

DOI: 10.13382/j.jemi.B2306625

基于百度飞桨的面向黑暗环境人员 行为检测与身份识别*

杜 闯 何贇泽 邓海平 常 珊 王耀南
(湖南大学电气与信息工程学院 长沙 410006)

摘 要:针对传统可见光在黑暗环境中难以实现人员行为检测与身份识别的问题,本文结合红外热成像技术基于百度飞桨深度学习框架研究了一种面向黑暗环境的人员行为检测与身份识别算法。首先经过实地采集,自主构建红外热成像人员行为数据集总计 10 900 张 9 种行为类别以及双光人脸数据集总计 3 000 张 30 位人员。针对行为检测方面,基于轻量化网络 PP-LCNet 改进 YOLOv5 骨干网络进行人员行为检测,大幅度减少模型参数并提高检测精度与推理速度。针对人脸识别方面,引入 CycleGAN 算法改进 InsightFace 实现将红外人脸转化为可见光人脸进行身份识别,提高在黑暗环境下人脸识别准确率。最后实现红外人员行为检测网络与人脸识别网络的级联工作,在黑暗环境下可以实时行为检测与身份识别,具有很好的应用效果。实验结果表明,基于 PPLCNet 轻量化改进的 YOLOv5 相对于原网络模型参数减少 56.4%,平均精度 mAP 由 89.1%提高至 94.7%,推理速度由 68 提高至 101 fps;基于 CycleGAN 算法改进 InsightFace 相对于原网络黑暗环境下识别准确率由 84%提高至 99%。

关键词: 黑暗环境;红外热成像;行为检测;跨模态人脸识别

中图分类号: TP183; TN29 **文献标识码:** A **国家标准学科分类代码:** 520.20

Human behavior detection and identification in dark environment based on Baidu Paddle

Du Chuang He Yunze Deng Haiping Chang Shan Wang Yaonan
(College of Electrical and Information Engineering, Hunan University, Changsha 410006, China)

Abstract: Aiming at the problem that traditional visible light is difficult to realize personnel behavior detection and identity recognition in dark environment, this paper combined with infrared thermal imaging technology to study an algorithm for personnel behavior detection and identity recognition in dark environment based on Baidu Paddle deep learning framework. First, after field collection, the behavioral dataset of infrared thermal imaging personnel totaled 10 900 pieces of 9 behavior categories and the double-light face dataset totaled 3 000 pieces of 30 personnel. In terms of behavior detection, the lightweight network PP-LCNet is used to improve the YOLOv5 backbone network for personnel behavior detection, reducing model parameters greatly and improving detection accuracy and reasoning speed. In terms of face recognition, CycleGAN algorithm is introduced to improve InsightFace to transform infrared faces into visible faces for identity recognition and improve face recognition accuracy in dark environments. Finally, the cascade of infrared human behavior detection network and face recognition network is realized, and real-time behavior detection and identity recognition can be achieved in the dark environment, which has a good application effect. The experimental results show that compared with the original network model, the parameters of YOLOv5 based on PPLCNet are reduced by 56.4%, the average precision mAP is increased from 89.1% to 94.7%, and the reasoning speed is increased from 68 to 101 fps. Based on CycleGAN algorithm, the recognition accuracy of InsightFace is improved from 84% to 99% in the dark environment of the original network.

Keywords: dark environment; infrared thermal imaging; behavior detection; cross-modal face recognition

0 引言

随着深度学习技术的快速发展,人工智能技术已经广泛应用于计算机视觉、自然语言处理、语音识别等领域。身份信息识别和行为姿态检测是计算机视觉领域的重要应用,其能够确定目标的身份并分析出目标是否有异常或危险行为,在安防、交通、生产、医疗等众多领域有广泛应用^[1]。但基于传统可见光的人员行为检测与身份识别算法已渐渐无法满足日益复杂的应用环境,而且易受到复杂环境和恶劣天气的影响,不能在夜间黑暗的环境中继续工作。红外热成像技术可以在黑暗环境中被动成像,不易受到光照及环境因素干扰并能昼夜连续工作^[2]。因此在面对黑暗环境检测困难的问题上研究基于红外热成像的人员行为检测与身份识别技术具有重要的研究意义与应用价值^[3]。

针对行为检测方面,有研究人员提出一种增强型轻量级水上目标检测网络实现了基于红外热成像水域场景人员的 3 种简单行为检测^[4]。在交通领域有学者提出一种改进的 MobileNet 网络基于红外热成像设备实现夜间行人行为识别^[5],但仅适用于单目标无法解决多目标多种行为同时检测的问题。在安全监护领域有学者提出一种基于 K 最近邻算法通过检测环境中的温度分布和变化情况识别人体行为^[6],但其仅适用于简单环境无法应用于复杂环境。针对人脸识别方面,有研究人员提出了一种基于图像全部信息进行样本特征矩阵构造的压缩感知红外人脸识别方法^[7],但其仅能识别样本库内已存在的人脸,无法进行大规模的红外人脸识别应用。还有学者提出基于特征向量规整化的红外人脸识别方法^[8],通过改变红外人脸特征在特征空间的分布来提高红外人脸识别的准确率,但其对有表情以及角度变化的红外人脸识别效果不理想。因此本文提出一种基于轻量化网络改进 YOLOv5 的人员行为检测网络与一种基于 CycleGAN 算法改进 InsightFace 的跨模态人脸识别算法,将行为检测与身份识别级联,实现在黑暗环境下实时的行为检测与身份识别。

1 行为检测与人脸识别概述

1.1 行为检测算法

行为检测算法一般主要分为 3 类,分别是图像帧检测,视频检测以及骨骼关键点行为检测。图像帧行为检测主要采用目标检测算法,可以同时实现目标定位与行为分类,识别速度快且精度高。目前主流目标检测算法主要分为 One-Stage 与 Two-Stage 算法,前者不需要提取候选区域,直接预测目标的位置坐标和类概率,例如

YOLOv3-5^[9-10]等算法。后者包含两个阶段,首先提取目标可能存在的候选区域,然后对该区域进行更精确的分类与定位,例如 Faster R-CNN^[11]等算法。视频行为检测主要使用视频分类网络识别一段时间内的人员行为,例如 SlowFast^[12]等算法。但由于模型的复杂度较高并且部署到移动端困难,所以实时检测性能较差。骨骼关键点行为检测是通过人体关键点一段时间内变化程度实现行为检测识别,例如 STGCN^[13]等算法。但该算法受环境信息影响较大,只适用于简单环境和少数人员行为检测。

1.2 人脸识别算法

人脸识别算法近些年发展迅速,2014 年人脸识别的奠基之作 DeepFace^[14]提出了采用人脸检测、对齐、提取以及分类的过程进行识别。中国科学院于 2016 年提出的 center loss^[15]使每一类样本在中心位置都非常接近,并将 Softmax+center loss 结合起来进行训练,得到了较好的效果,且信息损失较小。2018 年,腾讯 AI 将 Softmax loss 通过对特征和权重向量进行 L2 正则化,得到了 CosFace^[16]。同年英国伦敦帝国学院提出在 AmSoftmax^[17]和 CosFace 的基础上设计得到具有区分性的损失函数 ArcFace^[18]。这些人脸识别算法在人脸识别数据集 LFW 上都已经达到了很高的识别精度,但在实际无约束的场景场景中,光照、模糊、遮挡、姿态以及表情等干扰因素都会降低图像的质量,影响识别精度,随着年龄的增长和体重的变化人脸图像还会出现较大差异,增加了识别的难度。现有方法大多过度追求准确率,忽略了识别准确率与效率之间的平衡。且现有方法大都基于光照良好的环境下训练得到,在黑暗环境可见光成像无法清晰地捕捉人脸特征的情况下,识别效果会大打折扣。

2 数据集

2.1 红外人员行为数据集

本文建立的红外人员行为数据集采用巨哥 MAG62 热像仪进行数据采集,该设备支持以图片或视频形式保存数据,采集地点选取某公司工厂车间并对多名工作人员进行行为采集。数据集共计 11 000 张,其中共包含 9 种行为类别,分别是玩手机(playphone)、打电话(call)、打架(fight)、吸烟(smoke)、持危险品(danger)、躺倒(lie)、操作(operate)、分拣(pick)以及处理(process),其中前 6 种属于异常或危险行为,后 3 种属于正常工作行为。红外图像分辨率为 640×512,所有图像数据经视频抽帧、去重、筛选标注后得到,标注采用 Labelimg 标注工

具进行处理。数据集构成如表 1 所示。

表 1 红外人员行为数据集

Table 1 Infrared human behavior dataset			
类别	标签名	图片数	标签数
玩手机	playphone	1 840	3 381
打电话	call	485	523
打架	fight	802	802
吸烟	smoke	1 779	2 051
持危险品	danger	1 157	1 163
躺倒	lie	425	425
操作	operate	263	263
分拣	pick	360	360
处理	process	243	243

红外人员行为数据集种图像示例如图 1 所示,红外图像可以在光线不好或者复杂环境中清晰显示工作人员的各种行为。

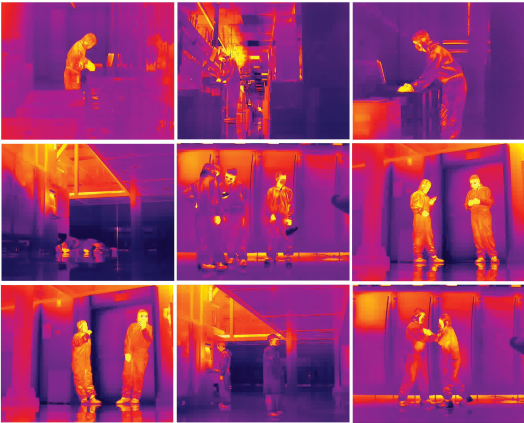


图 1 红外人员行为数据集样本展示

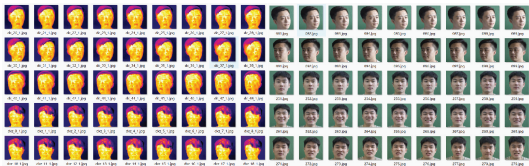
Fig. 1 Display of infrared human behavior dataset

2.2 双光人脸数据集

本实验需要对红外人脸进行跨模态转换,所以需采集红外与可见光两种模态的人脸图像。实验室先后拍摄了 30 名志愿者的红外和可见光各 1 500 张、总共 3 000 图片,红外图像与可见光图像相对应。在每个人的 50 张图片中包含表情、角度、遮挡的变化。拍摄情境分有眼镜遮挡平行进行,面部表情包括正常、微笑、皱眉、拍摄角度分左半侧脸、右半侧脸、抬头、低头。

在训练之前需要对数据集进行处理,首先进行标注、裁剪,使用 Labelimg 工具对采集的图像进行标注,生成.xml 文件再根据标签文件将人脸裁剪出来。最后进行缩放,保证数据集内的图片大小一致。处理好后的人脸图像的形状大小统一为 512×512,随后将两个不同模态的图像分别放置于两个不同的文件夹,红外图像与可见光图像保持对应关系。双光人脸数据集如图 2 所示,用于训练模态转换环节中所使用的算

法 CycleGAN,产生能够将红外图像转换成可见光图像的模型。为满足黑暗环境下识别的要求,本实验使用热像仪采集红外图像进行识别,因此本次实验的初始输入图像为红外图像,数据集中的可见光图像仅用于训练算法模型,中间生成的可见光图像最终输入到人脸识别算法进行识别。



(a) 红外人脸数据集 (b) 可见光人脸数据集
(a) Infrared face dataset (b) Visible face dataset

图 2 双光人脸数据集

Fig. 2 Dual light face dataset

由于本实验的目的是实现黑暗环境下的身份识别,因此还需要采集黑暗环境下的人脸,来进行实验的验证与对比。对前面采集双光人脸数据集的 30 名志愿者,在黑暗环境下再拍摄一组可见光图像,要求与前面一致,保证 3 种情况下的人脸相对应。黑暗环境下拍摄图像组成黑暗环境人脸数据集,如图 3 所示。

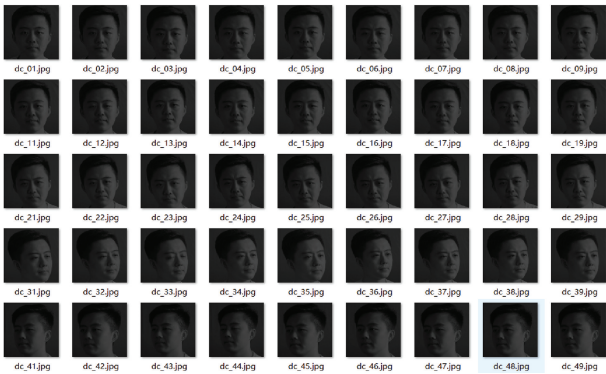


图 3 黑暗环境下的人脸数据集

Fig. 3 Face dataset in dark environment

3 黑暗环境人员行为检测与身份识别

3.1 基于 PP-LCNet 改进的 YOLOv5 算法

人员的行为多种多样而且具有短时性和不规律性,所以行为检测识别对算法的实时性和可靠性具有很高的要求。近年来 YOLO 系列目标检测算法发展迅速,其中 YOLOv5 是目前应用最广,使用频率最高的网络之一。因此本文选择 YOLOv5 网络作为 YOLO 算法代表进行相应的研究和改进。YOLOv5 的网络结构主要分为骨干网络(backbone)、增强特征提取网络(Neck)和检测层(Head)。原骨干网络采用 CSPDarknet53,结合残差网络

Residual 和跨阶段局部网络 CSPNet 构成 BottleneckCSP 模块。增强特征提取网络包括特征金字塔网络 FPN 和路径聚合网络 PAN。FPN 是自顶向下将高层特征通过上采样与低层特征进行融合, PAN 在其基础上自底向上进一步融合增强。检测层采用 YOLO Head 可以预测目标的位置、类别和概率。

综合考虑行为检测与身份识别的实时性与模型投入实际使用的部署问题,本文采用轻量级网络 PP-LCNet^[19]对 YOLOv5 进行改进。PP-LCNet 是百度最新提出的基于 Paddle 深度学习框架的轻量级卷积神经网络。该网络使用 MobileNetv1 中提出的深度可分离卷积 DepthSepConv 作为基本块,由深度卷积 DW 和逐点卷积 PW 组成。深度可分离卷积相对于传统卷积它通过分解卷积操作可以显著减少模型参数数量和计算复杂度,在保持类似水平性能的前提下大大降低计算成本。同时引入 H-Swish 激活函数和 SE 注意力机制。H-Swish 可以在不增加推理时间的前提下通过替换原网络的 ReLU 激活函数来获得更好的网络性能,在输入接近于 0 的区域, H-Swish 更加平滑所以在反向传播过程中梯度计算更加准确从而有利于网络的优化。而 ReLU 函数在输入小于 0 的情况下会完全抑制梯度传播,导致训练过程中可能会出现梯度消失问题。同时 H-Swish 函数引入非线性阈值,可以更好地模拟神经元激活过程。与 ReLU 函数相比, H-Swish 函数在整个输入范围内都能产生非零输出值,这有助于网络在处理复杂数据时更好地捕捉到非线性特征。其表达式为:

$$H-Swish(x) = x \frac{\text{ReLU6}(x+3)}{6} \quad (1)$$

SE 模块由全局平均池化层、两个全连接层、激活函数 ReLU 和 Sigmoid 组成, ReLU 函数在计算通道的重要性权重时可以确保权重非负, Sigmoid 函数则可以将权重约束在 0~1 的范围内表示通道的相对重要性, 通过结合 ReLU 和 Sigmoid 满足 SE 模块的设计需求, 从而能够有效地计算通道注意力并提高在网络性能。将 SE 模块放置在最后两个 DepthSepConv 层中可以在不增加推理时间的情况下有效提高网络的特征提取能力。本文采用 PP-LCNet 网络轻量化改进后的网络结构如图 4 所示。

该改进后网络输出 3 个特征层的特征图的尺寸分别为 $80 \times 80 \times (C+4+1)$ 、 $40 \times 40 \times (C+4+1)$ 和 $20 \times 20 \times (C+4+1)$, 即特征图深度为 C 个类别参数, 4 个位置参数和 1 个置信度参数。

3.2 结合 CycleGAN 改进的 InsightFace 算法

人脸识别算法发展迅速, 新的算法不断出现, InsightFace 是目前为止识别精度最高、使用较为方便的算法之一。InsightFace 算法对于光照条件良好的环

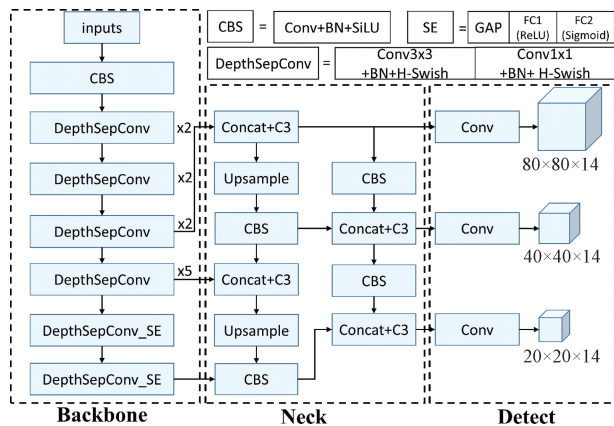


图 4 轻量化改进 YOLOv5 网络结构

Fig. 4 Lightweight improved YOLOv5 network structure diagram

境下的可见光人脸识别能达到非常高的准确率, 使用多种网络结构在 LFW 上实验均达到了 99.5% 以上准确率。然而对于黑暗环境下的可见光人脸识别难度较高, 很难清晰的识别人脸特征。因此对于黑暗环境下的人脸, 本文采取的办法为使用热像仪采集识别对象的红外图像, 红外图像不受光照的影响, 在黑暗环境下也能清晰的成像。

但红外图像是基于人体热辐射成像, 其人脸特征较为模糊很难分辨, InsightFace 算法无法准确的识别。因此需要对采集到的红外图像进行模态上的处理, 使其与可见光图像建立联系。考虑到身份识别的实时性以及数据集的处理难度, 本文采用基于生成对抗网络的风格迁移算法 CycleGAN^[20]来对 InsightFace 进行改进。改进的 InsightFace 算法结构图如图 5 所示, 在原来的 InsightFace 网络的输入端, 添加 CycleGAN 的前向转换生成器 GAB, 使输入的红外图像转换成可见光图像, 随后进行后面的可见光人脸识别步骤, 从而间接实现了对红外人脸的识别。改进的算法在输入红外图像之后立即对其进行模态的转换, 实现在保证运算速度和识别精度的情况下, 增加红外人脸识别的功能。

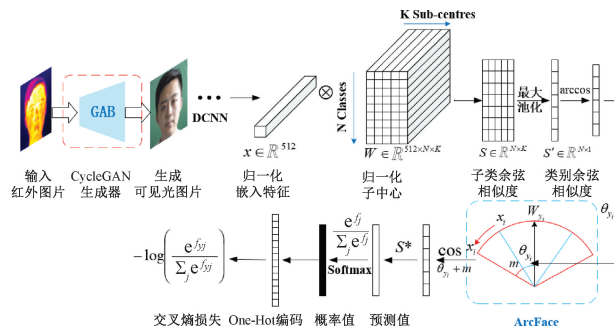


图 5 改进的 InsightFace 网络结构

Fig. 5 Improved InsightFace network structure diagram

3.3 网络级联

基于多线程架构实现人员行为检测算法与身份识别算法二者级联并实时检测识别即线程 1 实现将红外热像仪实时采集数据传输至 PC 端或嵌入式移动端,线程 2 实现将实时采集数据导入轻量化行为检测网络 PPLCNet-YOLOv5 进行行为检测识别,线程 3 实现检测并采集人脸图像数据导入跨模态人脸识别模型进行身份识别。

网络级联整体框架图如图 6 所示。首先通过服务器进行模型训练与算法优化,然后红外热像仪连接服务器实时拍摄红外数据并传输至 PC 端导入行为检测模型,在实时输出行为检测结果的基础上检测并采集人脸图像进行跨模态的身份识别。

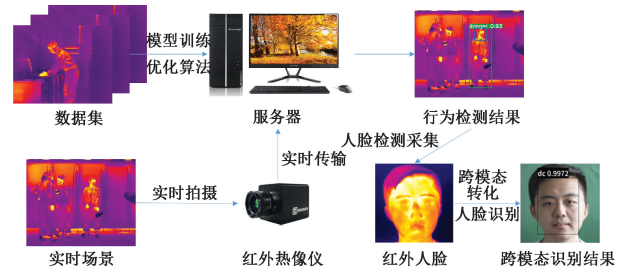


图 6 网络级联
Fig. 6 Network cascade

4 实验与结果分析

本文模型训练及测试实验环境为:CPU 处理器 i9-10900X、GPU 处理器 NVIDIA GeForce RTX3090、软件环境 Pycharm、Python3.8、Cuda11.1、Cudnn8.0、Pytorch1.9.1 及 PaddlePaddle2.4.0。

4.1 行为检测实验

1) 模型训练及测试

将数据集按照 8 : 1 : 1 比例划分为训练集、验证集和测试集。为了更好的证明本文基于 PP-LCNet 轻量化改进 YOLOv5 网络的优越性,在相同的参数设置下对原 YOLOv5 (骨干网络 CSPDarkNet53) 以及 PP-LCNet、MobileNetv3、ShuffleNetv2、GhostNet 和 EfficientNet 等轻量化网络作为 YOLOv5 主干网络进行对照实验。根据各轻量化网络结构定义修改 YOLOv5 网络的配置文件,将轻量化网络逐层替换到原 YOLOv5 网络中然后加载相同的预训练模型权重进行训练和评估。对照实验中训练参数设定均采用原 YOLOv5 网络的默认参数,设置训练参数如下所示:图像大小 640,训练次数 100,批量大小 128,学习率 0.001,优化器 Adam,学习率调整策略 CosineAnnealingLR,损失函数 CIOU 损失。其中

CosineAnnealingLR 学习率调整策略可以使模型在训练后期更加稳定并且具有较好的收敛性能,CIOU 是原 YOLOv5 网络采用的损失函数,它可以更好地衡量目标框之间的距离和重叠程度,相比于传统的 IOU 损失函数能够更准确地指导模型优化从而提高检测精度。训练过程的损失曲线 Loss 和平均精度 mAP 曲线如图 7~8 所示。从图中可以看出本文提出的基于 PP-LCNet 轻量化改进网络相对于原 YOLOv5 网络以及其他轻量化改进网络具有更好的性能表现,随着训练次数的增加网络损失值和平均精度值均优于其他网络。

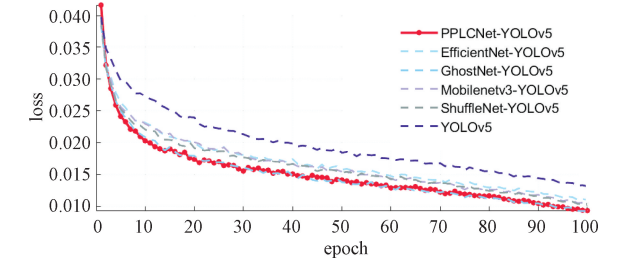


图 7 训练过程 Loss 曲线
Fig. 7 Loss curve of training process

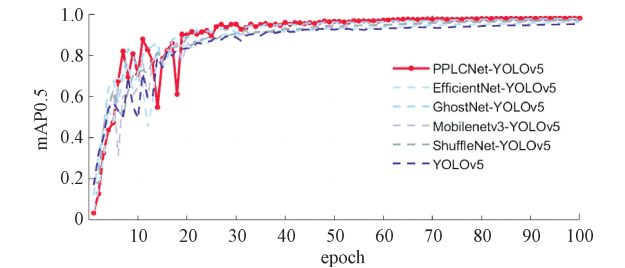


图 8 训练过程 mAP 曲线
Fig. 8 mAP of the training process

2) 实验结果与分析

为了定量评估各检测方法的性能,对测试集测试使用以下指标进行评估:参数量、平均精度 mAP 和检测帧率 FPS。参数量表示模型的复杂程度,模型的参数量越大即计算量越大,消耗内存越多。平均精度 mAP 具体是指在不同召回率下的精确度平均值,衡量模型在所有类别中检测性能。检测帧率 FPS 表示算法每秒可处理的图像帧数,可用于衡量模型的实时检测能力。性能指标对比情况如表 2 所示。

从表 2 中可以看到,与原 YOLOv5 模型相比本文提出的基于 PP-LCNet 网络轻量化改进算法的参数量减少 56.4%而且模型训练时间减少 60.6%,平均精度和推理速度分别提高了 5.6%和 32.2%,具有很好的性能表现。相对于其他轻量化网络改进算法,本系统轻量化改进算法在识别精度、模型复杂度和推理速度方面均优于其他模型,非常适合在以英伟达开发板为硬件平台的环

境实现部署。

表 2 网络性能对比

Table 2 Network performance comparison

网络	参数量/M	mAP@ 0.5	帧率/fps
CSPDarkNet	46.1	89.10	68
PP-LCNet	20.1	94.70	101
MobileNetv3	14.3	92.20	85
ShuffleNetv2	13.8	93.70	97
GhostNet	21.8	93.60	77
EfficientNet	21.9	91.50	83

各类别数据检测精度如图 9 所示,由图可知工作行为处理和操作类别检测精度最高达到 99.5%,而吸烟类别检测精度最低为 80.9%,平均精度达到 94.7%,足以支持实际工程检测需求。

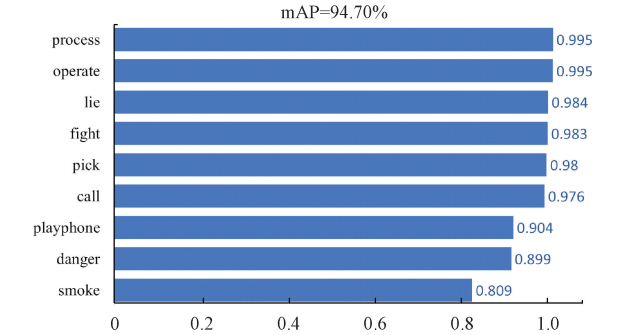


图 9 网络各类检测结果

Fig. 9 Network detection results for each category

对测试集进行测试,识别效果如图 10 所示,从图中可以看出该轻量化改进网络可以对红外热成像人员行为进行检测识别,对红外条件下的人员工作行为,异常及危险行为均有较好的识别效果。通过该实验可以表明基于轻量化网络 PP-LCNet 改进的 YOLOv5 网络可以较好的实现黑暗环境下人员行为检测功能。

4.2 人脸识别实验

1) 模型训练及测试

将双光人脸数据集分别以 8 : 1 : 1 的比例分为训练集、验证集以及测试集,然后用红外人脸图像以及可见光图像分别训练原始的 InsightFace 红外模型和 InsightFace 可见光模型,用来验证原始的 InsightFace 模型在黑暗环境下对红外人脸以及可见光人脸的识别效果。同时为了证明本文利用 CycleGAN 改进后的 InsightFace 在黑暗环境下识别的优越性,需在相同参数下利用可见光人脸数据集训练改进后的 InsightFace 模型。改进后的 InsightFace 模型本质上是识别模态转换后的可见光图像。具体实现方法为将 CycleGAN 网络作为模态转换模块添加至原始 InsightFace 算法的输入端,将输入的红外图像处理成可见光图像后再输入 InsightFace 算法进行识

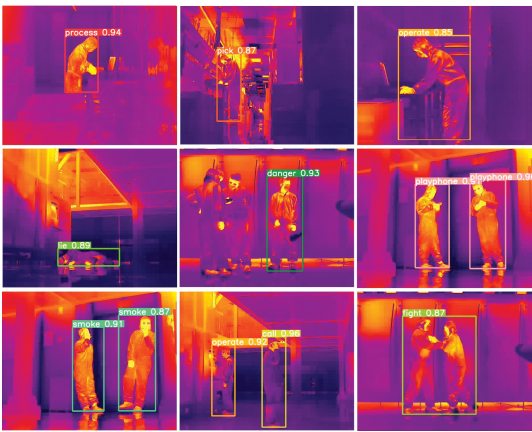


图 10 网络识别效果

Fig. 10 The effect of network recognition

别。CycleGAN 使用双光人脸数据集进行训练,模型训练完成后再添加至 InsightFace 算法,InsightFace 算法则用可见光图像训练即可,训练参数设置为:图片大小 512×512,训练次数 100,批量大小 128,学习率 0.001,优化器为 Adam,分类器 LargeScaleClassifier,损失函数 ArcFace。其中图片过小如 256×256 会导致图片细节丢失,在 CycleGAN 转换环节出现转换后的可见光人脸图像模糊、人脸特征缺少等现象,降低后续人脸识别的准确率;图片过大则会消耗过多计算资源,导致识别过程时间延长,降低实时性。训练批次、批量大小以及学习率的设置为多次重复训练后调优得到的结果。LargeScaleClassifier 分类器具有较强的表达能力和分类性能,可以提高人脸识别模型的准确性。损失函数 ArcFace 可以在特征空间中更好地区分不同别的目标,提高模型的判别能力和分类精度。另外设置 InsightFace 可见光模型识别光照良好的可见光图像实验作为对照组。

2) 实验结果与分析

为了定量比较各个模型的性能,对测试集测试引入如下 4 个参数:真正率 (TPR),假正率 (FPR),错误接受率 (FAR),还有准确率 (Accuracy,以下简称 Acc)。分类算法的结果通常都会用混淆矩阵来表示,其中包括真正例 (TP)、真负例 (TN)、伪正例 (FP)、伪负例 (FN) 4 个类别。而 TPR 表示实际为正,预测成正的比率,计算公式如下:

$$TPR = \frac{TP}{TP + FN} \tag{2}$$

式中:TP 表示实际为正,预测也为正的样本数,FN 表示实际为正,预测为负的样本数。

类似的,FPR 表示实际为负,预测成正的比率,计算如下:

$$FPR = \frac{FP}{FP + TN} \tag{3}$$

式中: FP 为实际为负, 预测为正的样本, TN 为实际为负, 预测为负的样本。 TPR 越接近 1, FPR 越接近 0, 算法性能越好。

Acc 计算公式为:

$$Acc = \frac{TP + TN}{TP + FN + FP + TN}$$

(4)

FAR 表示错误接受比例, 即在比对不同人的图像时把这些的图像误认为是同一个人的图像的比例, FAR 越小越好。

$$FAR = \frac{\text{非同人比较分数大于阈值 } T \text{ 的次数}}{\text{非同人比较的次数}}$$

(5)

设定相同阈值情况下, 不同模型测试过程中得到的各项性能指标如表 3 所示。

表 3 InsightFace 模型性能对比

Table 3 InsightFace Model performance comparison

模型	TPR	FPR	FAR	Acc
InsightFace 红外模型	0.997	0.778	0.001 7	84.7
InsightFace 可见光模型(黑暗)	0.996	0.984	0.001 9	83.7
InsightFace 改进模型	0.998	0.022	0.001 0	98.9
InsightFace 可见光模型(光照)	0.999	0.003	0.001 3	99.8

从表 3 中可以看出, InsightFace 改进后的模型识别精度 Accuracy 相比原始的 InsightFace 模型在红外图像以及黑暗环境下可见光图像上的表现提高了很多, 接近对照组中在光照良好环境下的识别精度。真正率 TPR 越大、假正率 FPR 越小代表分类算法性能越好, 改进的 InsightFace 表现明显优于原始的红外模型以及可见光黑暗模型, 改进后的模型错误接受率 FAR 也降低了很多。

将不同阈值下的 TPR 值和 FPR 值列表作图, 其中 FPR 为横坐标, TPR 为纵坐标, 得到 ROC 曲线, 如图 11 所示。ROC 曲线全称为受试者工作特征曲线, 是根据一系列不同二分类方式, 以真阳性率(TPR)为纵坐标, 假阳性率(FPR)为横坐标绘制的曲线, 常用于评价分类算法的性能。ROC 曲线越接近左上角, 该分类算法的性能越好, 当 $FPR=0, TPR=1$ 时, 分类效果最完美。用曲线与坐标轴包围的面积大小 AUC 来衡量算法的性能, AUC 越大性能越好。由下图可知, 改进后的 InsightFace 算法的 $AUC=0.999$, 大于原始的 InsightFace 算法模型, 因而改进后的 InsightFace 算法性能上得到了很大的提升, 接近在光照良好环境下的识别效果, 能够满足工程环境对黑暗环境下人脸识别的需求。

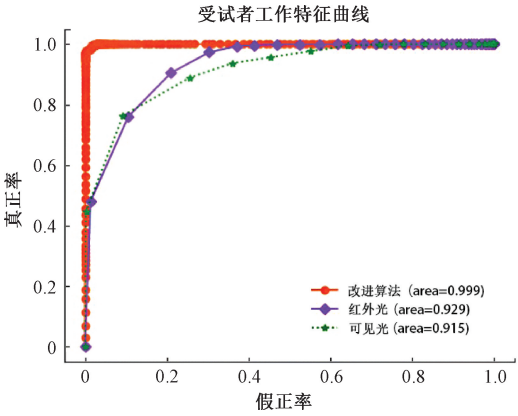


图 11 模型测试 ROC 曲线

Fig. 11 ROC curve for model testing

改进后的算法对测试集测试的效果如图 12 所示, 识别的过程可以拆分为先将红外图像转换成可见光图像, 然后进行人脸识别这两个步骤。改进后的算法成功解决了 InsightFace 在黑暗环境下识别效果不佳的问题, 提高了 InsightFace 对红外人脸识别的准确率。

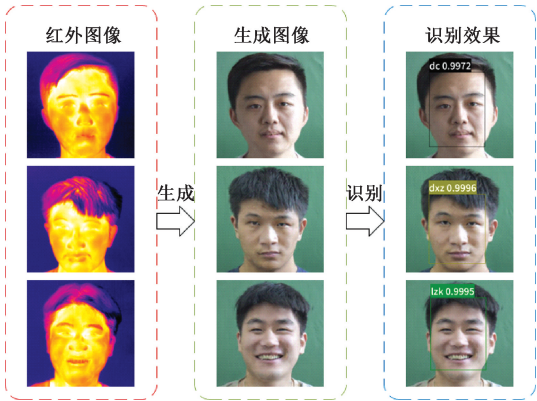


图 12 改进算法识别效果

Fig. 12 Improved algorithm recognition effect

原始的 InsightFace 算法识别速度可达 125 fps, 而基于 CycleGAN 改进后的 InsightFace 算法识别速度因为转换模块的加入有所降低, 为 36 fps, 在较大幅度提升了黑暗环境下人脸识别准确率的基础上仍然满足工业应用的实时性要求。

5 结 论

本文针对黑暗环境下人员行为检测与身份识别的难题, 提出了一种基于 PP-LCNet 网络轻量化改进 YOLOv5 的人员行为检测网络与一种基于 CycleGAN 算法改进 InsightFace 的跨模态人脸识别算法, 并将行为检测与身份识别级联, 实现在黑暗环境下实时的行为检测与身份

识别。基于 PP-LCNet 改进的 YOLOv5 算法相比原始 YOLOv5 模型参数数量减少 56.4% 而且模型训练时间减少 60.6%, 平均精度和推理速度分别提高了 5.6% 和 32.2%, 实现了黑暗环境下人员行为检测功能。基于 CycleGAN 改进的 InsightFace 算法相比原始 InsightFace 模型在黑暗环境下的识别准确率提升 15.2%, 实现了在黑暗环境下的实时身份识别。将二者级联, 同时达到黑暗环境下人员行为检测和身份识别的目的, 具有很好的应用效果。

参考文献

- [1] 彭磊. 工地监控中人脸与行为自动识别[D]. 成都: 电子科技大学, 2020.
PENG L. Automatic recognition of face and behavior in construction site monitoring [D]. Chengdu: University of Electronic Science and Technology, 2020.
- [2] 郑含博, 李金恒, 刘洋, 等. 基于改进 YOLOv3 的电力设备红外目标检测模型[J]. 电工技术学报, 2021, 36(7): 1389-1398.
ZHENG H B, LI J H, LIU Y, et al. The infrared target detection model based on the improvement of YOLOv3 [J]. Journal of Electrician Technology, 2021, 36 (7): 1389-1398.
- [3] 任丹彤, 何赟泽, 刘贤金, 等. 面向智慧工厂的双光融合车间人员行为识别方法[J]. 测控技术, 2022, 41(8): 9-15.
REN D T, HE Y Z, LIU X J, et al. Personnel behavior recognition method of dual light fusion for smart factory[J]. Measurement & Control Technology, 2022, 41 (8): 9-15.
- [4] 党相昭, 何赟泽, 程亮, 等. 面向水域场景的热成像人员识别算法研究[J]. 电子测量与仪器学报, 2022, 36(8): 187-193.
DANG X ZH, HE Y Z, CHENG L, et al. Research on thermal imager recognition algorithm for water scene [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(8): 187-193.
- [5] 卢海燕, 赵红东, 王添盟, 等. 基于轻量级卷积神经网络的红外行人行为识别[J]. 传感器与微系统, 2022, 41(9): 129-131, 135.
LU H Y, ZHAO H D, WANG T M, et al. Infrared pedestrian behavior recognition based on lightweight convolutional neural networks [J]. Sensors and Microsystems, 2022, 41(9): 129-131, 135.
- [6] 王召军, 许志猛, 陈良琴, 等. 基于红外阵列传感器的人体行为识别系统研究[J]. 红外技术, 2020, 42(3): 231-237.
WANG ZH J, XU ZH M, CHEN L Q, et al. Research on human behavior recognition system based on infrared array sensor [J]. Infrared Technology, 2020, 42(3): 231-237.
- [7] 杜梅, 曹蔚然. 基于压缩感知的红外人脸识别[J]. 软件工程, 2019, 22(1): 16-18.
DU M, CAO W R. Infrared face recognition based on compressed sensing [J]. Software Engineering, 2019, 22(1): 16-18.
- [8] 申景欣. 基于特征向量规整化的红外人脸识别研究[D]. 哈尔滨: 哈尔滨理工大学, 2022.
SHEN J X. Research on infrared Face recognition based on feature vector regularization [D]. Harbin: Harbin University of Science and Technology, 2022.
- [9] REDMON J, FARHADI A. YOLOv3: An incremental improvement [EB/OL]. [2018-04-08]. <https://arxiv.org/pdf/1804.02767.pdf>.
- [10] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal speed and accuracy of object detection [EB/OL]. [2020-04-23]. <https://arxiv.org/pdf/2004.10934.pdf>.
- [11] 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 & Machine Intelligence, 2017, 39 (6): 1137-1149.
- [12] FEICHTENHOFER C, FAN H, MALIK J, et al. Slowfast networks for video recognition[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019.
- [13] YAN S, XIONG Y, LIN D, et al. Spatial temporal graph convolutional networks for skeleton-based action recognition[C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2018.
- [14] TAIGMAN Y, YANG M, RANZATO M A, et al. Deepface: Closing the gap to human-level performance in face verification [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1701-1708.
- [15] WEN Y, ZHANG K, LI Z, et al. A discriminative feature learning approach for deep face recognition[C]. European Conference on Computer Vision, 2016.
- [16] WANG H, WANG Y, ZHOU Z, et al. Cosface: Large margin cosine loss for deep face recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [17] WANG F, CHENG J, LIU W, et al. Additive margin softmax for face verification[J]. IEEE Signal Processing

Letters, 2018,25(7):926-930.

[18] DENG J, GUO J, XUE N, et al. Arcface: Additive angular margin loss for deep face recognition [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.

[19] CUI C, GAO T, WEI S, et al. PP-LCNet: A lightweight CPU convolutional neural network[J]. ArXiv Preprint, 2021,ArXiv:2109.15099.

[20] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017.

作者简介



杜闯,2021 年于南京邮电大学获得学士学位,现为湖南大学硕士研究生,主要研究方向为红外机器视觉、深度学习。
E-mail: duchuang_hnu@163.com

Du Chuang received his B. Sc. degree in 2021 from Nanjing University of Posts and Telecommunications. Now he is a M. Sc. candidate in Hunan University. His main research interests include infrared machine vision and deep learning.



何贇泽(通信作者),2006 年于西安交通大学获得学士学位,2008 年于国防科学技术大学获得硕士学位,2012 年于国防科学技术大学获得博士学位,现为湖南大学教授,主要研究方向为嵌入式人工智能与边缘计算、红外热成像与机器视觉。
E-mail: yhe@vip.163.com

He Yunze (Corresponding author) received his B. Sc. degree from Xi'an Jiaotong University in 2006, M. Sc. degree from National University of Defense Technology in 2008, and Ph. D. degree from National University of Defense Technology in 2012, respectively. Now he is the professor of Hunan University. His main research interests include embedded artificial intelligence and edge computing, infrared thermal imaging and machine vision.



邓海平,2023 年于湖南大学获得学士学位,现为湖南大学硕士研究生,主要研究方向为红外机器视觉、深度学习。
E-mail: d19896200948@163.com

Deng Haiping received his B. Sc. degree in 2023 from Hunan University. Now he is a M. Sc. candidate in Hunan University. His main research interests include infrared machine vision and deep learning.