

DOI: 10.19333/j.mfkj.20221001607

基于 YOLOv5 算法的女性职业装款式自动识别方法

焦 哲¹, 夏乾龙¹, 王巧铃¹, 卢致文^{1,2}, 路丽莎³, 刘 锋¹

(1. 太原理工大学 轻纺工程学院, 山西 晋中 030600; 2. 安徽天助纺织科技集团股份有限公司, 安徽 阜阳 236000;

3. 武汉纺织大学 服装学院, 湖北 武汉 430073)

摘 要:针对目前自动识别方法存在的图像背景处理繁琐、形式单一、轮廓特征提取复杂、单张图片只有一种款式可以识别等问题,运用了一种单阶段目标检测(YOLOv5)算法,实现了基于服装图片的职业装款式分类。以办公室女性职业装为例,首先建立一个包含衣裤套装、衣裙套装、连衣裙的3类女性职业装经典款式数据集并对图片进行标注,使用Mosaic数据增强操作提升模型的训练速度和网络的精度,选用CIOU_Loss做定位损失的损失函数,最终实现女性职业装3类款式的自动检测识别。实验表明,基于YOLOv5算法的款式识别方法可以实现对复杂背景下的女性职业装的有效分类,可为服装商品的可视化分类识别提供新的思路,提高服装产品相关调研的效率。

关键词:YOLOv5 算法; 女性职业装; 款式识别; 图片处理; 识别效率

中图分类号:TS 941.2

文献标志码:A

Automatic recognition method of women's professional dress style based on YOLOv5 algorithm

JIAO Zhe¹, XIA Qianlong¹, WANG Qiaoling¹, LU Zhiwen^{1,2}, LU Lisha³, LIU Feng¹

(1. College of Textile Engineering, Taiyuan University of Technology, Jinzhong, Shanxi 030600, China;

2. Anhui Tianzhu Textile Science Technology Co., Ltd., Fuyang, Anhui 236000, China;

3. School of Fashion, Wuhan Textile University, Wuhan, Hubei 430073, China)

Abstract: In view of the problems existing in the current automatic recognition methods, such as complicated image background processing, single form, complicated contour feature extraction and only one style of a single picture can be recognized, a single-stage object detection (YOLOv5) algorithm was applied to realize the classification of professional clothing styles based on clothing pictures. Taking office women's professional dress as an example, firstly, a sample set of three classic styles of women's professional dress, including pants, dresses and skirts, was established and the pictures were marked. The Mosaic data was used to enhance the training speed of the model and the accuracy of the network. CIOU_Loss was selected as the loss function of positioning loss, and finally the automatic detection and identification of three types of women's professional dress was realized. The experiment shows that the style recognition method based on YOLOv5 algorithm can effectively classify women's professional wear under complex background, provide new ideas for visual classification and recognition of clothing products, and improve the efficiency of related research on clothing products.

Keywords: YOLOv5 algorithm; women's business attire; style recognition; picture processing; identification efficiency

职业装是上班族职场生活中必不可少的服装, 对企业形象、工作效率、员工情感都会有一定的影

收稿日期:2022-10-09

基金项目:山西省哲学社会科学规划课题(2022YY029);山西省回国留学人员科研资助项目(2022-090)

第一作者:焦哲,硕士生,主要研究方向为服装数字化设计,E-mail:jiaozhe0930@link.tyut.edu.cn。

通信作者:刘锋,副教授,主要研究方向为服装文化研究与数字化技术,E-mail:liufeng@tyut.edu.cn。

响^[1]。女士职业装由传统的西服西裤发展到西装裙,款式越来越多元化。越来越多的企业开始注重自身企业文化及员工仪表^[2]。计算机技术广泛应用于电商行业,其中基于计算机视觉的自动分类技术有助于零售商快速地实现服装商品的自动分类,便于消费者跨越语义鸿沟、准确地检索到自己心仪的服装商品。

目前国内外学者对于计算机视觉技术应用到服装行业的研究主要集中在人体体型和服装商品2个方面。Jiang等^[3]提出了一种基于YOLOv3(单阶段目标检测算法YOLO的第3个版本)的特定场景的服装识别方案;Chang等^[4]研究了一种轻量级学习算法YOLOv5s(单阶段目标检测算法YOLO的第5个版本的一种)用于服装风格的识别;倪世明等^[5]通过将MEA(一种全局优化算法)引入BP神经网络预测模型实现对青年女性体型的识别;吴欢等^[6]采用卷积神经网络CaffeNet模型自动提取形状特征,实现了对女裤廓形的分类;尹光灿等^[7]基于Alex Net卷积神经网络,运用Softmax回归分类器来实现服装领型的分类;邓莹洁等^[8]提出了一种带有Inception v2模组的快速区域卷积神经网络模型的女装半身裙多特征分类识别方法;贾小军等^[9]基于卷积神经网络实现了对蓝印花布纹样基元的分类。然而BP神经网络和卷积神经网络模型需要对图片样本背景进行繁琐的处理,女裤廓形的识别需要选取白色背景的平铺图,手动搜索样本浪费时间长,而实际上服装图片背景均较为复杂^[10],不易识别,而且对于单张图片的识别只有一种结果,不能实现对单张图片中多种款式的分类。

针对以上问题,本文拟利用YOLOv5算法使用python工具实现服装图片的款式分类。YOLOv5算法使用Mosaic数据增强操作提升模型的训练速度和网络精度,选用CIOW_Loss做定位损失的损失函数,以实现衣裤套装、衣裙套装、连衣裙3种款式的女性职业装自动检测识别。此方法对于样本的选择局限小,一张图片中的多种款式可以明确分类,视觉检测识别更加多样化。

1 单阶段目标检测算法(YOLOv5)

YOLO是一种采用一个单独的卷积神经网络模型的目标检测算法,核心思想是利用整张图片作为网络的输入,直接在输出层回归 bounding box(边界框)的位置及其所属的类别。YOLO算法的优点是能够对大区域面积的图像进行识别,对于复杂背景图片具有较好的检测效果。

YOLO算法经过YOLOv1, YOLOv2, YOLOv3,

YOLOv4的发展和一次次的优化改进,形成完整版本的YOLOv5, YOLOv5有YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 5个模型。这些不同的模型使得YOLOv5能很好地在精度和速度中权衡,方便用户选择。本文选取速度和精度适中的YOLOv5s和YOLOv5m 2类算法进行实验。

1.1 输入端

YOLOv5在训练模型阶段的输入端共有3部分,分别是Mosaic数据增强方法、自适应锚框计算、自适应图片缩放。Mosaic数据增强方法是使用一系列操作将4张图片拼成一张图片进行网络训练;自适应锚框计算是在训练过程中根据数据集特点计算最佳初始锚框值,网络训练会在初始锚框的基础上得到预测框,与真实框比较得到差值,进一步迭代训练;自适应图片缩放是将不同长宽的图片缩放成统一大小,便于网络训练计算。

1.2 主干网络和颈部网络特征提取和特征融合

主干网络和颈部网络分别是Backbone网络和Neck网络,其中Backbone网络包括Focus结构、CSP结构,用于特征提取。Focus结构中重要的是切片操作,4×4×3的图像经切片后会变成2×2×12的特征图。以YOLOv5s的结构为例,在原始640×640×3的图像中输入Focus结构,采用切片操作,图像会先变成320×320×12的特征图,再经过一次32个卷积核的卷积操作,最终变成320×320×32的特征图。而在YOLOv5m结构中,有48个卷积核,所以经过卷积操作后输出为320×320×48的特征图^[11]。YOLOv5的Neck网络包括FPN结构和PAN结构,用于特征融合,采用CSPNet设计的CSP2结构,从而加强网络的特征融合能力,PAN加入了自底向上的路线,对特征层进行上采样,弥补并加强定位信息^[12]。Neck网络图片特征融合流程图如图1所示。

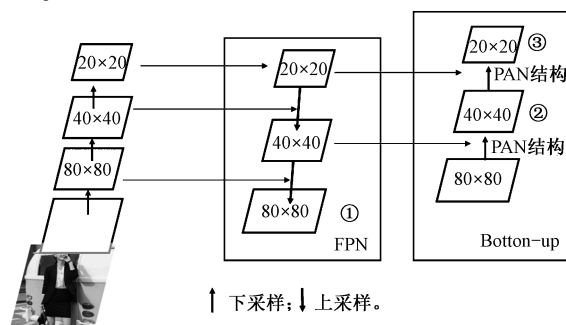


图1 Neck网络图片特征融合流程图

Fig.1 Neck network picture feature fusion flow chart

1.3 输出端

输出端是预测层,如图 1 所示,用于预测的 3 个特征图是 20×20 、 40×40 、 80×80 。损失函数用来评估预测准确性。YOLOv5 的损失有 3 种,分别是分类损失(cls_loss)、定位损失(box_loss)、置信度损失(obj_loss)。分类损失函数和置信度损失函数采用的是 BCEWithLogits_Loss 函数,定位损失函数采用的是 CIOU_Loss 函数^[13]。采用标准的非极大值抑制操作(NMS)以滤除多余的预测框,得出理想的预测结果。

2 YOLOv5 实验

2.1 数据集选取

女性职业装多指从事脑力劳动的白领及其他精英阶层女性在公务、商务活动时所穿着的服装^[14],是现代企业文化的重要组成部分^[15]。然而目前并没有系统的职业装分类样本库,大类上衣裤套装、衣裙套装、连衣裙 3 种款式是女性职业装的常见经典款式。本文从淘宝、京东、阿里巴巴等购物网站搜集 3 种款式各 1 200 张共 3 600 张女士职业装图片,文件格式为 jpg,女性职业装代表图像如图 2 所示。



图 2 3 种款式女性职业装代表图像

Fig.2 3 kinds represent images of womens' professional dress.(a) Suit pants;(b) Suit dress;(c) Suit skirt



2.2 数据集预处理

用 jpeg resizer 图片处理软件统一图片大小和分辨率,图片分辨率为 $640 \text{ dpi} \times 640 \text{ dpi}$,大小在 65 k 以内。用 labeling1.8.6 标注工具对数据标注,标注格式为 YOLO 格式,标注为 3 类:suit pants(衣裤套装)、suit dress(衣裙套装)、suit skirt(连衣裙)。标注出最需要识别的图片内容主题,避免背景或图片中其他服装影响图片识别准确率。用分类程序将数据集分为训练集、验证集、测试集,分配比例为 3 : 1 : 1。最终得到训练集 2 160 张,验证集和测试集各 720 张。标注完成后,每一张图片都对应着一个与该图片名称相同的 txt 文件,txt 文件中的每 1 行表示 1 个标记实例,1 个标注框对应 1 行数据,每 1 行有 5 列,从左到右分别表示标签类别、标记框中心横坐标与图片宽度的比值、标记框中心纵坐标与图片高度的比值、标记框宽度与图片宽度的比值、标记框高度与图片高度的比值^[16]。标注过程如图 3 所示,服装定位识别训练过程如图 4 所示。



图 3 女性职业装图像标注

Fig.3 Part of women's professional dress image annotation frame



图 4 服装定位识别训练过程

Fig.4 Training process of clothing positioning recognition

2.3 实验配置

测试和验证训练实验进行机器学习的操作系统

为 window11 64 位系统,GPU 为 NVIDIA GeForce RTX 3060 Laptop,搭建深度学习框架为

Pytorch1.12.1 和程序语言为 python3.8 的运算环境平台,图片标注框架配置处理包为 labelme,PC 界面

使用 pyqt5 来实现。配置实验环境后的算法流程图如图 5 所示。

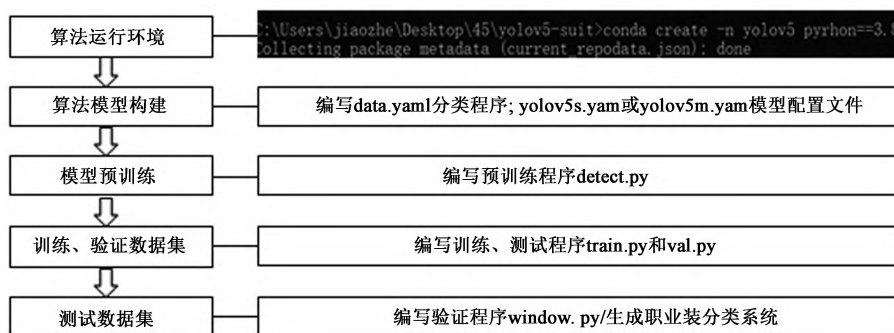


图 5 YOLOv5 算法流程图

Fig.5 YOLOv5 algorithm flow chart

2.4 实验训练和验证

设置不同的训练轮数来进行实验训练 (train) 和验证 (val) (epoch), 训练轮数分别为 15、30、100、300。用 YOLOv5s 和 YOLOv5m 做对比实验,二者具体的损失值、精准率 (precision, P)、召回率 (recall, R)、不同交并比 IoU (intersection over union) 阈值 (0.5 : 0.95) 上的综合准确程度 mAP (mean average precision) 变化过程如图 6、7 所示。精准率和召回

率的计算公式如式 (1) (2) 所示。

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

式中: TP 表示将正例预测为正例的个数; FP 表示将反例预测为正例的个数; FN 表示将反例预测为反例的个数^[17]。

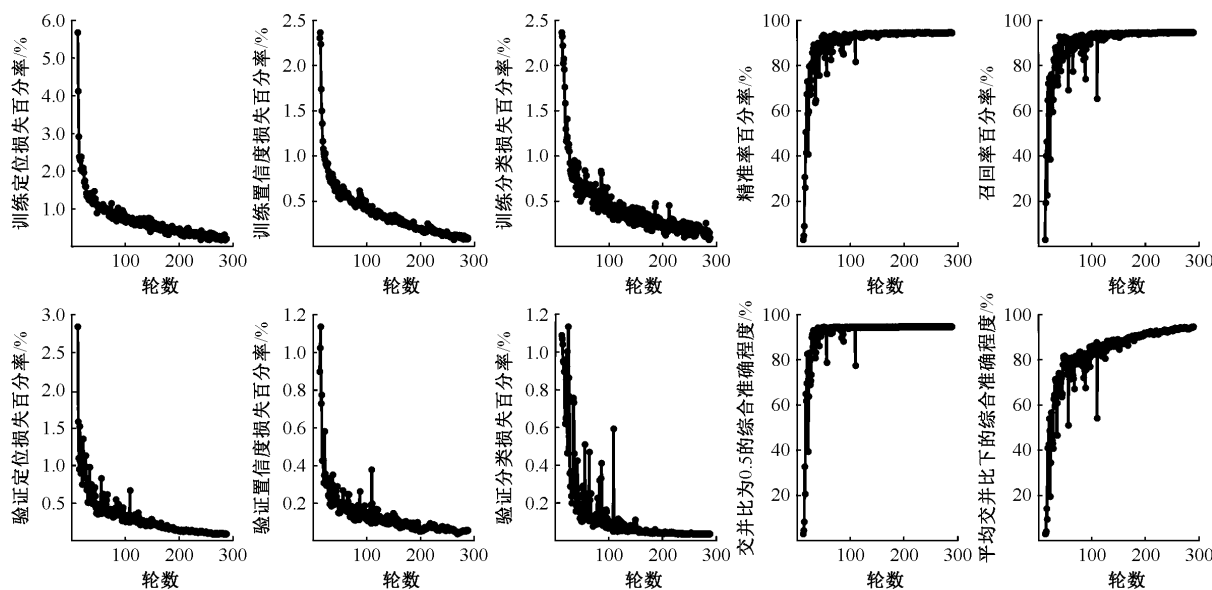


图 6 YOLOv5s 算法训练 300 轮过程图

Fig.6 Training process diagram of YOLOv5s algorithm

由图 6 可知,在数据集训练和验证过程中,损失值 (定位损失、置信度损失、分类损失) 逐渐降低,误差不断下降,精准率和召回率逐渐提高。最终损失值接近于 0,精准率接近于 1,说明预测信息越来越接近理想状态。由图 7 可知,在数据集训练和验证过程中,定位损失和置信度损失呈平稳下降状态,最终接近于

0。训练模型的最优化可以使职业装分类窗口页面获得更高的测试准确度。训练 300 轮时,通过图 6、7 的对比发现,YOLOv5s 的精准率和召回率较低,误检漏检现象相对较重。YOLOv5s 和 YOLOv5m 模型训练在 300 轮时达到最优,YOLOv5m 的 3 类损失、精准率、召回率、 mAP 训练效果更好。

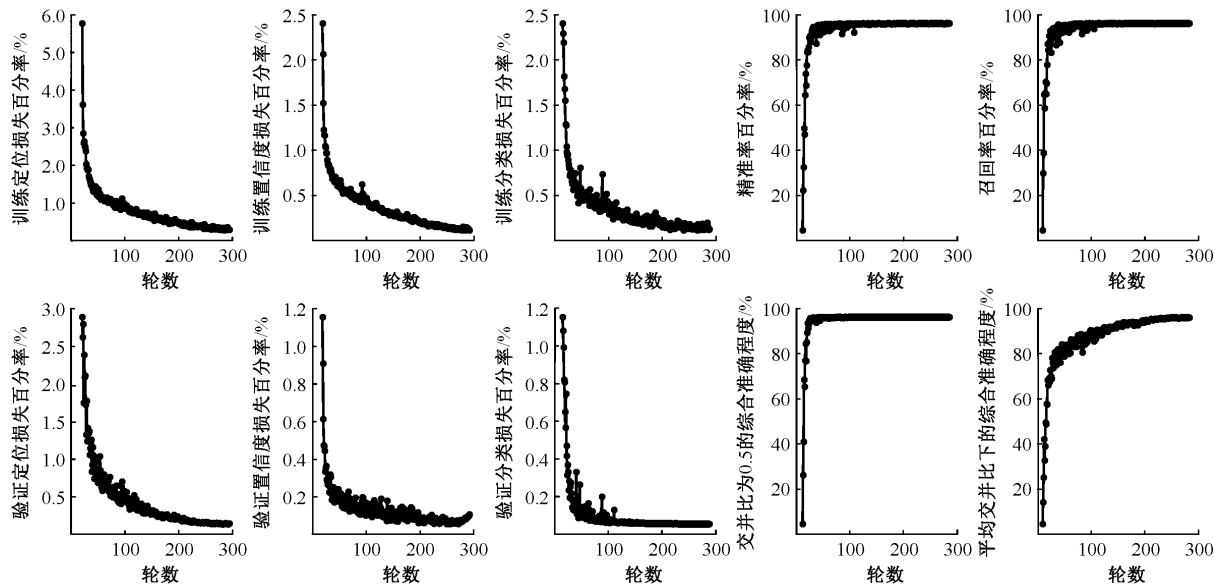


图 7 YOLOv5m 算法训练 300 轮过程图

Fig.7 Training process diagram of YOLOv5m algorithm

3 结果与分析

最终的训练验证结果用 AP 值和 mAP 值来表示,其中 AP (average precision) 表示各类职业装的训练验证准确程度, mAP 表示 3 类职业装的综合准确程度。评价指标计算公式如式(3)(4)所示。具体训练结果见表 1、2。

$$AP = \int_0^1 P(R) dR \quad (3)$$

$$mAP = \frac{1}{N} \sum_{i=0}^{N-1} AP_i \quad (4)$$

式中: N 表示类别,本文为 3; $P(R)$ 即精准率 P 。

表 1 不同训练轮数下的综合准确程度

Tab.1 mAP of different training rounds %		
训练轮数	YOLOv5s	YOLOv5m
15	56.4	58.4
30	62.3	68.3
100	92.7	93.8
300	95.7	96.5

表 2 训练 300epoch 下的结果

Tab.2 Training results under 300epoch						
训练模型	训练时间/h	模型精准率/%	mAP /%	连衣裙 AP /%	衣裙套装 AP /%	衣裤套装 AP /%
YOLOv5s	9.2	95.5	95.7	92.7	98.4	96.0
YOLOv5m	13.5	96.2	96.5	93.1	98.8	97.6

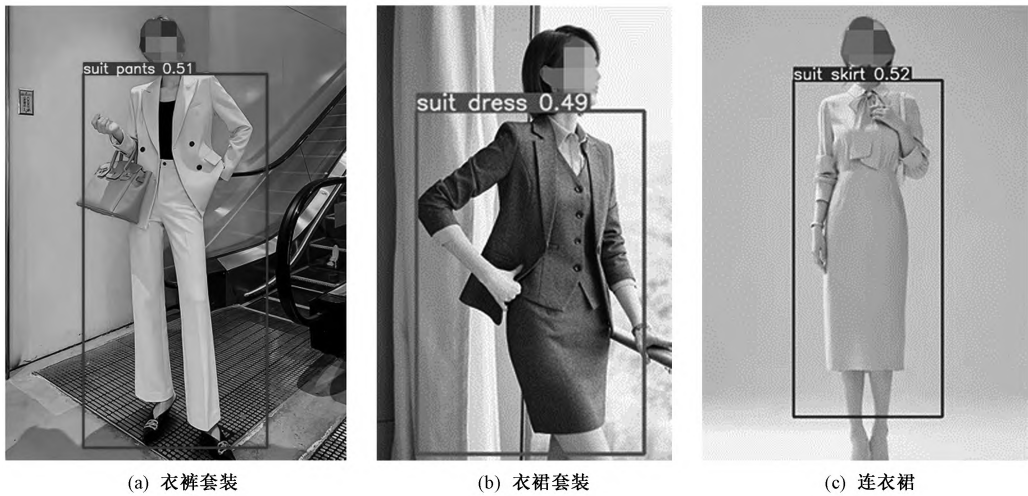


图 8 部分职业装分类系统图片检测结果图

Fig.8 Picture inspection results of professional wear classification system.(a) Suit pants;(b) Suit dress;(c) Suit skirt

由表 1 可知,不同的训练轮数,直接影响训练结果的 mAP 值。训练轮数越高,训练验证的职业装分类准确程度越高,当训练轮数为 15 时,YOLOv5m 模型训练的 mAP 值为 58.4%,比 YOLOv5s 模型训练的 mAP 值高 2%;训练轮数为 300 时,YOLOv5m 模型训练的 mAP 值达到 96.5%,比 YOLOv5s 模型训练的 mAP 值高 0.8%。由表 2 可知,YOLOv5m 模型的训练时间较长,模型精准率和综合准确程度(mAP 值)较高,其中衣裙套装的训练准确程度 AP 较高,最高达到 98.8%,衣裤套装次之,整个训练和验证实验能够实现对职业装良好的分类效果。用样本集中的 720 张测试集图片做职业装分类系统的测试,部分职业装分类系统图片检测结果如图 8 所示。

通过测试样本测试集中的 720 张图片,除去 3 张无效识别图片,剩余 717 张图片中仅仅有 15 张图片识别错误。其中密集服装图片的识别效果低,单张图片一套服装识别效率高。半身照片、人体不能完全呈现的照片识别效率低。职业装分类系统图片测试准确率达到 97.9%。由此可见,基于 YOLOv5 的职业装分类系统可以对大数据下的女性职业装图片实现有效的分类。

4 结 论

基于 YOLOv5m 算法和 YOLOv5s 算法做对比实验,利用 python 语言编写了基于图片分类的职业装款式识别程序。通过分类程序以 3:1:1 比例将数据集分为训练集、验证集以及测试集,以实验训练和验证来设置不同的训练轮数,得出训练轮数为 300 时,YOLOv5m 模型训练验证出的综合准确程度 mAP 值达到 96.5%,比 YOLOv5s 模型训练的综合准确程度 mAP 值高 0.8%,分类效果较好。采用 YOLOv5m 模型对职业装分类系统图片测试准确率达到 97.9%。实现了对大数据下的女性职业装图片有效的分类,可为服装电商人员图片搜集处理提供一定的技术支持。针对测试集中对密集服装的图片识别效果低的问题,后续可对 YOLOv5 算法做改进,提高职业装分类系统的综合准确率。

参考文献:

- [1] 范福军,吴俊,郭彤彤. 职业装色彩对 IT 职员情感影响的探讨[J]. 纺织学报,2009,30(10):120-123.
FAN Fujun, WU Jun, WU Rongrong. Influence of business wear color on feelings of employees in IT enterprises [J]. Journal of Textile Research, 2009, 30(10):120-123.
- [2] 赵永刚. 职业装袖缝模板创新工艺设计与应用[J]. 毛纺科技,2020,48(6):77-80.
ZHAO Yonggang. Innovative process design and

- application of professional sleeve seam template [J]. Wool Textile Journal, 2020,48(6):77-80.
- [3] JIANG M Q, YANG X, FANG X. Scene-specific clothing recognition based on YOLOv3 algorithm [J]. Journal of Research in Science and Engineering, 2020(2):1-10.
- [4] CHANG YH, ZHANG YY. Deep learning for clothing style recognition using YOLOv5 [J]. Micromachines, 2022, 13(10):1678-1678.
- [5] 倪世明,白云龙,蒋益群. 基于 MEA-BP 神经网络的青年女性臀部体型识别[J]. 丝绸,2022,59(5):42-49.
NI Shiming, BAI Yunlong, JIANG Yiqun. Young female arm body shape recognition based on MEA-BP neural network [J]. Journal of Silk, 2022,59(5):42-49.
- [6] 吴欢,丁笑君,李秦曼,等. 采用卷积神经网络 CaffeNet 模型的女裤廓形分类[J]. 纺织学报,2019,40(4):117-121.
WU Huan, DING Xiaojun, LI Qinman, et al. Classification of women's trousers silhouette using convolution neural network CaffeNet model [J]. Journal of Textile Research, 2019,40(4):117-121.
- [7] 尹光灿,罗戎蕾. 基于卷积神经网络的服装领型识别与分类研究[J]. 现代纺织技术,2020,28(3):48-53.
YIN Guangcan, LUO Ronglei. Research on recognition and classification of garment collar type based on convolutional neural network [J]. Advanced Textile Technology, 2020,28(3):48-53.
- [8] 邓莹洁,罗戎蕾. 基于卷积神经网络的半身裙款式特征分类识别[J]. 现代纺织技术,2021,29(6):98-105.
DENG Yingjie, LUO Ronglei. Classification and recognition of bust skirt style and common features based on convolutional neural network [J]. Advanced Textile Technology,2021,29(6):98-105.
- [9] 贾小军,叶利华,邓洪涛,等. 基于卷积神经网络的蓝印花布纹样基元分类[J]. 纺织学报,2020,41(1):110-117.
JIA Xiaojun, YE Lihua, DENG Hongtao, et al. Elements classification of vein patterns using convolutional neural networks for blue calico [J]. Journal of Textile Research, 2020,41(1):110-117.
- [10] 杜娟,胡静. 基于变分自编码器的现代服饰局部中国风格迁移[J]. 毛纺科技,2021,49(9):72-77.
DU Juan, HU Jing. Local Chinese style migration of modern clothing based on variational auto encoder [J]. Wool Textile Journal, 2021,49(9):72-77.
- [11] 钱坤,李晨瑄,陈美杉,等. 基于 YOLOv5 的舰船目标及关键部位检测算法[J]. 系统工程与电子技术,2022,44(6):1823-1832.
QIAN Kun, LI Chenxi, CHEN Meishan, et al. Ship target and key parts detection algorithm based on YOLOv5 [J]. Systems Engineering and Electronics, 2022,44(6):1823-1832.
- [12] 谢团结,林贤伟,胡连信,等. 基于改进 YOLOv5 算法的织物疵点检测系统[J]. 棉纺织技术,2022,50(11):15-20.
XIE Tuanjie, LIN Xianwei, HU Lianxin, et al. Fabric defect detection system based on improved YOLOv5 algorithm [J]. Cotton Textile Technology, 2022, 50(11):15-20.

- [13] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression [C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: Association for the Advancement of Artificial Intelligence, 2020.
- [14] 解晓君, 刘瑞璞, 陈果. 女性职业套装 3 种格式考议[J]. 服装学报, 2019, 4(5): 430-434, 451.
XIE Xiaojun, LIU Ruipu, CHEN Guo. Textual research on three formats of women's professional suits [J]. Journal of Clothing Research, 2019, 4 (5): 430-434, 451.
- [15] 陈国强, 李倩倩. 职业装色彩研究[J]. 毛纺科技, 2017, 45(3): 52-55.
CHEN Guoqiang, LI Qianqian. Research of uniform color [J]. Wool Textile Journal, 2017, 45(3): 52-55.
- [16] 郭磊, 王邱龙, 薛伟, 等. 基于改进 YOLOv5 的小目标检测算法[J]. 电子科技大学学报, 2022, 51(2): 251-258.
GUO Lei, WANG Qiulong, XUE Wei, et al. A small object detection algorithm based on improved YOLOv5 [J]. Journal of University of Electronic Science and Technology of China, 2022, 51(2): 251-258.
- [17] 李雪, 吴圣明, 马丽丽, 等. 基于 Ghost 卷积和 YOLOv5s 网络的服装检测[J]. 计算机系统应用, 2022, 31(7): 203-209.
LI Xue, WU Shengming, MA Lili, et al. Clothes detection using ghost convolution and YOLOv5s network [J]. Computer Systems & Applications, 2022, 31(7): 203-209.