焦小目标特征的能力。使用自适应空间特征融合(ASFF)网络改进路径聚合网络(PANet), 使模型通过学习获得多尺度特征图的融合权重, 从而充分利用浅层特征与深层特征, 提高算法对小目标瑕疵的检测精度。采用K-means++算法确定先验框尺寸, 并用Focal Loss函数修改模型损失函数, 降低正、负样本不平衡对检测结果的影响, 解决不同种类瑕疵长宽比差异大及类别不平衡的问题。实验结果表明, 相较于YOLOv4算法, 所提算法的平均精度均值提高了2.3个百分点, 检测速度提升了12 frame/s, 能较好地应用于织物瑕疵检测。

关键词: 织物瑕疵检测; 自适应空间特征融合; CoordAttention模块; YOLOv4网络; MobileNetv2网络

开放科学(资源服务)标志码(OSID):



中文引用格式: 杨毅, 桑庆兵. 多尺度特征自适应融合的轻量化织物瑕疵检测[J]. 计算机工程, 2022, 48(12): 288-295.

英文引用格式: YANG Y, SANG Q B. Lightweight-fabric defect detection based on adaptive fusion of multiscale features[J]. Computer Engineering, 2022, 48(12): 288-295.

Lightweight-Fabric Defect Detection Based on Adaptive Fusion of Multiscale Features

YANG Yi, SANG Qingbing

(School of Artificial Intelligence and Computer, Jiangnan University, Wuxi, Jiangsu 214122, China)

[Abstract] Fabric defect detection is an essential process in the textile industry for ensuring product quality. A algorithm for detecting lightweight fabric defects based on multiscale feature adaptive fusion is proposed to minimize small-target defect detection difficulties in fabric defect detection and the significant differences in the aspect ratios of different defects and high requirements for real-time performance. Based on the YOLOv4 network, lightweight network MobileNetv2 is used as the backbone network to effectively reduce the total number of model parameters and the cost of calculations to satisfy real-time requirements. In the inverse residual structure of MobileNetv2, a new attention mechanism is added. CoordAttention module, which embeds the spatially accurate position information necessary for detecting small-target defects into the channel attention, is used to enhance the ability of the network to focus on small-target features. Second, the Adaptive Spatial Feature Fusion (ASFF) network is used to improve Path Aggregation Network (PANet) to enable the model to obtain the fusion weights of multiscale feature maps through learning, fully utilize shallow and deep features, and further improve the detection accuracy of small-target defects. For different defect types and problems with significant aspect ratio differences and unbalanced categories, the K-means++ algorithm is used to determine the prior frame size, and the Focal Loss function is used to modify the model loss function to reduce the impact of the positive and negative sample imbalance on the detection result. The experimental results show that compared with YOLOv4, the mean Average Precision (mAP) of the proposed algorithm increases by 2.3 percentage points, and the detection speed increases by 12 frame/s. The proposed approach can be effectively applied to fabric defect detection.

[Key words] fabric defect detection; Adaptively Spatial Feature Fusion (ASFF); CoordAttention module; YOLOv4 network; MobileNetv2 network

DOI: 10.19678/j.issn.1000-3428.0063507

基金项目: 国家自然科学基金面上项目(52172324); 陕西省交通厅重点项目(20-38T); 西安市未央区科技计划项目(202121); 长安大学实验教学改革研究项目(20211811)。

作者简介: 杨毅(1996—), 男, 硕士研究生, 主研方向为计算机视觉; 桑庆兵, 副教授。

收稿日期: 2021-12-13 **修回日期:** 2022-02-13 **E-mail:** lmwen@chd.edu.cn

0 概述

织物质量检测一直以来都是纺织产业生产加工中的重要环节,织物瑕疵的出现会直接影响企业的利润。随着企业生产效率不断提升,传统人工检测的方法已不能适应大批量自动化生产检测的要求^[1]。因此,采用自动疵点检测技术已成为提高织物质量、降低人工成本的必然趋势^[2],但是机器自动检测织物瑕疵的效率和准确性仍面临较大挑战。

传统织物瑕疵检测方法主要分为结构法、统计法、频谱法、模型法、基于学习的方法等^[3]。MIRMAHDAVI等^[4]将无瑕疵样本输入高斯混合模型训练,通过实验设定欧氏距离阈值进行瑕疵检测。文献[5]利用织物纹理的周期性,通过匹配缺陷位置的相对偏移量和傅里叶位移定理曲线,分割出织物瑕疵部分。然而,传统算法存在缺陷分割性能差、噪声敏感、漏检率高等问题。随着计算机算力的快速提升,深度学习被广泛应用于织物视觉计算与处理领域。LIU等^[6]提出一种基于CGAN^[7]的多判别器模型,研究潜在匹配规则和服装属性之间关系,ZHAO等^[8]开发了一个基于双向长短期记忆Bi-LTSM^[9]模型的端到端服装搭配学习框架,ZHANG等^[10]提出使用GAN^[11]从人体穿着服装图像生成具有真实纹理的平铺服装图像。在织物瑕疵检测任务中,深度学习方法具有较强的特征表征能力。文献[12]采用深度学习方法,通过改进ResNet50^[13]网络提取特征,使用特征金字塔网络过滤冗余特征,实现对织物瑕疵的检测。以上检测方法解决了部分织物瑕疵检测问题,但目前织物检测存在缺陷种类多,尤其是小目标缺陷较多,对实时性要求高的问题,给深度学习方法的应用带来一定困难。

为满足织物瑕疵数据检测任务实时性需求,本文将YOLOv4^[14]网络的主干网络替换为MobileNetv2^[15],

并使用CoordAttention^[16]模块对MobileNetv2的逆残差结构进行改进,增强其对小目标瑕疵的特征提取能力。采用自适应特征融合(Adaptively Spatial Feature Fusion, ASFF)^[17]网络改进路径聚合网络(Path Aggregation Network, PANet)^[18]结构,提高小目标缺陷检测精度。针对不同种类瑕疵长宽比差异大、原始先验框尺寸不适用的问题,使用K-means++^[19]算法进行瑕疵尺寸聚类,确定锚框尺寸,并采用Focal Loss^[20]函数降低正、负样本不平衡对检测效果的影响。

1 YOLOv4与MobileNetv2结构

目标检测算法根据有无区域建议主要分为两类:一类是无区域建议算法,通过一个卷积神经网络直接对图像各个位置上的候选框进行预测和分类,例如YOLOv4^[14]与SSD^[21];另一类是基于区域建议的算法,由区域提议和检测2个步骤组成,相较于无区域建议算法,其精度较高但速度慢,代表网络有Faster R-CNN^[22]等。

YOLOv4是无区域建议目标检测算法的代表之一,其结构如图1所示。其主干网络CSPDarknet-53在YOLOv3^[23]网络的Darknet53基础上引进了CSPNet^[24]模块结构,相较于Darknet53网络,模型预测的准确率和推理速度均有所提高。颈部网络包括空间金字塔池化(Spatial Pyramid Pooling, SPP)^[25]和PANet两个部分,SPP通过使用不同大小的滑动核,有效增加了网络感受野,并将局部与全局的特征进行融合,增加了特征的丰富性。PANet相较于FPN^[26]网络增加了自下而上的增强路径,通过保留空间信息以增强实例分割过程。头部网络采用YOLOv3网络检测头,分别提取主干特征网络最后3层的图像特征,生成边界框和类别预测。

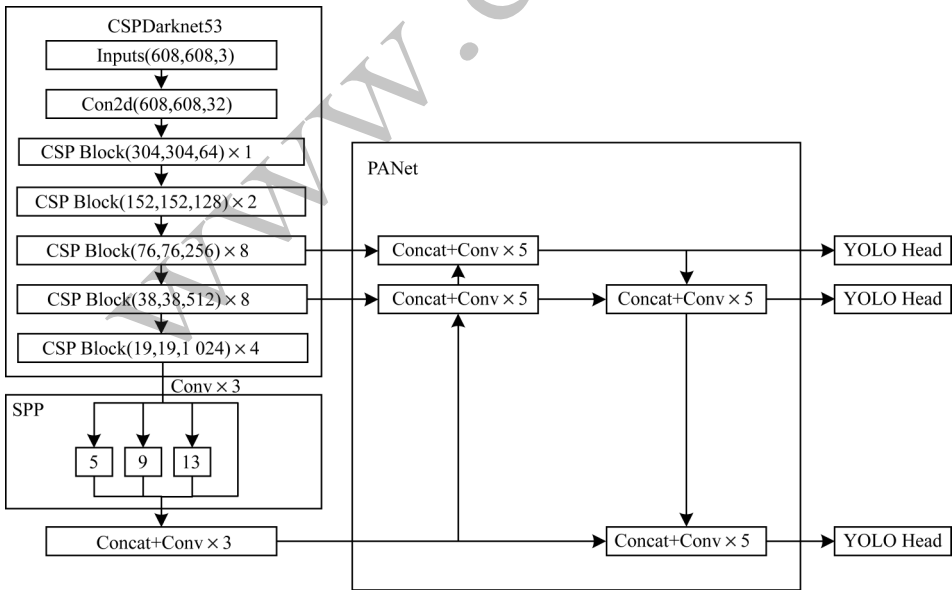


图1 YOLOv4网络结构

Fig.1 Structure of YOLOv4 network

YOLOv4 原始主干网络 CSPDarknet-53 虽然能有效提取深度特征信息,但结构复杂且参数数量多,导致难以应用于计算资源较差、实时性要求高的织物瑕疵检测任务中。

MobileNetv2 是一种轻量化的特征提取网络,其核心为深度可分离卷积和具有线性瓶颈的逆残差结构,具体如下:

1) 深度可分离卷积。相较于普通卷积,深度可分离卷积能有效减少模型的参数数量和计算量,其流程如图 2 所示。

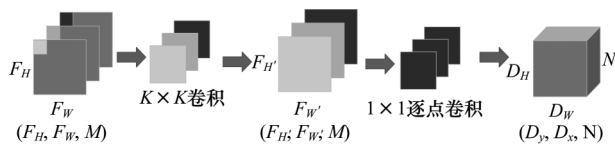


图2 深度可分离卷积流程

Fig.2 Procedure of depth separable convolution

对于输入为 $F_H \times F_W \times M$ 的特征图, F_H 、 F_W 分别为特征图的高和宽, M 为通道数,对特征图进行卷积操作,卷积核大小为 $K \times K$,卷积核数量为 N ,输出特征图大小为 $D_H \times D_W \times N$,在对特征图进行标准卷积的过程中,计算量 C_1 和参数量 P_1 的表达式如下所示:

$$C_1 = K \times K \times M \times N \times F_H \times F_W \quad (1)$$

$$P_1 = K \times K \times M \times N \quad (2)$$

对特征图进行深度可分离卷积运算,计算量 C_2 和参数量 P_2 的表达式分别如下所示:

$$C_2 = K \times K \times M \times F_H \times F_W + M \times N \times F_H \times F_W \quad (3)$$

$$P_2 = K \times K \times M \times N + K \times K \times N \quad (4)$$

深度可分离卷积与标准卷积计算量之比如式(5)所示:

$$\frac{C_2}{C_1} = \frac{1}{N} + \frac{1}{K^2} \quad (5)$$

若将深度可分离卷积的卷积核大小设为 3×3 ,步长设置为 1,按式(5)进行计算,其计算量下降为普通卷积的 $1/9$,模型检测速度得到大幅提高。

2) 具有线性瓶颈的逆残差结构。普通残差块先通过 1×1 卷积将特征图通道数降维,经过卷积提取特征后再扩张通道数。逆残差结构与其相反,由于输入通道数的增加有利于深度可分离卷积提取特征,所以先对特征图通道数进行升维,以提取更多特征,再压缩通道数。同时,为了消除 Relu 函数对被压缩特征的影响,在 1×1 卷积降维后使用 Linear 函数代替 Relu 函数。

表 1 所示为 MobileNetv2 网络的结构,可以看到 MobileNetv2 网络由多个具有线性瓶颈的逆残差结

构组成。

表 1 MobileNetv2 网络结构

Table 1 Structure of MobileNetv2 network

输入尺寸/ 像素	模块	升维 倍数	输出通道 数	重复 次数	卷积 步长
608×608	Conv2d	—	32	1	2
304×304	bottleneck	1	16	1	1
304×304	bottleneck	6	24	2	2
152×152	bottleneck	6	32	3	2
76×76	bottleneck	6	64	4	2
38×38	bottleneck	6	96	3	1
38×38	bottleneck	6	160	3	2
19×19	bottleneck	6	320	1	1
19×19	Conv2d	—	1 280	1	1

2 模型结构与优化

2.1 MobileNetv2 逆残差结构优化

在轻量级网络上的研究表明,通道注意力会给模型带来较显著的性能提升。为提高模型对瑕疵特征提取能力,本文在 MobileNetv2 逆残差块中加入新型注意力机制 CoordAttention 模块,如图 3 所示,通过将位置信息嵌入到通道注意力中,使网络能够更好地获取空间方向的特征信息,使小目标检测定位更加精确,同时也能避免产生大量的计算开销。

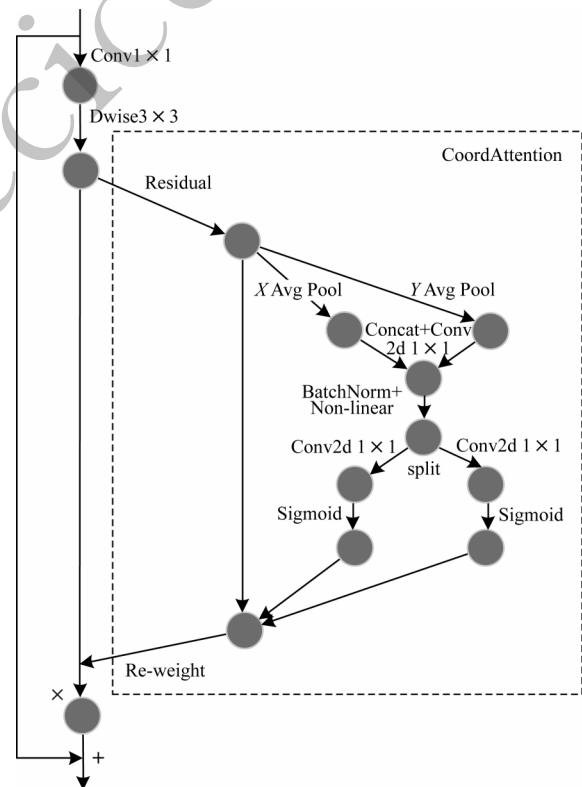


图3 加入 CoordAttention 模块的逆残差结构

Fig.3 Inverse residual structure with CoordAttention module

CoordAttention 的运算过程由坐标信息嵌入和坐标注意力生成两步组成, 具体过程如下所示:

1) 坐标信息嵌入。相对于全局池化压缩全局空间信息, 导致位置信息难以保存, CoordAttention 模块使用一对一特征编码操作替换全局池化, 并分别沿 2 个空间方向聚合特征, 得到一对具有空间信息的特征图, 从而有利于网络更准确地定位感兴趣的目标。

2) 坐标注意力生成。将生成的特征图进行拼接, 然后使用 1×1 卷积生成具有水平和垂直方向空间信息的特征图, 如式(6)所示:

$$f = \delta(F([\mathbf{z}^h, \mathbf{z}^w])) \quad (6)$$

其中: $[\cdot, \cdot]$ 为沿空间维度的 concatenate 操作; F 为 1×1 卷积变换函数; δ 为非线性激活函数; \mathbf{z}^h 、 \mathbf{z}^w 分别为水平与竖直方向特征图。得到同时具有水平、竖直方向空间信息的特征图之后, 按空间维度将 f 分解为两个独立的张量, 并使用 1×1 卷积与 sigmoid 激活函数对张量分别进行处理, 将结果 g^h 、 g^w 分别展开作为水平与竖直方向的注意力权重, 其表达式如下所示:

$$g^h = \sigma(F_h(f^h)) \quad (7)$$

$$g^w = \sigma(F_w(f^w)) \quad (8)$$

其中: σ 是 sigmoid 激活函数; F_h 和 F_w 为 1×1 卷积变换。将权重与输入的特征图相乘, 使原始特征图具有注意力权重, 增强聚焦特征能力。使用嵌入 CoordAttention 模块的逆残差替换原网络中的逆残差结构, 提升 MobileNet 网络提取精确位置信息和感兴趣领域的能力, 且不过多增加计算开销。相较于 SE 模块^[27]只考虑通道间信息的编码而忽视位置信息, CBAM 模块^[28]仅能捕获局部相关性而无法捕获

更大区域的依赖关系, 而 CoordAttention 模块对于特征信息的提取更全面。

2.2 PANet 网络的改进

深层特征图具有较强的语义信息, 但分辨率低, 细节感知能力差, 适合大目标检测。浅层特征图感受野小, 包含更多位置、信息细节, 且几何表征能力强, 适合小目标检测。PANet 网络结构如图 4 所示, YOLOv4 使用该结构进行多尺度特征融合,

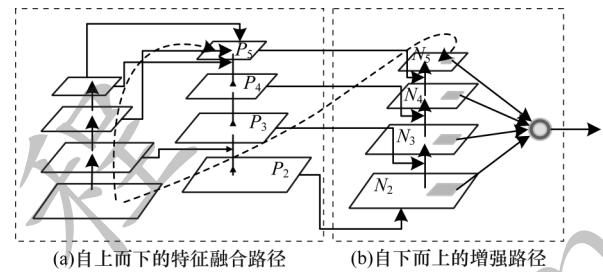


图 4 PANet 网络结构

Fig.4 Structure of PANet network

PANet 网络在原有 FPN 自上而下的特征融合路径 (如图 4(a) 所示) 的基础上, 增加自下而上的增强路径 (如图 4(b) 所示), 通过对浅层与深层特征图的反复提取, 融合不同尺度特征图的特征信息。但 PANet 网络的融合方式只是简单地将特征图转换成相同尺度再相加, 不能充分利用不同尺度的特征。使用 ASFF 网络改进 PANet 结构, 通过给不同尺度的特征融合分配自适应的权重参数, 实现多尺度特征图的高效融合, 改进后的 PANet 结构如图 5 所示。

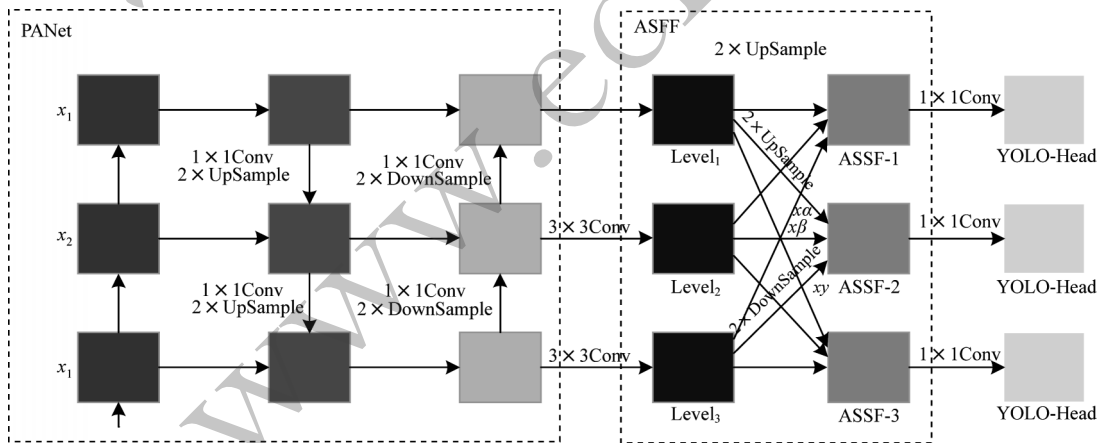


图 5 改进后的 PANet 网络结构

Fig.5 Structure of improved PANet network

以 ASFF-2 的计算过程为例, ASFF-2 由特征图 Level 1、Level 2 和 Level 3 融合得到, Level 1 通过 1×1 卷积将通道数调整到与 Level 2 一致, 再用插值方式调整尺寸与 Level 2 一致, 得到 $X^{1 \rightarrow 2}$; 对 Level 3 特征图进行 stride=2 的 3×3 卷积操作, 并调整通道数与 Level 2 一致, 得到 $X^{3 \rightarrow 2}$ 。Level 2 特征层即 $X^{2 \rightarrow 2}$, 最

后利用卷积核为 1×1 的卷积操作压缩通道数, 沿通道方向拼接, 再使用卷积将通道数降为 3, 对 3 个通道上的特征图使用 softmax 函数处理, 分别得到 $X^{1 \rightarrow 2}$ 、 $X^{2 \rightarrow 2}$ 、 $X^{3 \rightarrow 2}$ 的融合权重 α_{ij} 、 β_{ij} 、 γ_{ij} 。ASFF 网络的融合过程表达式如式(9)所示:

$$y_{ij}^l = \alpha_{ij}^l \times X_{ij}^{1 \rightarrow l} + \beta_{ij}^l \times X_{ij}^{2 \rightarrow l} + \gamma_{ij}^l \times X_{ij}^{3 \rightarrow l} \quad (9)$$

其中: X_{ij}^n 表示第 n 层特征图经变换后的特征图, 尺寸通道数均相同; $\alpha_{ij}, \beta_{ij}, \gamma_{ij}$ 分别表示 3 个不同尺寸特征图的融合权重, 通过 softmax 函数使得 $\alpha_{ij} + \beta_{ij} + \gamma_{ij} = 1$, $\alpha_{ij}, \beta_{ij}, \gamma_{ij} \in [0, 1]$; y'_{ij} 表示融合后的层特征图。

ASFF 网络实现了 $\alpha_{ij}, \beta_{ij}, \gamma_{ij}$ 这 3 个比例参数的自适应调整, 将其作为网络模型的可更新参数, 通过误差梯度反向传播更新融合权重, 使模型的多尺度特征融合更加充分, 能更有效地利用特征, 提高模型对小目标瑕疵的检测能力。

2.3 聚类先验框优化

YOLOv4 的先验框是在 PASCAL VOC 数据集中采用 K-means 算法聚类而来, 织物瑕疵数据集中小目标瑕疵较多, 粗经、轧痕等瑕疵长宽比较大, 因此需要针对织物缺陷数据集聚类出先验框。传统 K-means 聚类算法结果不稳定且易陷入局部最优, 导致聚类先验框与实际目标匹配度低, 影响检测精度。本文使用 K-means++ 聚类算法稳定初始化聚类中心选择, 具体过程如下:

1) 在数据点之间随机选择一个聚类中心;

2) 计算每个样本与当前已有聚类中心最短距离 $D(x)$, 数据点被选取作为聚类中心的概率与 $D(x)$ 大小正相关, 使用轮盘法选出下一个聚类中心;

3) 重复步骤 2 直到选择出 k 个聚类中心点。

K-means++ 能够优化初始聚类中心点的选取, 从而改善分类结果误差, 修改后的 anchor box 大小分别为 (4, 15)、(8, 10)、(12, 48)、(22, 52)、(36, 55)、(46, 52)、(55, 396)、(65, 118)、(228, 38)。

2.4 Focal Loss 函数改进

由于织物瑕疵数据中瑕疵区域占整个织物图像的比例很小, 若将不含瑕疵的区域当做负样本, 含瑕疵的区域为正样本, 则不含瑕疵区域的负样本数量远多于含瑕疵的正样本数量。正负样本数量严重失衡可能导致给模型提供的瑕疵特征较少, 影响模型收敛, 因此采用 Focal Loss 函数优化分类损失。类别损失、置信度损失和位置损失共同构成 YOLOv4 的损失函数, 其中, 原始 YOLOv4 的类别损失为交叉熵损失函数, 其表达式如式 (10) 所示:

$$L = - \sum_{j=1}^C y \log_a y' \quad (10)$$

其中: L 表示交叉熵; C 表示分类种类数; y 为期望输出, 表示经过激活函数后的输出。改进后的 Focal loss 函数表达式如式 (11) 所示:

$$L_f = -\alpha[1 - y]^\beta \log_a y' \quad (11)$$

其中: α 为矫正系数, 作用是平衡样本数量; β 为超参数, 一般取 2。通过引入 α 和 β 两个约束参数, 可以在网络训练过程中自动对正负样本不平衡进行调整。

3 实验结果与分析

本次实验操作系统为 Windows10, CPU 为 Intel Core i7-8700K, 内存 16 GB, GPU 为 NVIDIA GeForce GTX1080, 显存 8 GB, 选用 Pytorch 作为深度学习框架, 开发工具为 Pycharm2019.1。初始学习率设定为 0.001, Momentum 动量参数设为 0.2, IOU 阈值设置为 0.5, 共训练 300 个 epoch。

3.1 实验数据集及评价指标

本次实验所使用的数据集来自阿里天池布匹疵点数据集, 为验证本文改进算法在织物瑕疵检测任务, 尤其是对小目标瑕疵检测任务中的性能提升, 从其中选取实际生产较为常见的 6 种缺陷, 包括结头、三丝、破洞、粗经、纬缩和轧痕图片共 2 026 张, 其中结头、三丝、破洞为小目标瑕疵。采用图像翻转、图像缩放、图像融合等数据增强方式扩充至 8 104 张, 处理训练图片大小为 608×608 像素, 各类瑕疵图片数量如表 2 所示, 训练集、测试集和验证集的划分比例为 8:1:1。

表 2 阿里天池布匹疵点数据集的类别统计结果

Table 2 Statistics results of alibaba tianchi fabric defect dataset category

类别	数量
结头	2 000
三丝	1 728
破洞	1 040
粗经	968
纬缩	1 248
轧痕	1 120

本文实验的评价指标为平均精度 (Average Precision, AP), 平均精度均值 (mean Average Precision, mAP), 帧率 (Frame Per Second, FPS), FPS 指模型每秒能够检测的图片数量, 可以作为检测速度的衡量指标。平均精度指标的表达式如下所示:

$$A_{AP} = \int_0^1 p(r) dr \quad (12)$$

其中: p 为准确率; r 为召回率。准确率和召回率的表达式如下所示:

$$p = \frac{T_{TP}}{T_{TP} + F_{FP}} \quad (13)$$

$$r = \frac{T_{TP}}{T_{TP} + F_{FN}} \quad (14)$$

其中: T_{TP}, F_{FP} 分别表示正确、错误判断的瑕疵样本个数, F_{FN} 为被错误检测的负样本数量。

平均精度均值的表达式为:

$$m_{AP} = \frac{1}{N} \sum_{i=1}^N A_{AP}^{(i)} \quad (15)$$

3.2 主流算法与本文算法对比

为验证本文算法在织物瑕疵检测任务中的性

能,将主流目标检测算法 Faster-RCNN^[20]、YOLOv4、YOLOv4-Tiny、YOLOv4-MobileNet 与本文改进 YOLOv4 算法在相同的数据集上进行训练,结果如表 3 所示。由表 3 可知,Faster-RCNN 算法的 mAP 较高,但检测速度较慢,模型体积大,对算力要求高,难以满足织物瑕疵检测实时性的要求。将本文改进 YOLOv4 算法与 YOLOv4 算法相比,mAP 提升 2.3 个百分点,检测速度提升了 12 frame/s,且本文模型体积相较于 YOLOv4 算法大幅减少,能够满足工业生产部署需求。YOLOv4-Tiny 算法是轻量级检测算法的代表之一,参数量仅为 YOLOv4 算法的 1/10,检测速度满足实时性需求,但结构简单,无法满足检测精度要求。

表 3 主流算法与本文算法的对比
Table 3 Comparison between mainstream algorithm and improved algorithm in this paper

算法	mAP/%	模型大小/ MB	帧率/ (frame·s ⁻¹)
Faster-RCNN 算法	87.55	108.0	5
YOLOv4 算法	89.37	244.0	14
YOLOv4-Tiny 算法	78.61	22.6	58
YOLOv4-MobileNetv2 算法	83.79	46.5	29
本文算法	91.67	55.0	26

由表 3 还可知,本文算法的 mAP 相较于 YOLOv4-Tiny 算法高 9.06 个百分点。将 YOLOv4 原特征提取网络 CSPDarkNet53 替换为 MobileNet 网络后,检测速度提升了 15 frame/s,mAP 减少 5.58 个百分点,原因是特征提取网络层数减少,导致特征提取的效果下降,但模型计算量减小,检测速度有所提高。与仅将 YOLOv4 主干网络替换为 MobileNet 的算法对比,本文算法 mAP 提升了 7.88 个百分点,这充分说明 ASFF 算法对 PANet 网络改进的积极作用。在 MobileNet 逆残差块中加入 CoordAttention 模块,在没有增加较多模型体积的前提下,提高了模型的 mAP,检测速度仅降低 3 frame/s。

通过对比实验数据,本文算法在织物瑕疵检测任务中相较于大型网络检测 mAP 相近,检测速度更快,模型体积更小。相较于轻量级网络,本文算法瑕疵检测能力更强,mAP 更高,更适用于织物瑕疵检测任务。

3.3 消融实验

为验证 ASFF 算法、CoordAttention 注意力模块和 Focal Loss 损失函数对模型性能的影响,以将主干网络替换为 MobileNet 的 YOLOv4 算法作为基准,对加入的模块进行消融实验。实验结果如表 4 所示,其中“×”表示不加入,“√”表示加入。

表 4 消融实验结果

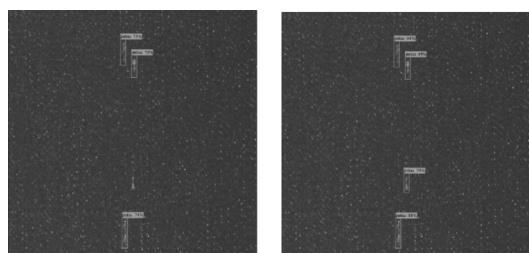
Table 4 Results of ablation experiment

基准算法	CoordAttention 模块	ASFF 算法	Focal Loss 函数	AP/%						mAP/%
				结头	三丝	破洞	纬缩	粗经	轧痕	
√	×	×	×	80.92	85.16	84.56	79.40	85.67	87.05	83.79
√	√	×	×	84.75	88.91	89.10	82.87	87.38	89.31	87.22
√	√	√	×	88.60	92.94	92.58	85.69	89.36	91.84	90.17
√	√	√	√	89.68	94.40	93.75	87.95	90.85	93.38	91.67

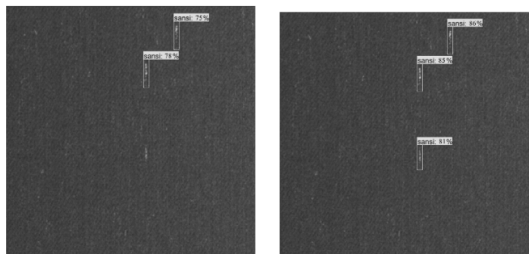
由表 4 可知,在 MobileNet 网络的逆残差结构中插入 CoordAttention 注意力模块后,小目标瑕疵的 AP 分别提升了 3.83、3.75、4.54 个百分点,纬缩、粗经、轧痕三种缺陷的 AP 也有所提升,这说明加入 CoordAttention 注意力模块提高了 MobileNet 的特征提取能力,可以捕获更精确的位置信息与缺陷特征,对于小目标瑕疵检测效果较好。在此基础上加入使用 ASFF 网络改进的 PANet,相较于原始的 PANet,改进后的 PANet 在结头、三丝、破洞等小目标缺陷上的 AP 分别提高了 3.85、4.03、3.48 个百分点,表明相较于原始的 PANet 网络,使用 ASFF 网络改进 PANet 结构,利用多尺度特征图将模型训练学习到的权重进行融合,能大幅提升对小目标瑕疵的检测能力。最后,加入 Focal Loss 损失函数,能使算法整体 mAP 和精度提高。

3.4 与原始 YOLOv4 算法检测效果对比

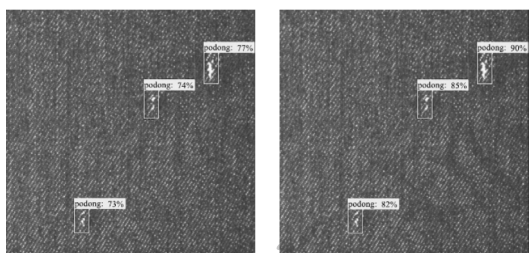
为验证本文算法相较于原始 YOLOv4 算法在织物瑕疵检测任务中的优势,选取结头、三丝、破洞图片各一张,进行检测结果对比,YOLOv4 与本文算法检测结果对比如图 6 所示。图 6 中 6 组检测效果图中,左侧为 YOLOv4 检测效果,右侧为本文算法的检测效果。从中可以看出,原始 YOLOv4 算法在结头、三丝等小目标瑕疵检测中存在漏检情况,本文算法有效改善了漏检情况,并且检测置信度值得到了明显提高。这充分说明 CoordAttention 模块对逆残差结构的改进,增强了模型对于小目标瑕疵的聚焦能力。ASFF 网络对 PANet 的改进,充分融合与利用多尺度特征图的信息,有效增强对于小目标瑕疵的检测性能。同时,通过 K-means++ 聚类算法聚类先验框使模型检测框更加适用于粗经、轧痕等长宽比较大的瑕疵类型检测。



(a)结头检测效果



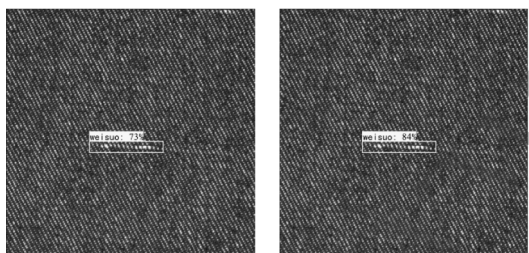
(b)三丝检测效果



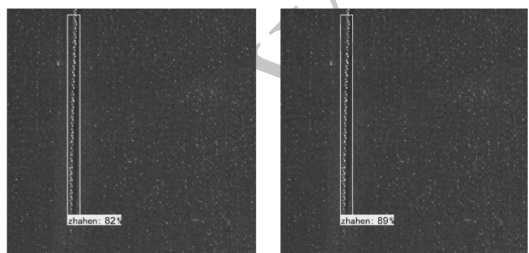
(c)破洞检测效果



(d)粗经检测效果



(e)纬缩检测效果



(f)轧痕检测效果

图6 YOLOv4算法与本文算法检测效果对比

Fig.6 Comparison of detection effect between YOLOv4 algorithm and algorithm in this paper

4 结束语

本文提出一种基于多尺度特征自适应融合的轻量化织物瑕疵检测算法,通过在 MobileNetv2 网络的逆残差块中加入 CoordAttention 注意力模块,并使用该网络作为 YOLOv4 主干网络,减少参数总量与计算量,增强网络聚焦小目标瑕疵特征能力。使用自适应空间特征融合结构改进 PANet 网络,提高小目标瑕疵检测精度,采用 K-means++ 算法获得适合的候选框,并使用 Focal Loss 函数优化分类损失降低正、负样本不平衡对检测结果的影响。实验结果表明,本文算法相较于原始 YOLOv4 算法的平均精度均值提升了 2.3 个百分点,帧率提高了 12 frame/s,模型体积大幅减少。下一步将通过知识蒸馏和模型剪枝技术,在保证检测精度的基础上降低参数数量和计算量,提高模型检测速度,并针对瑕疵特征改进模型,提高算法的检测精度。

参考文献

- [1] KUMAR A. Computer-vision-based fabric defect detection: a survey[J]. IEEE Transactions on Industrial Electronics, 2008, 55(1): 348-363.
- [2] DING S M, LIU Z F, LI C L. AdaBoost learning for fabric defect detection based on HOG and SVM[C]//Proceedings of International Conference on Multimedia Technology. Washington D. C., USA: IEEE Press, 2011: 2903-2906.
- [3] HANBAY K, TALU M F, ÖZGÜVEN Ö F. Fabric defect detection systems and methods—a systematic literature review[J]. Optik, 2016, 127(24): 11960-11973.
- [4] MIRMAHDARI S A, AHMADYFARD A, SHAHRABI A A, et al. A novel modeling of random textures using Fourier transform for defect detection[C]//Proceedings of the 15th International Conference on Computer Modelling and Simulation. Washington D. C., USA: IEEE Press, 2013: 470-475.
- [5] 任欢欢,景军锋,张缓缓,等. 应用 GIS 和 FTDT 的织物错花缺陷检测研究[J]. 激光与光电子学进展, 2019, 56(13): 94-99.
REN H H, JING J F, ZHANG H H, et al. Cross-printing defect detection of printed fabric using GIS and FTDT[J]. Laser & Optoelectronics Progress, 2019, 56(13): 94-99. (in Chinese)
- [6] LIU L L, ZHANG H J, XU X F, et al. Collocating clothes with generative adversarial networks cosupervised by categories and attributes: a multidiscriminator framework [J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(9): 3540-3554.
- [7] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Conditional generative adversarial nets[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [8] ZHAO M B, LIU Y, LI X R, et al. An end-to-end framework for clothing collocation based on semantic feature fusion[J]. IEEE MultiMedia, 2020, 27(4): 122-132.
- [9] SCHUSTER M, PALIWAL K K. Bidirectional recurrent neural networks [J]. IEEE Transactions on Signal Processing, 1997, 45(11): 2673-2681.

- [10] ZHANG H J, WANG X H, LIU L L, et al. WarpClothingOut: a stepwise framework for clothes translation from the human body to tiled images[J]. IEEE MultiMedia, 2020, 27(4): 58-68.
- [11] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of Advances in Neural Information Processing Systems. New York, USA: ACM Press, 2014: 2672-2680.
- [12] 许玉格, 钟铭, 吴宗泽, 等. 基于深度学习的纹理布匹瑕疵检测方法[J]. 自动化学报, 2020, 45(7): 1-15.
XU Y G, ZHONG M, WU Z Z, et al. Detection of detecting textured fabric defects based on deep learning. [J]. Acta Automatica Sinica, 2020, 45(7): 1-15. (in Chinese)
- [13] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016: 770-778.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. [2021-11-03]. <https://arxiv.org/abs/2004.10934>.
- [15] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2018: 4510-4520.
- [16] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2021: 13708-13717.
- [17] LIU S T, HUANG D, WANG Y H. Learning spatial fusion for single-shot object detection[EB/OL]. [2021-11-03]. <https://arxiv.org/abs/1911.09516>.
- [18] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2018: 8759-8768.
- [19] ARTHUR D, VASSILVITSKII S. K-means++: the advantages of careful seeding[C]//Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms. New York, USA: ACM Press, 2007: 1027-1035.
- [20] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2017: 2999-3007.
- [21] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Proceedings of European Conference on Computer Vision. Berlin, Germany: Springer, 2016: 21-37.
- [22] 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.
- [23] 吴涛, 王伟斌, 于力, 等. 轻量级 YOLOV3 的绝缘子缺陷检测方法[J]. 计算机工程, 2019, 45(8): 275-280.
WU T, WANG W B, YU L, et al. Insulator defect detection method for lightweight YOLOV3 [J]. Computer Engineering, 2019, 45(8): 275-280. (in Chinese)
- [24] WANG C Y, MARK LIAO H Y, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Washington D. C., USA: IEEE Press, 2020: 1571-1580.
- [25] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [26] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 936-944.
- [27] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2018: 7132-7141.
- [28] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision. Berlin, Germany: Springer, 2018: 3-19.