

基于改进 YOLOv5 的指纹二级特征检测方法

高梦婷¹, 孙晗², 唐云祁^{1*}, 杨智雄¹

¹中国人民公安大学侦查学院, 北京 100038;

²江苏省公安厅刑警总队, 江苏 南京 210000

摘要 马德里指纹错案的出现使得现行指纹鉴定体系不断受到挑战和质疑。以指纹二级特征的统计规律为基础的量化鉴定技术成为了新的研究难点与热点,而指纹二级特征的自动检测与分类是实现指纹二级特征自动统计的基础。因此,提出一种基于 YOLOv5 的指纹二级特征检测方法。首先,建立指纹二级特征数据集,共包含 4000 张带标注的指纹图像;其次,根据指纹二级特征点尺寸小且分布密集的特点,对 YOLOv5 网络结构进行改进,删除原有的 32 倍下采样大目标特征检测层,添加新的微小特征融合层;再使用 Feature Pyramid Networks (FPN)、Pyramid Attention Network (PAN) 和 Spatial Pyramid Pooling (SPP) 结构通过融合多种尺度的方式实现局部特征和全局特征提取;最后,添加 Squeeze-and-Excitation (SE) 通道注意力机制模块,有效增强模型的鲁棒性和密集小目标的检测能力。实验结果表明,相比于原模型,改进后 YOLOv5s_FI 模型,在检测速度基本不变的情况下,平均精度均值 (mAP@0.5) 从 93.0% 提高到 97.4%,且权重缩减了 3/4。

关键词 图像处理; 目标检测; 指纹特征识别; YOLOv5; 注意力机制

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP213375

Fingerprint Second-Order Minutiae Detection Method Based on Improved YOLOv5

Gao Mengting¹, Sun Han², Tang Yunqi^{1*}, Yang Zhixiong¹

¹School of Investigation, People's Public Security University of China, Beijing 100038, China;

²Jiangsu Provincial Criminal Police Corps, Nanjing 210000, Jiangsu, China

Abstract The fingerprint identification system has been challenged and questioned following the erroneous fingerprint individualization in the Madrid train bombings case. Therefore, the quantitative identification technology based on the statistical law of fingerprint secondary features is now a prevalent research topic, for which the automatic detection and classification of fingerprint second-order minutiae serve as foundations. In this paper, a YOLOv5 based fingerprint second-order minutiae detection method was proposed. First, a fingerprint second-order minutiae dataset was established, which contained 4000 fingerprint images with annotations. The structure of the YOLOv5 network was improved based on the characteristics of small size and dense distribution of fingerprint second-order minutiae. More specifically, the original feature detection layer of 32 times down-sampled large target was deleted, and a new micro-scale detection layer was added. Feature Pyramid Networks (FPN), Pyramid Attention Network (PAN), and Spatial Pyramid Pooling (SPP) structures were used to extract local and global features through multi-scale fusion. Finally, the Squeeze-and-Excitation (SE) channel attentional mechanism was added to effectively enhance the robustness of the model and the detection ability of dense small targets. The experimental results reveal that compared with the original model, the mean average precision (mAP@0.5) value of the improved YOLOv5s_FI model increases from 93.0% to 97.4% under the condition that the detection speed is basically unchanged, and the weight of the improved YOLOv5s_FI model is reduced by three quarters.

Key words image processing; object detection; fingerprint minutiae identification; YOLOv5; attentional mechanism

1 引言

指纹是人体唯一性极强的生物特征,是公安机关

打击刑事犯罪的重要依托。依据作案人遗留在犯罪现场的指纹,经过指纹查询比对与检验鉴定,可确定作案人身份,进而为司法审判提供证据支撑。1896年,阿

收稿日期: 2021-12-28; 修回日期: 2022-01-29; 录用日期: 2022-02-14; 网络首发日期: 2022-02-24

基金项目: 中央高校基本科研业务费项目(2021JKF203)、上海市现场物证重点实验室开放课题基金资助(2021XCWZK04)

通信作者: *tangyunqi@ppsuc.edu.cn

根廷警方最早使用指纹证据进行人身认定。为解决鉴定标准问题,1911年,法国法庭科学家Edmond Locard最早提出进行人身同一认定必须满足12个特征匹配点,但并未对匹配点数目进行科学解释^[1]。1974年,Osterburg通过论证指出2枚指纹出现8处相同特征但却不属同一人的概率约为十万亿分之一,然而并未正面回答指纹鉴定匹配点数量的标准问题^[2]。指纹鉴定匹配点数量标准的缺失,使得各个国家关于指纹鉴定特征符合点的最低数目规定并不相同,少则8个、多则16个,鉴定结论也仅限于定性,分为“同一”“不同一”和“不满足鉴定条件”^[3]这3种情况。

近年来,指纹错案的出现使得现行指纹鉴定体系不断受到挑战和质疑。最知名的指纹错案是2004年马德里电车爆炸案,该案中警方在现场提取到一枚残缺指纹,依据这枚现场指纹美国警方错误地将布兰登·梅菲尔德认定为作案人。2014年,美国迈阿密警局对指纹错误鉴定进行统计,结果表明指纹鉴定伪真率为3.0%、伪否率为7.5%^[4],这说明定性的指纹鉴定结论并不可靠。

司法错案的出现,使得司法审判对法庭证据的准确性、可靠性和科学性提出了更高的要求^[5],指纹证据的检验评价也开始向以量化评价为核心的似然比框架模式转变^[6]。指纹鉴定所依据的二级特征分别是端点(ridge ending)、分叉点(bifurcation)、小勾(spur)、小桥(crossover)、小点(island)、小棒(independent ridge)和小眼(lake),其中端点可细分为起点和终点,分叉点可细分为分歧点和结合点。而这些二级指纹特征分布并不均衡,端点和分叉点最为普遍,其鉴定价值远低于其他各类。要实现指纹鉴定结论的量化,就必须统计各类指纹二级特征点的分布规律,但目前并无基于大数据的统计结果。现有指纹识别技术仅能将二级指纹特征简化为带方向的点线特征,无法对上述7类二级特征进行精确识别。

本文从指纹鉴定量化角度出发,基于YOLOv5研究指纹二级特征自动检测方法,为自动统计指纹二级特征分布规律奠定技术基础。首先建立指纹二级特征数据集,进而针对指纹二级特征目标尺寸小、分布密集且多有重叠的特点,在原模型YOLOv5的基础之上,多次实验选取最佳检测层,融合浅层特征与深层语义特征,并添加注意力机制模块,实现指纹二级特征准确识别和精准定位。

2 相关工作

随着深度学习和卷积神经网络不断发展,很多研究者将其应用于目标检测领域,旨在解决目标的定位和分类问题。目标检测方法可分为2类:两阶段(Two stage)和单阶段(One stage)^[7]。两阶段目标检测先定位、后进行分类,代表算法有R-CNN^[8]、Fast R-CNN^[9]和R-FCN^[10]等。单阶段目标检测将目标检测视为一

个回归问题,同时进行定位和分类,代表算法有YOLO^[11]、SSD^[12]和RetinaNet^[13]等。相比于两阶段,单阶段检测方法精度稍低,但检测速度较快,可满足实时检测需要,应用更加广泛。

在目标检测各种模型中,YOLO系列最具代表性,是性能高且普适性强的目标检测模型。YOLO是由Redmon等^[11]于2016年提出的,使用端到端的网络结构,同时完成目标和分类这2个任务,缺点是检测精确度不高。之后,Redmon等^[14-15]分别于2017年和2018年提出后续版本YOLO900和YOLOv3,在保持高检测速度的同时提高检测精度。2020年,文献^[16]基于YOLOv3提出了YOLOv4,使目标检测在性能普通的GPU(如RTX1080Ti)进行训练成为可能。同年,Ultralytics^[17]提出了YOLOv5,该模型检测准确度高于以往的两阶段目标检测模型,且检测速度快,可以很好地应用于嵌入式设备和移动端进行检测。因此,YOLOv5成为目前目标检测表现最好的网络模型之一。

随着人工智能的发展,深度学习也逐渐应用在指纹识别领域,例如:李伯男等^[18]提出一种基于胶囊网络的指纹纹型分类网络Cap-FingerNet;Song等^[19]用深度卷积神经网络来学习表示指纹局部脊结构,提出一种新的指纹聚合方法,以提高检索效率。但关于指纹特征点提取检测相关的研究很少,如Tang等^[20]提出一种基于卷积神经网络模型的复杂背景下指纹特征提取网络,但并未区分二级特征的种类,具有一定的局限性。

为实现指纹鉴定量化评估,将YOLOv5算法应用到指纹识别领域,实现5类指纹二级特征的检测与定位,为日后建立以数据为基础的鉴定概率评估方法打下基础。

3 所提方法

3.1 YOLOv5简介

YOLOv5共包含4种不同整深度宽度的模型,由C3模块中参数区分,网络由浅至深分别为YOLOv5s、YOLOv5m、YOLOv5l和YOLOv5x,目标检测性能也依次递增,网络使用更加灵活多变,以满足不同检测需求。YOLOv5网络结构如图1所示,由输入(input)、主干(backbone)、颈部(neck)和输出(output)这4部分组成。

1)输入部分。对输入图片尺寸批量归一化。使用Mosaic数据增强,通过对4张图片进行随机的旋转、翻转和缩放再拼接成一张图片作为训练数据,提升模型训练速度。并使用自适应锚框(anchor)计算方法,每次训练时,自动根据所使用的数据集,用聚类算法计算出最佳的一组锚框值。

2)主干网络部分。由Focus、4次卷积Conv、3个C3模块和1个Spatial Pyramid Pooling(SPP)结构构成,从输入图像中提取不同大小的特征图(feature map)。Focus

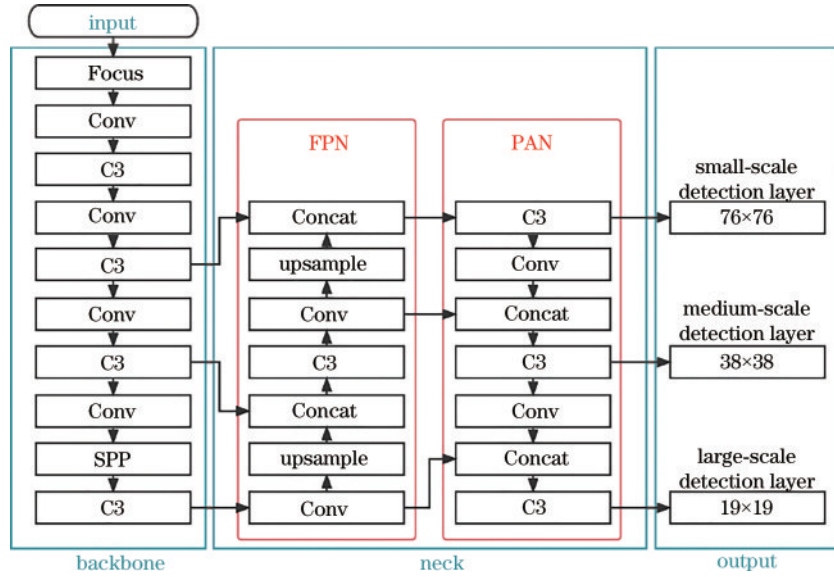


图 1 YOLOv5网络结构

Fig. 1 Network structure of YOLOv5

通过切片操作对输入图片进行裁剪后堆叠,对输入图片进行一次下采样操作。C3模块则是在YOLOv4主干网络 Cross Stage Partial Network(CSP)模块的基础上加以改进,由 Conv、batchnorm 和 SiLU 激活函数这 3 部分组成的。延续使用 SPP 模块进行特征融合,SPP 模块结构如图 2 所示。通过 3 种尺度的池化,将任意大小的特征图固定为相同长度的特征向量,传输给全连接层,实现多重感受野融合。

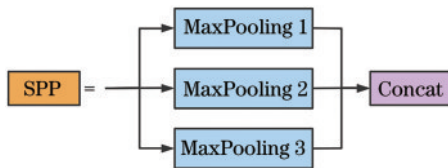


图 2 SPP 结构

Fig. 2 Structure of SPP

3) 颈部网络部分。采用 Feature Pyramid Networks (FPN)^[21] 和 Pyramid Attention Network (PAN)^[22] 结构融合不同层次的特征图。FPN 由下至上把深层的语义特征传到浅层,从而增强多个尺度上的语义表达;而 PAN 则相反,由上至下把浅层的定位信息传导到深层,增强多个尺度上的定位能力。这种结构共同增强了颈部网络的特征融合能力,获得更多的上下文信息,减少信息丢失。

4) 输出部分。经过 8 倍下采样、16 倍下采样和 32 倍下采样,网络输出端共生成了 3 个特征图。特征图越小说明每个网格单元对应的图像区域就越大,19x19 的输出适用于检测大尺寸目标,而 76x76 的输出适用于检测小尺寸目标。YOLOv5 中采用 CIOU_Loss 作为 Bounding box 的损失函数。根据这些新的特征映射,网络输出端进行检测和分类。

3.2 注意力机制

近年来,注意力机制被广泛使用在计算机视觉、自然语言处理等各种深度学习任务中,取得了不少突破性的进展,成为神经网络研究的一大热点。其中,比较有代表性的是 Hu 等^[23]提出的 Squeeze-and-Excitation (SE) 注意力机制模块和 Woo 等^[24]提出的 Convolutional Block Attention Module(CBAM)。

SE 模块是针对通道的注意力机制,包括压缩(squeeze)和激励(excitation)这 2 部分。压缩部分就是对输入进行全局平均池化(global average pooling),输入尺寸为 $W \times H \times C$ 时,特征图池化后输出大小为 $1 \times 1 \times C$ 的向量。激励部分是由 2 个全连接层组成的,为减少通道个数和参数量,添加 R_{SE} 缩放参数。第 1 个全连接层神经元个数为 $C \times R_{SE}$,输出为 $1 \times 1 \times C \times R_{SE}$;第 2 个全连接层神经元个数为 C ,输出为 $1 \times 1 \times C$ 。Scale 操作将 SE 模块计算出来的各通道的权值分别和原输入 $W \times H \times C$ 对应通道相乘,完成在通道维度上的原始特征重标定。SE 模块结构如图 3(a)所示。

CBAM 先后从通道(channel)和空间(spatial)这 2 个维度提取具有意义的注意力特征^[25]。CBAM 通道注意力和 SE 模块大致相同,区别在于 CBAM 在通道压缩采取了最大池化和全局平均池化。CBAM 空间注意力结构如图 3(b)所示。通道注意力模块的输出作为空间注意力模块的输入 $W \times H \times C$,再次采用最大池化和全局平均池化,得到 2 个 $W \times H \times 1$ 的特征图。经过 7x7 卷积核卷积和 Scale 操作得到双重注意力机制调整过的特征图。CBAM 通道注意力机制 M_c 和空间注意力机制 M_s 公式分别如下:

$$M_c(\mathbf{F}) = \sigma \left\{ \text{MLP} \left[\text{AvgPool}(\mathbf{F}) \right] + \text{MLP} \left[\text{MaxPool}(\mathbf{F}) \right] \right\} = \mathbf{F}', \quad (1)$$

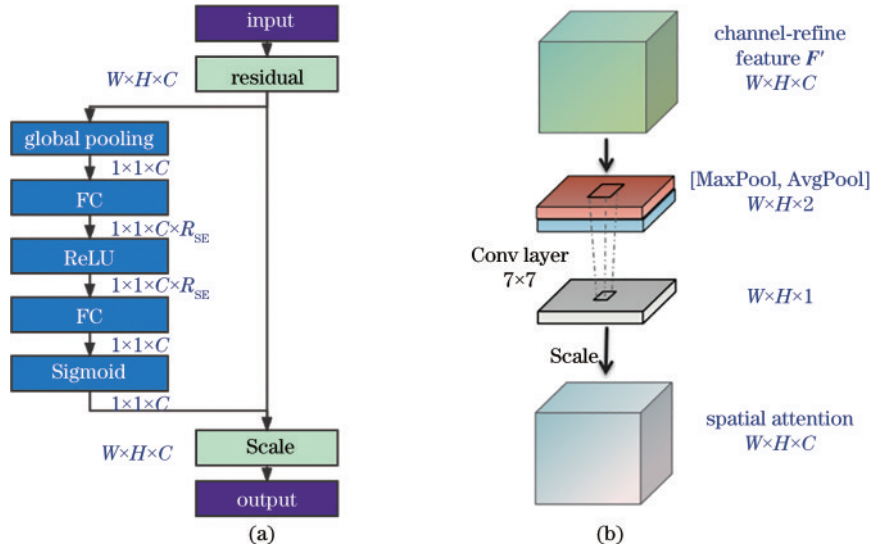


图 3 注意力机制模块结构。(a) SE 模块结构；(b) CBAM 中空间注意力机制结构

Fig. 3 Structure of attention mechanism. (a) SE model structure; (b) spatial attention model structure in CBAM

$$M_s(\mathbf{F}') = \sigma \left\{ f^{7 \times 7} \left[\text{AvgPool}(\mathbf{F}'); \text{MaxPool}(\mathbf{F}') \right] \right\}, \quad (2)$$

式中： \mathbf{F}' 表示通道注意力机制输入特征图； σ 表示 Sigmoid 激活函数；MLP 表示多层感知器参数； \mathbf{F}' 表示通道注意力的输出，也是空间注意力的输入； $f^{7 \times 7}$ 表示卷积层使用 7×7 的卷积核。

3.3 面向指纹二级特征检测的 YOLOv5 模型改进

指纹图像中二级特征尺寸微小、数量多、分布密集

且多有重叠，直接应用 YOLOv5 等目标检测方法来检测二级特征结果并不理想。为了解决这些问题，对 YOLOv5 原网络结构进行了 3 处改进：1) 删除 32 倍下采样大尺寸特征融合层，添加 4 倍下采样的微小特征融合层；2) 将 FPN 和 PAN 结构迁移到剪枝之后的网络中，并选取合适的 SPP 池化核参数；3) 添加 SE 注意力机制模块。改进后的 YOLOv5_FI 网络结构如图 4 所示。

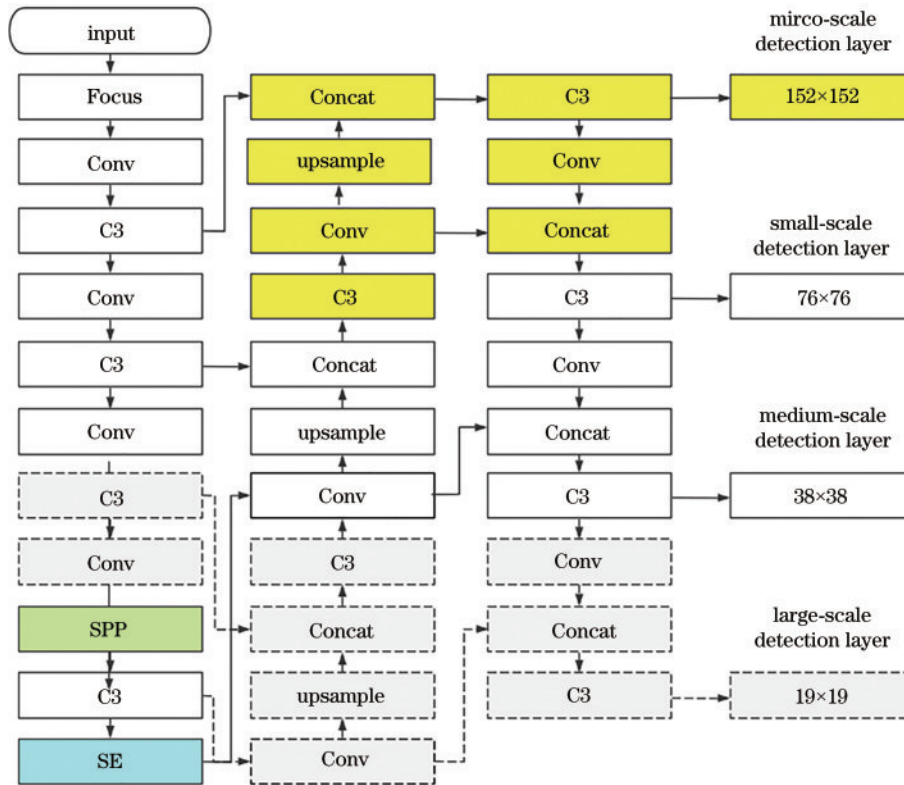


图 4 改进后 YOLOv5_FI 网络结构

Fig. 4 Improved YOLOv5_FI network structure

首先,为了提高YOLOv5检测小尺寸目标的性能,添加了一个新的微小特征融合层。该融合层由主干网络4倍下采样输出,后与8倍下采样特征图融合,生成一个尺寸为 152×152 的特征图,分割的网格更加细密,有助于微小目标检测识别。因为指纹图片中并无大尺寸二级特征,删除原结构32倍下采样大尺寸特征融合层及其对应的主干和颈部网络结构(虚线部分),能够极大地减小算法复杂度和模型的参数量。

其次,为了提高模型对密集重叠目标的检测性能,将FPN和PAN结构迁移到剪枝之后的网络中,FPN结构将强语义特征从上至下传递,PAN结构则是从下到上融合,增强多个尺度上的定位能力。延续使用SPP模块,将特征图从3个不同尺度进行最大池化,有效增加主干特征的接收范围,实现多重感受野融合,有利于检测目标尺寸差异大和重叠的情况。

最后,由于指纹二级特征数量分布并不均衡,在主干和颈部网络之间添加SE注意力机制模块。通道注

意力机制中每个特征图代表一个特征通道,有助于筛选原图中更具有意义的部分特征,聚焦于特征通道权重分配。

4 实验及结果分析

4.1 指纹数据集构建

4.1.1 数据集的采集

数据集的质量对目标检测算法设计和训练有极大的影响^[26]。由于目前开源指纹数据集图像质量不高、部分指纹残缺,实验数据来自于公安实战指纹图像,尺寸为 680×680 ,共500张。指纹完整、纹线清晰,有利于后续的预处理,减少误差。

4.1.2 数据集预处理

直接使用深度学习对指纹图像进行降维和特征提取,会极大影响实验精度,需要对指纹图像进行预处理^[27]。目前针对指纹的预处理方法已经相对完善,本实验预处理步骤分别是背景分离、计算局部脊线方向、纹线增强和二值化,如图5所示。

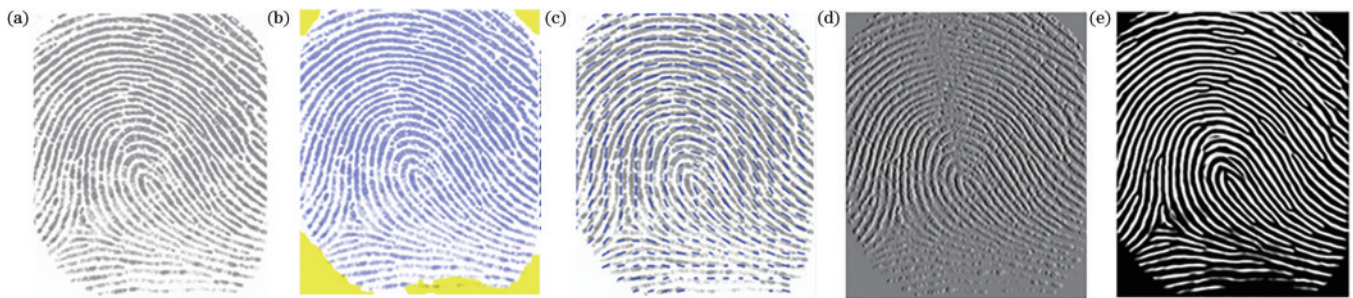


图5 指纹图片预处理。(a)原图;(b)背景分割;(c)局部脊线方向;(d)纹线增强;(e)二值化

Fig. 5 Pre-processing of fingerprint image. (a) Original image; (b) background segmentation; (c) local ridge direction; (d) ridge enhancement; (e) binarization

4.1.3 数据集的标注

使用LabelImg软件制作检测标签,对预处理后的500张指纹图片中二级特征进行手工标注,尽可能准确地框选特征点的全貌,避免框住无关纹线。标签格

式设置为PASCAL VOC数据集的格式,共标注5类二级特征,标签名称分别为bifurcation(标签0)、spur(标签1)、independent ridge(标签2)、lake(标签3)和crossover(标签4)。图像标注情况如图6所示。



图6 图像标注情况

Fig. 6 Image annotation

4.2 实验环境及参数配置

实验平台操作系统为Intel(R) Xeon(R) CPU E5-

1650 v3@3.50 GHz,内存为16 GB,GPU为NVIDIA GeForce GTX2080Ti,显存为11 GB。软件配置为

Windows 10, CUDA 1.2 GPU 并行计算库, 深度学习框架为 Pytorch1.9.0 版本。

本次实验训练和测试中图片设置为 640×640 JPG 格式, batch_size 设置为 36, 整个训练过程为 400 个 epoch, 选取平均精度均值 (mAP0.5, mAP0.5:0.95)、权重大小 (weight) 和实际检测速度 (FPS) 作为模型评价指标进行对比。

由于 YOLOv5 模型预设的超参数是根据 COCO 数据集优化而来的, 不具有普适性。因此, 使用超参数进化 (hyperparameter evolution) 得到更适用于本数据

集的超参数数值。超参数进化算法使用遗传算法根据评价指标等调节优化超参数, 将训练的过程重复 300 个 epoch, 得到初始学习率 (lr0) 为 0.0128, 循环学习率 (lrf) 为 0.256, SGD 学习率动量 (momentum) 为 0.905。

考虑到数据集样本不充足和人工标注的时间成本问题, 为增加样本的多样性对数据集进行数据增强。已知指纹图像的翻转和旋转并不影响特征点的种类和数量, 在竖直方向对每张图片进行翻转扩增至 2 倍, 然后分别顺时针旋转 90°、180° 和 270°, 扩增效果如图 7 所示。



图 7 数据集扩增效果

Fig. 7 Dataset amplification effect

经过以上步骤, 数据集扩增至 8 倍, 共 4000 张图片, 共制作标签 119768 个。5 类二级特征数量及在训练集和测试集分布情况如表 1 所示。

表 1 各类目标在训练集和测试集的分布情况

Table 1 Distribution of various targets in training set and testing set

Object	Bifurcation	Crossover	Independent ridge	Lake	Spur
Traning set	74985	2451	17811	3584	8948
Testing set	8319	285	1981	408	996

Mosaic 数据增强是 YOLOv5 的一大亮点, 通过对 4 张图片进行随机的旋转、翻转和缩放再拼接到一张图片作为训练数据。对于本数据集, 数据增强时部分图片已经过旋转、翻转, 使用 Mosaic 数据增强会造成过拟合, 而且缩放拼接的方式并不利于指纹特征点小目标检测, 因此本实验并未使用 Mosaic 数据增强。

4.3 实验结果与分析

4.3.1 YOLOv5 基础模型结果对比

YOLOv5 目标检测网络结构共有 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 等 4 个模型, 网络深度和宽度逐次递增。本实验将 YOLOv5s 针对本数据集超参数进化后的模型命名为 YOLOv5s_A。在自行组建并标注的指纹特征数据集对 YOLOv5 基础模型算法对比测试, 各项指标如表 2 所示。

YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 结构复杂度依次递增, 参数越多, 相应训练产生模型的权重越大, 训练的时间越长。在公开数据集上测试结果显示, YOLOv5 结构越复杂, 深度越深, 检测效果越好^[28]。然而对于指纹特征点目标检测却出现了不同的结果, 经分析, 过深的网络和过多的卷积操作不适合检

表 2 YOLOv5 基础模型各项性能指标

Table 2 Various performance indexes of YOLOv5 basic model

Model	mAP0.5 / %	mAP0.5:0.95 / %	Weight / 10 ⁶
YOLOv5s	93.0	57.1	14.8
YOLOv5m	90.3	56.6	42.9
YOLOv5l	85.7	52.1	93.8
YOLOv5x	83.8	53.5	175.2
YOLOv5s_A	94.7	57.6	14.8

测细密且小的指纹特征点。由表 2 可知, 相比于 YOLOv5s, 超参数进化后的 YOLOv5s_A 模型 mAP0.5 提高了 1.7 个百分点, 有效增强了检测性能, 后续实验均使用超参数进化后数值。

4.3.2 网络检测层对检测性能的影响

YOLOv5s 原结构共有 3 个检测层, 分别经历了主干网络 8 倍、16 倍、32 倍下采样, 输出特征图大小对应为 76×76、38×38、19×19, 分别实现小、中、大尺寸的目标检测。为了探究不同深度网络检测层对检测性能的影响, 本实验由浅至深选取了 8 倍、16 倍、32 倍下采样层作为检测层, 分布对应 YOLOv5s_8、YOLOv5s_16、YOLOv5s_32 这 3 个模型, 其结构如图 8 所示。

通过表 3 可知, 8 倍下采样作为检测层的 YOLOv5s_8 模型性能最佳, mAP0.5 为 67.8%, 16 倍下采样次之, 32 倍下采样精度显著降低。经分析, 当主干网络下采样层数较浅时, 提取的主要是较低的空间特征, 有助于小目标的检测。指纹特征点尺寸较小且分布密集, 若主干网络过深, 获取更深层次的语义特征, 反而就会丢失细节特征, 导致小目标漏检率大大提升、精度大幅下降, 而且会增加实验计算量。因此, 本次实验删减掉 32 倍下采样层, 保留 8 倍和 16 倍下采样检测层。

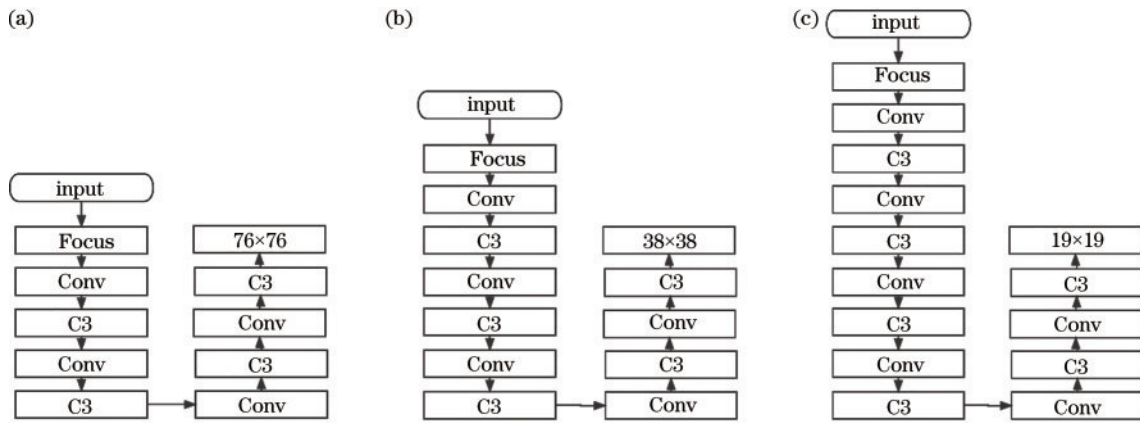


图 8 模型结构。(a) YOLOv5s_8; (b) YOLOv5s_16; (c) YOLOv5s_32

Fig. 8 Network structure. (a) YOLOv5s_8; (b) YOLOv5s_16; (c) YOLOv5s_32

表 3 不同深度检测层性能对比

Table 3 Performance comparison of detection layers at different depths

Model	mAP0.5 / %	mAP0.5:0.95 / %	Weight / 10^6
YOLOv5s_8	67.8	33.0	1.4
YOLOv5s_16	66.9	31.8	2.7
YOLOv5s_32	31.3	13.0	8.1

4.3.4 特征融合对检测性能的影响

在确定主干网络深度之后,在 YOLOv5_16 的基础之上添加 SPP 金字塔池化模块,颈部使用 FPN 和 PAN 结构将 8 倍与 16 倍下采样层进行特征融合,将模型命名为 YOLOv5s_B。为探究 SPP 最优池化效果,实验了 4 种常见 SPP 池化内核:(3,5,7)、(5,7,9)、(7,9,13) 和 (9,11,13),分别命名为 YOLOv5s_B_a、YOLOv5s_B_b、YOLOv5s_B_c 和 YOLOv5s_B_d。其结构如图 9 所示。

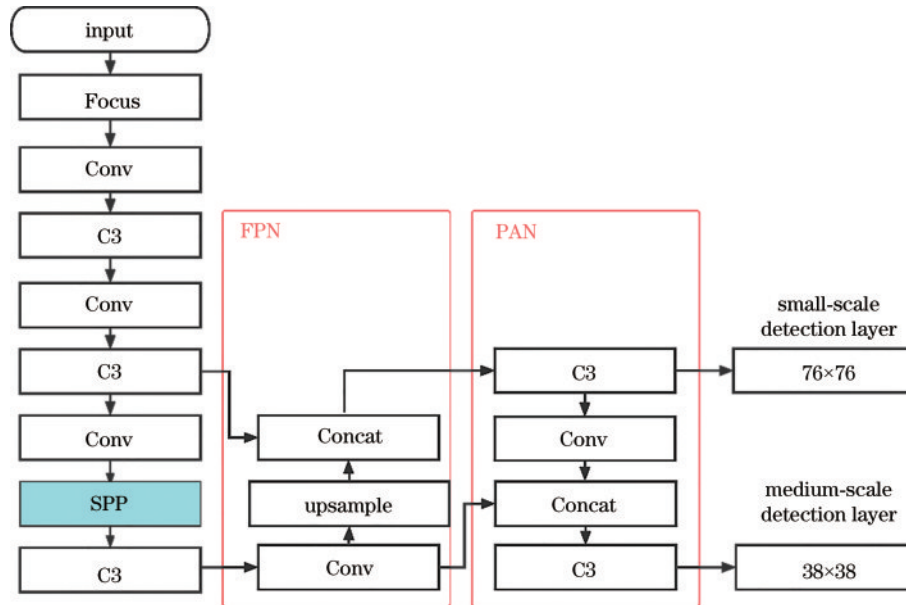


图 9 YOLOv5s_B 网络结构

Fig. 9 Network structure of YOLOv5s_B

由实验可知,SPP 池化核为(7,9,13)的 YOLOv5s_B_c 检测性能最佳,mAP0.5 为 93.7%,相比 YOLOv5s_16 提升了 24.3%,模型权重仅增加 0.7×10^6 ,如表 4 所示。为何特征融合之后模型检测性能会有如此明显的提升?经研究,SPP 模块分别使用 3 个不同尺度的最大池化进行处理,更加有效地增加了主干特征接收范围,实现多重感受野融合,这有效,弥补了删掉

32 倍下采样所损失的深层语义信息,并且多层次提取特征还增强了网络的鲁棒性。FPN 结构将强语义特征从顶部特征映射传递到下部特征映射,在多尺度上进行预测。同时,PAN 结构是从下到上的融合,将强定位特征从较低的特征映射传递到较高的特征映射,增强多个尺度上的定位能力。通过多种尺度的融合方式实现了局部特征和全局特征提取,增强了网络的表

表 4 不同 SPP 池化核性能指标对比

Table 4 Comparison of performance indexes of different SPP pooled nuclei

Model	mAP0.5 / %	mAP0.5:0.95 / %	Weight / 10^6
YOLOv5s_16	66.9	31.8	2.7
YOLOv5s_B_a	90.8	54.1	3.4
YOLOv5s_B_b	92.4	55.6	3.4
YOLOv5s_B_c	93.7	57.7	3.4
YOLOv5s_B_d	91.2	54.9	3.4

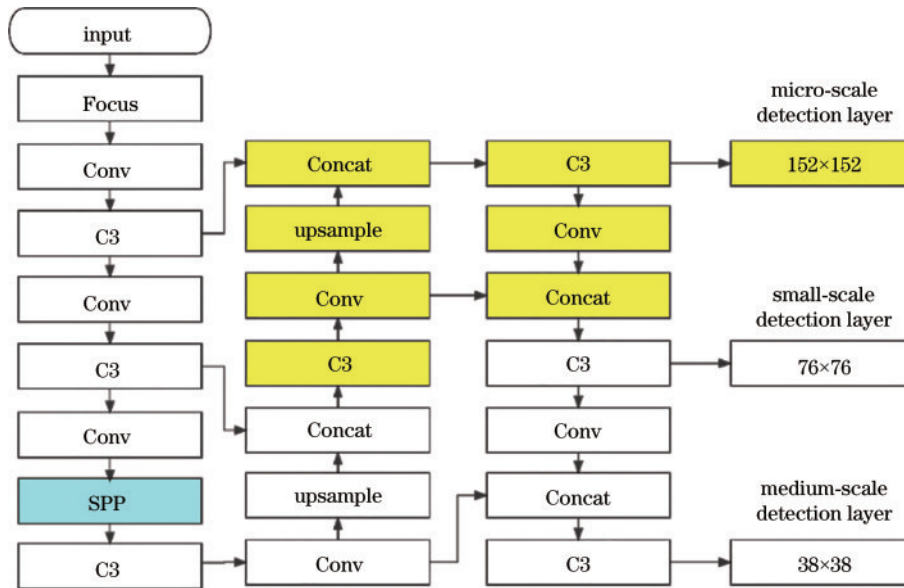


图 10 YOLOv5s_C 网络结构

Fig. 10 Network structure of YOLOv5s_C

由实验结果可知,添加微小尺度检测层后, mAP0.5 为 95.2%, mAP0.5:0.95 提升 1 个百分点,不可避免地导致模型权重小幅增加,如表 5 所示。新增 4 倍下采样检测层使检测网络结构更加广泛和细致,通过提取较低的空间特征并将其与深层语义特征融合生成特征图,适用于检测指纹图像中微小、重叠的目标。

表 5 添加微小尺度检测层后性能对比

Table 5 Performance comparison after adding micro-scale detection layer

Model	mAP0.5 / %	mAP0.5:0.95 / %	Weight / 10^6
YOLOv5s_B_c	93.7	57.7	3.4
YOLOv5s_C	95.2	58.7	4.0

4.3.6 添加注意力机制对检测性能的影响

注意力机制是一种通过调整不同的权重来强制学习过程聚焦于输入对象的重要通道和区域的方法,为探究添加注意力机制是否可以优化检测性能,将注意力机制模块 CBAM 和 SE 依次添加在 YOLOv5s_C 的主干网络和颈部之间,分别将模型命名为 YOLOv5s_

达能力,有利于检测目标尺寸差异大和特征重叠的情况。之后的实验将基于 YOLOv5s_B_c 模型进行改进。

4.3.5 添加微小尺度检测层对检测性能的影响

由于指纹特征点尺寸小且分布密集,而 YOLOv5 的小尺度 76×76 检测层对指纹特征点的适用性不足,本实验添加了一个新的微小尺度检测层(micro-scale detection layer)。该检测层经过 4 倍下采样输出,该模型命名为 YOLOv5s_C,结构如图 10 所示。

CBAM 和 YOLOv5s_SE,结构如图 11(a)所示。

由表 6 可知,添加 SE 注意力模块检测效果最佳, mAP0.5 为 97.3%,与 YOLOv5_C 模型相比提升了 1.4 个百分点,且 mAP0.5:0.95 提升了 3.2 个百分点,权重只是略有增多。通道注意力机制的每个特征图代表一个特征通道,有助于筛选原图具有意义的部分特征。空间注意力机制每个特征图中一个像素代表原图中某个区域的特征,有助于训练网络应该注意原图中哪个区域的特征。SE 仅关注通道权重分配, CBAM 既考虑不同通道像素的重要性,又考虑同一通道不同位置像素的重要性。但为何精度却不如 SE 注意力机制? 经分析有 2 个方面原因:1)由图 11(b)、(c)可知,指纹特征点在指纹图像中心及上半部分分布较多,添加空间注意力机制后,网络更加关注图像中心及上半部分,导致对图像其他空间注意力下降,漏检其余空间的目标;2)由图 11(a)可知,图像经过池化核为 (7,9,13) 的 SPP 模块池化之后,传入注意力机制,空间注意力机制使用了 7×7 的卷积核进行卷积操作、最大池化和全局平均池化的通道压缩方式, SPP 和 CBAM 冗余的重复多次卷积造成特征图有用信息的损失,从而影响检测效果。因此,针对指纹二

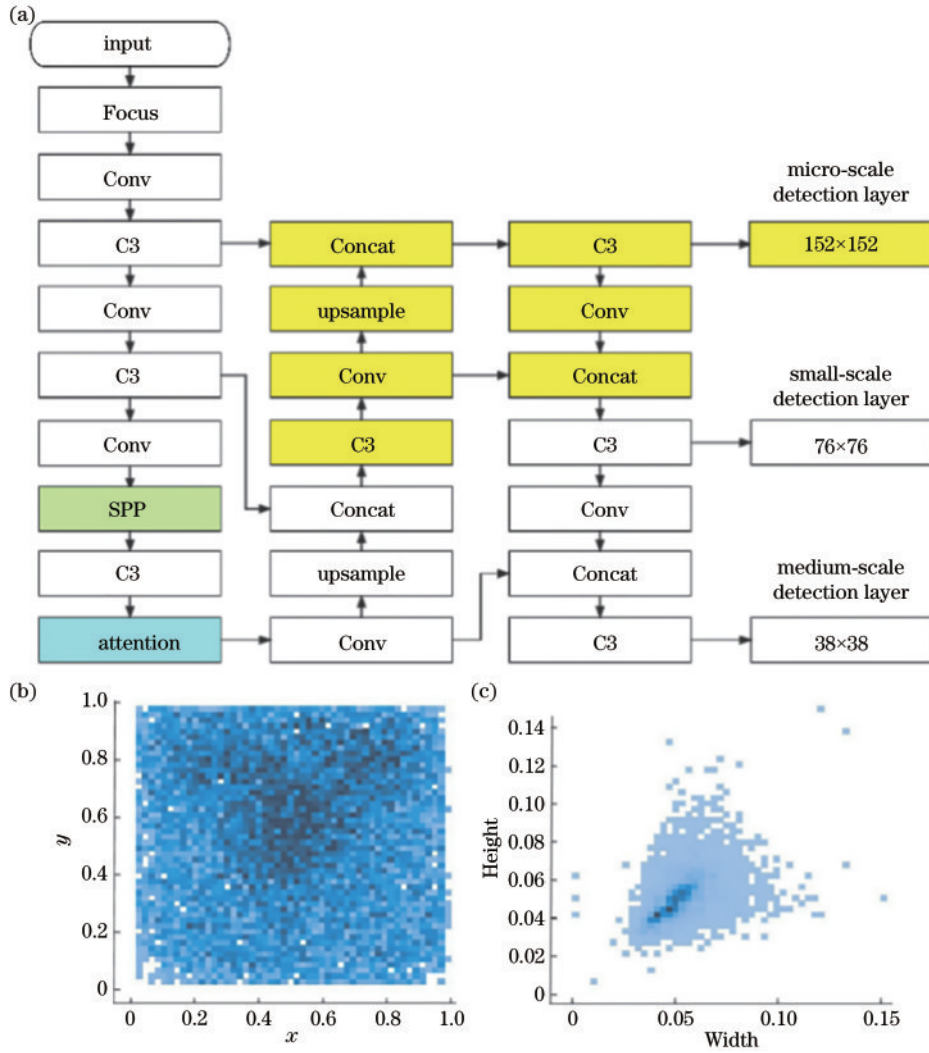


图 11 YOLOv5s_Attention 网络及数据集相关情况。(a) YOLOv5s_Attention 网络结构；(b) 数据集目标位置分布情况；(c) 数据集目标尺寸比例情况

Fig. 11 YOLOv5s_Attention network and dataset related information. (a) Network structure of YOLOv5s_Attention; (b) distribution of target locations in dataset; (c) proportion of target size of dataset

表 6 添加注意力机制性能对比

Table 6 Performance comparison of add attention mechanism

Model	mAP0.5 / %	mAP0.5:0.95 / %	Weight / 10 ⁶
YOLOv5_C	95.2	58.7	4.0
YOLOv5_CBAM	94.7	58.5	4.1
YOLOv5_SE	97.4	61.9	4.1

级特征检测，单一的通道注意力机制 SE 模块比 CBAM 检测效果更佳。

由实验数据可知，YOLOv5s_SE 表现最佳，将模型命名为 YOLOv5s_Fingerprints Identification (YOLOv5s_FI)。

4.3.7 改进模型与其他检测模型性能对比

将 YOLOv5s_FI 与目前性能优异的几种算法进行比较，并在表 7 中对各项性能指标进行对比。

由表 7 可知，针对指纹二级特征点识别，所提

表 7 不同检测算法性能对比

Table 7 Performance comparison of various detection algorithms

Model	Prediction / %	Recall / %	mAP0.5 / %	Weight / 10 ⁶	FPS
SSD	10.8	59.9	21.2	94.9	21.8
YOLOv4	80.4	46.6	60.6	250.3	32.2
YOLOv5s	93.6	87.0	93.0	14.8	27.3
YOLOv5s_FI	96.5	94.0	97.4	4.1	26.6

YOLOv5s_FI 模型在准确率、召回率、mAP0.5 上均优于经典算法 SSD、YOLOv4、YOLOv5s。YOLOv5s_FI 的 FPS 略低于 YOLOv5s，但模型训练权重仅为 4.1×10^6 ，相当于 YOLOv5s 模型的 1/4。YOLOv5s_FI 在综合性能上取得了较好的表现。

4.3.8 改进模型效果与性能评估

改进的 YOLOv5s_FI 基础模型在 400 个 epoch 训练时的目标检测损失 (obj_loss) 曲线和分类损失

(cls_loss)曲线如图 12 所示。

YOLOv5s_FI 与 YOLOv5s 模型在验证集上检测

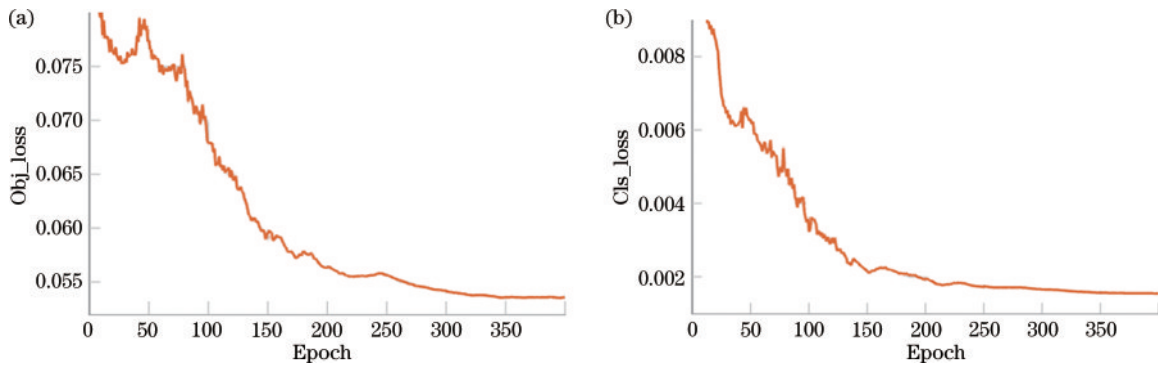


图 12 YOLOv5s_FI 的训练情况。(a) 目标损失变化曲线; (b) 分类损失变化曲线

Fig. 12 Training of YOLOv5s_FI. (a) Obj_loss change curve; (b) cls_loss change curve



图 13 检测结果对比。(a) YOLOv5s_FI; (b) YOLOv5s

Fig. 13 Comparison of detection results. (a) YOLOv5s_FI; (b) YOLOv5s

5 结 论

针对目前指纹鉴定量化评估的难点,为实现指纹 5 类二级特征点数理统计,建立了 5 类指纹二级特征数据集。基于此数据集进行训练和对比实验,并对 YOLOv5s 算法进行改进,删除 32 倍下采样检测层,添加 4 倍下采样微小特征融合层,有效地获取更多指纹图像的微小特征信息。使用 FPN、PAN 和 SPP 结构,通过多种尺度的融合方式实现了局部特征和全局特征提取,增强了网络的表达能力。此外,还添加 SE 通道注意力机制模块,减少了无用特征信息对模型的干扰,增强重要特征的通道权重,从而提高检测的效果。实验结果表明,所提 YOLOv5s_FI 算法的 mAP0.5 精度达到 97.4%,在检测速度基本不变的情况下,模型权重较 YOLOv5s 缩减了 72.3%,有效增加了模型的鲁棒性和密集小目标的检测能力,实现了 5 类指纹二级特征点的准确识别和定位。

参 考 文 献

[1] 谭铁君. 指纹证据的量化评价模式[J]. 刑事技术, 2020,

45(6): 616-621.

Tan T J. On quantitative evaluation of fingerprint evidence[J]. Forensic Science and Technology, 2020, 45(6): 616-621.

- [2] Pankanti S, Prabhakar S, Jain A K. On the individuality of fingerprints[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(8): 1010-1025.
- [3] Smith A M, Neal T M S. The distinction between discriminability and reliability in forensic science[J]. Science & Justice, 2021, 61(4): 319-331.
- [4] 刘世权, 罗亚平, 蔡伟思, 等. 指纹鉴定能力测试的鉴定结果准确性问题研究[J]. 中国人民公安大学学报(自然科学版), 2015, 21(1): 20-24.
- Liu S Q, Luo Y P, Cai W S, et al. Research on the accuracy of fingerprint identification ability test[J]. Journal of People's Public Security University of China (Science and Technology), 2015, 21(1): 20-24.
- [5] Garrett B L, Gardner B O, Murphy E, et al. Judges and forensic science education: a national survey[J]. Forensic Science International, 2021, 321: 110714.
- [6] Horsman G. The different types of reports produced in digital forensic investigations[J]. Science & Justice, 2021, 61(5): 627-634.

- [7] 杨智雄, 唐云祁, 张家钧, 等. 基于改进YOLOv4的行人鞋部检测算法[J]. 激光与光电子学进展, 2022, 59(8): 0810007.
Yang Z X, Tang Y Q, Zhang J J, et al. Detection algorithm of pedestrian shoe area based on improved YOLOv4[J]. Laser & Optoelectronics Progress, 2022, 59(8): 0810007.
- [8] 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, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 580-587.
- [9] Girshick R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1440-1448.
- [10] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN [C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2980-2988.
- [11] 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, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [12] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[EB/OL]. (2015-12-08) [2021-05-04]. <https://arxiv.org/abs/1512.02325>.
- [13] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2999-3007.
- [14] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [15] Redmon J, Farhadi A. Yolov3: an incremental improvement[EB/OL]. (2018-04-08) [2021-05-04]. <https://arxiv.org/abs/1804.02767>.
- [16] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2021-04-05]. <https://arxiv.org/abs/2004.10934>.
- [17] Ultralytics. YOLOv5[EB/OL]. (2020-06-03) [2021-04-15]. <https://github.com/ultralytics/yolov5>.
- [18] 李伯男, 赵彤, 吴敏. 基于胶囊网络的海量指纹纹型精准分类算法[J]. 中国科学院大学学报, 2020, 37(3): 387-397.
Li B N, Zhao T, Wu M. Fast exact classification algorithm of massive fingerprint patterns based on capsule network[J]. Journal of University of Chinese Academy of Sciences, 2020, 37(3): 387-397.
- [19] Song D H, Tang Y, Feng J F. Aggregating minutia-centred deep convolutional features for fingerprint indexing[J]. Pattern Recognition, 2019, 88: 397-408.
- [20] Tang Y, Gao F, Feng J F, et al. FingerNet: an unified deep network for fingerprint minutiae extraction[C]//2017 IEEE International Joint Conference on Biometrics, October 1-4, 2017, Denver, CO, USA. New York: IEEE Press, 2018: 108-116.
- [21] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [22] Li H C, Xiong P F, An J, et al. Pyramid Attention Network for Semantic Segmentation[EB/OL]. (2018-05-25) [2021-04-05]. <https://arxiv.org/abs/1805.10180v1>.
- [23] Hu J, Shen L, Sun G, et al. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 2011-2023.
- [24] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [25] Wang K, Liu M Z. YOLOv3-MT: a YOLOv3 using multi-target tracking for vehicle visual detection[J]. Applied Intelligence, 2022, 52(2): 2070-2091.
- [26] 冉蓉, 徐兴华, 邱少华, 等. 基于深度卷积神经网络的裂纹检测方法综述[J]. 计算机工程与应用, 2021, 57(9): 23-35.
Ran R, Xu X H, Qiu S H, et al. Review of crack detection methods based on deep convolutional neural networks[J]. Computer Engineering and Applications, 2021, 57(9): 23-35.
- [27] Liu W, Zhou C, Yan P L, et al. Efficient preprocessing algorithms of fingerprint images[J]. WSEAS Transactions on Information Science and Applications, 2006, 3(6): 1021-1027.
- [28] 汤寓麟, 边少锋, 翟国君, 等. 侧扫声纳检测沉船目标的改进YOLOv5法[J/OL]. 武汉大学学报(信息科学版): 1-11 [2021-10-28]. <https://doi.org/10.13203/j.whugis20210353>.
Tang Y L, Bian S F, Zhai G J, et al. Improved YOLOv5 method for shipwreck detection by side scan sonar[J/OL]. Geomatics and Information Science of Wuhan University: 1-11 [2021-10-28]. <https://doi.org/10.13203/j.whugis20210353>.