

基于百度飞桨的热成像人员身份与行为识别教学案例开发*

何贇泽 杜闯 常珊 刘敏 王洪金

湖南大学电气与信息工程学院, 湖南 长沙 410082

摘要 大数据时代, 计算机视觉作为一门新兴技术, 对其课程培养体系各高校还在探索建设阶段。当前课程体系中缺少自主可控的热成像人员身份和行为识别教学案例。针对目前身份识别和行为分析技术主要利用可见光信息易受到恶劣天气以及环境影响、基于国外深度学习平台开发等问题, 基于国产百度飞桨深度学习框架, 结合红外热成像技术与深度学习算法, 通过对热像仪二次开发, 提出一种基于热像仪实时识别的目标行为与身份识别技术的教学案例, 在实现行为实时识别的基础上增加人脸检测与采集人脸图像功能并与二级网络人脸识别网络级联进行人脸识别。该教学案例提高了学生的思考动手能力, 教学效果良好, 有利于学生学习其他人工智能有关课程。

关键字 百度飞桨, 红外热成像, 身份识别

Teaching Case Development of Thermal Imaging Personnel Identity and Behavior Identification Based on Baidu Paddle Paddle

He Yunze Du Chuang Chang Shan Liu Min Wang Hongjin

College of Electrical and Information Engineering
Hunan University
Changsha, 410082, China

Abstract—In the era of big data, computer vision, as an emerging technology, is still in the stage of exploration and construction for its curriculum training system. There is a lack of autonomous and controllable teaching cases of thermal imaging personnel identity and behavior recognition in the current curriculum system. In view of the problems that the current identity recognition and behavior analysis technology mainly uses visible light information, is vulnerable to bad weather and environmental impact, and is developed based on foreign deep learning platforms, based on the domestic Baidu PaddlePaddle deep learning framework, combined with infrared thermal imaging technology and deep learning algorithm, through the secondary development of thermal imager, a teaching case of target behavior and identity recognition technology based on real-time recognition of thermal imager is proposed. On the basis of real-time behavior recognition, face detection and face image acquisition functions are added, and face recognition is cascaded with the secondary network face recognition network. This teaching case improves students' thinking and hands-on ability, and the teaching effect is good, which is conducive to students' learning other artificial intelligence related courses.

Keywords—Baidu PaddlePaddle, Infrared thermal imaging, Identity identification

1 引言

目前, 计算机视觉领域发展迅猛, 在日常生活中和特定产业需求中都有着广泛的应用, 呈现出宽领域、多层次、多形式的特点^[1]。全国高校应顺应时代发展需要, 大力推进对“人工智能+X”复合型专业人才的培养^[2], 向社会输送更多具备专业知识和工程实践能力的优质本科生。多目标视觉检测与行为分析技术可以确定目标身份, 检测异常和危险行为, 在防护、安保、边检等众多领域有广泛应用。但是目前该技术主要存在以下不足: 主要针对可见光条件下图像或视频,

对识别环境条件要求高; 主要基于国外深度学习平台开发, 易受限制。因此基于国内深度学习框架利用几乎不受光照影响的红外热成像进行人员身份识别和行为分析成为研究热点, 社会对人员身份和行为识别方向的毕业生需求越来越多。目前进行AI开发框架主要用的是Pytorch^[3], Tensflow^[4]等国外深度学习开发平台。近几年国产的AI框架比如飞桨也在教育界引起重视。因此基于国产百度深度学习框架^[5]开展红外热成像的人员行为与身份的实时识别具有重要的教学价值。

国内高校已经开始探索把百度飞桨开发平台引入到教学实践中。吉林大学的陈玫玫等人^[6]针对目前机器学习教学仍停留在理论方面, 学生无法深入理解现实场景下应用, 提出以百度飞桨 AI Studio 平台为载体的线上线下混合式理论教学模式和基于项目式的实

* **基金资助:** 本文得到湖南省普通高等学校教学改革研究重点项目(HNJG-2021-0028), 湖南省学位与研究生教育改革研究重点项目(2021JGSZ020), 教育部产学研合作协同育人教学内容和课程体系改革项目(202101007008)。

验教学模式。燕山大学的冯建周等人^[7]针对如何结合大数据与人工智能技术的时代特征和机器学习课程的自身特点,设计了一种三阶递进式机器学习教学模式,从课堂、实验、项目层层递进,基于百度飞桨框架介绍具体教学实践过程,最后说明教学效果。吉林化工学院的董如意等人^[1]从机器学习课程的教学方法、教学内容及考核方式三个角度思考,对机器学习课程建设进行了探索,通过引入百度云平台的方式解决计算机算力不足的问题,在讲解经典算法的理论知识同时注重实践结合,并将新技术引入课堂教学,以校企合作的方式,进行项目化教学,巩固学生理论学习成果。马文彬等人^[2]提出CDIO(构思、设计、实现、运作)一体化教学模式选择最具代表性的算法,详细解释其理论从推导到应用实践,尽可能使学员达到融会贯通,增强学习效果。

红外热成像已在物理、化学、生物及信息技术等学科教学中实践应用。如在物理学科中,杨舒茜等人^[8]采用红外热成像技术创新设计表征物质溶解吸放热实验,将利用温度计表达转化为视觉感知明显的颜色变化;在化学学科中,吴梅芬等人^[9]引入智能手机热成像仪观察某些物理、化学过程的热运动和热现象;在生物学科中,龙仲英等人^[10]借助红外热成像在人体、植物和动物方面的应用案例,阐述热成像应用于生物体表面测温、呈现温度分布、实时测量等功能。

据以上调研,目前教学领域几乎没有飞桨结合热成像的人工智能教学案例,所以本课题组联合百度,结合自身在红外领域的科研优势开发相关案例。让学生学习理解计算机视觉的图像分类、目标检测、视频理解和图像转化四大领域的相关知识。通过将百度飞桨和红外热成像结合起来做实验,培养学生对人工智能与红外领域的学习兴趣,培养学生将理论应用于实践的能力。

2 实验目的与任务

多目标视觉检测与行为分析技术作为人工智能重要研究项目方向,其能够确定目标身份,检测异常和危险行为,在防护、安保、智慧生活等众多领域有广泛应用。但当前身份识别和行为分析技术主要是针对合作目标,利用可见光信息,存在着易受到恶劣天气以及环境影响,主要基于国外深度学习平台进行开发,缺少自主可控AI平台等问题。在课题组承担CCF-百度松果基金、百度-教育部产学研合作协同育人项目以及湖南省重点研发计划的基础上,开展基于百度飞桨的热成像人员身份识别与姿态行为分析,本实验的主要任务安排如下。

(1) 本实验以国产百度飞桨Paddle开源框架为基础,基于红外热成像背景,综合采用了飞桨图像识

别套件PaddleClas、飞桨目标检测开发套件PaddleDetection、飞桨视频理解套件PaddleVideo和飞桨生成对抗网络开发套件PaddleGAN,让学生学习理解计算机视觉的图像分类、目标检测、视频理解和图像转化四大领域的相关知识。



(a)



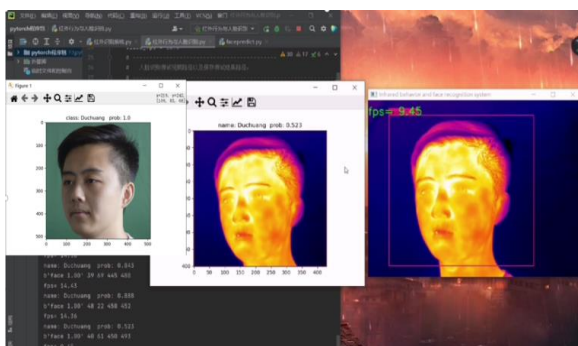
(b)

图 1 热成像行为识别效果示例

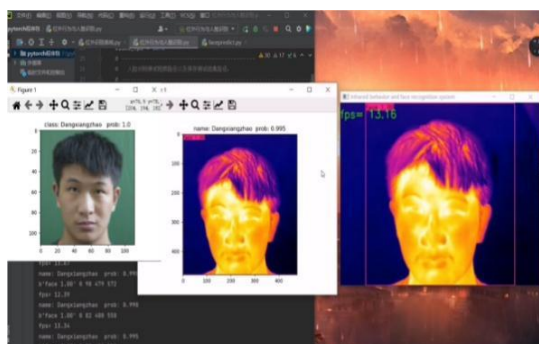
(2) 学生通过该实验应学会使用PaddleClas的图像分类网络如PP-LcNet、PP-HGNet等实现图像分类即人脸识别功能,使用PaddleDetection的目标检测网络如PP-YOLOE^[11]、PP-Picodet^[12]等实现图片级人员定位和行为分析,使用PaddleVideo的PP-TSM、SlowFast^[13]等实现视频级人员行为分析识别,学会使用PaddleGAN的Pix2Pix、CycleGan实现红外图像转化为可见光图像。对图片级,视频级人员行为识别实验完成效果示例截图如图1a, b所示。人脸识别及红外人脸转换可见光识别实验完成效果示例截图如图2a, b所示。

(3) 本实验以百度飞桨深度学习框架和红外热成像为基础,人员行为分析与身份识别为实验目的,

将图像分类、目标检测及视频识别，图像迁移转化巧妙结合在一起，让学生更好地理解计算机视觉相关知识。学生在进行实验的过程中，需要记录实验过程及数据，分析遇到的报错信息，思考寻找解决方案，撰写完成实验报告。



(a)



(b)

图 2 热成像人脸识别效果示例

3 实验过程及要求

3.1 实验过程要求

(1) 自主预习 Python 语言与 Paddle 框架的基本语法指令，能够完成进行网络训练的基本操作。

(2) 仿照教学示例的热成像数据集自制红外数据集用于训练。用于行为识别的红外热成像数据集部分内容展示如图 3 所示。用于人脸识别的红外热成像和可见光数据集部分内容展示如图 4a, b 所示。

(3) 思考讨论用于图像分类，目标检测，视频识别及图像迁移转化的网络结构原理，并了解各个网络的代码实现来明确各种网络参数对网络训练效果的影响。

(4) 使用 Paddle 框架搭建实验所需的网络模型，掌握通过调节网络参数来更好地训练网络，掌握使用训练好的网络进行模型评估和预测以及模型部署，实现对图片级及视频流中红外人员行为分析与身份识别。

(5) 记录完成实验过程中的主要操作命令，利用 VisualDL 工具完成实验指标数据的可视化记录，整理进行实验过程中遇到的报错信息以及对应的解决方案，撰写实验报告准备答辩。

3.2 实验验收答辩要求

(1) 本实验要求验收时学生搭建的行为检测与身份识别的飞桨网络识别准确率须满足基本要求，且具有一定的推理速度可以满足实时性识别要求。

(2) 学生在验收答辩时，每组需准备一份 PPT，展示自制数据集的文件信息及实验网络在自制数据集上的识别测试效果，阐述网络原理及在飞桨平台上的网络代码实现，如何通过调节相关参数提升模型性能，并在答辩时做到表达流畅，逻辑清晰，能够较好地回答老师提出的问题，方能达到基本的验收要求。

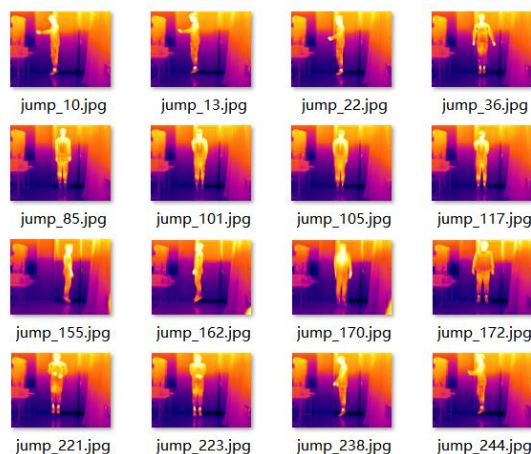


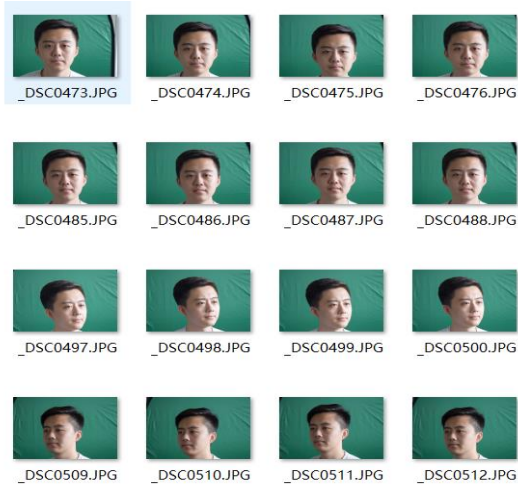
图 3 红外行为数据集部分内容展示

4 实验原理与结果讨论

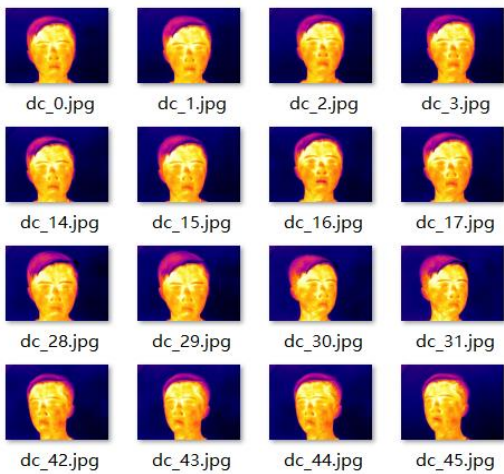
本实验综合采用了飞桨图像识别套件 PaddleClas、飞桨目标检测开发套件 PaddleDetection 以及飞桨视频识别分析套件 PaddleVideo，实现基于黑暗条件下红外热成像进行人员图片级和视频级的行为识别与身份判断。每位同学可根据自身数据集需求通过红外热像仪拍摄人员行为与人脸数据，通过 Labelimg 标注工具对行为进行标注。用于完成基于百度飞桨平台实现红外人员行为与身份识别实验各个网络模块组合起来的整体思路方法设计如图 5 所示。实现红外人员行为与身份识别整体框架图如图 6 所示。各网络模块详解如下。

(1) 基于红外图像的人体行为识别。学生可以采用飞桨目标检测开发套件 PaddleDetection 的最新目标检测网络 PPYOLO 或 PP-PicoDet 网络为核心用于红外人体图像姿态识别，基于百度飞桨 PaddlePaddle 深度学习框架，识别红外姿态类别包括站立、行走、

跳跃、出拳、踢腿等十余类单人及多人人体行为。采用相同的训练模式以及参数设置，对于相同的数据集进行训练与测试，对比结果如下表，基于 Paddle 框架的目标检测系列网络的 MAP 和推理速度优于其他网络，具有更好的识别效果。详细的指标对比如表 1 所示。



(a)



(b)

图 4 红外和可见光人脸数据集部分内容展示

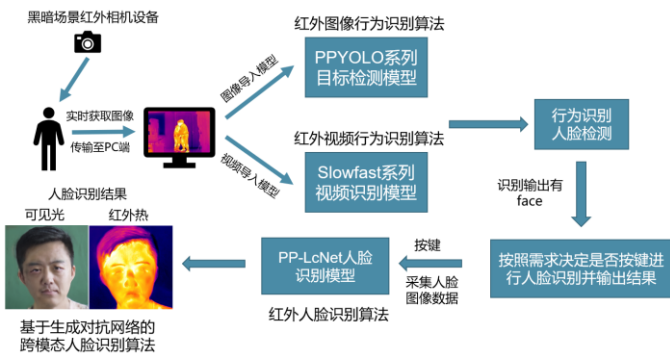


图 5 实现红外人员行为与身份识别整体流程图

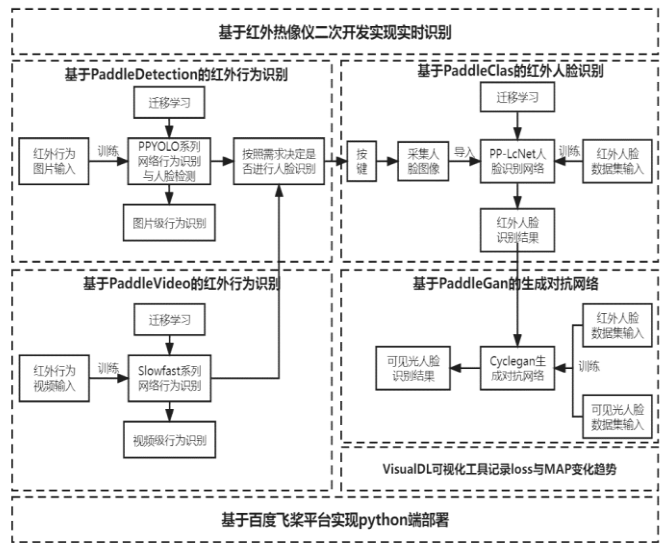


图 6 红外人员行为与身份识别整体框架图

(2) 基于红外热视频的人体行为识别。学生可以采用飞桨视频识别分析套件 PaddleVideo 的 Slowfast 网络为核心实现行为识别目标。SlowFast 网络使用了一个慢速高分辨率 CNN (Slow 通道) 来分析视频中的静态内容，同时使用一个快速低分辨率 CNN (Fast 通道) 来分析视频中的动态内容。基于 PaddlePaddle 框架和飞桨平台搭建 Slowfast 网络模型，通过使用 Facebook 团队训练 AVA 数据集得到的预训练权重进行迁移学习，加快训练速度，从而实现热视频的行为识别技术，目前识别动作包括行走、对话、携带物体等 80 余类单人及多人人体动作。采用相同的训练模式以及参数设置，对于相同的数据集进行训练与测试，详细的指标对比结果如表 2 所示，基于 Paddle 框架的视频识别系列网络的 MAP 和推理速度优于其他网络，具有更好的识别效果。

(3) 基于红外图像的人员身份识别。对于已经录入热图像数据库的人员，利用其热图像在数据库范围内进行匹配识别，学生可以基于飞桨图像识别套件 PaddleClas 的 PP-LcNet、PP-HGNet 网络提取人脸图片特征，得到特征向量，与数据库的数据进行比较，通过相似度函数的阈值判断是否与库内数据匹配。阈值设定综合准确率和召回率两方面数据情况进行考量。

(4) 基于热图像与可见光图像的跨模态人员身份识别。对于未被录入热图像数据库中的非在库人员，学生可以基于飞桨生成对抗网络开发套件 PaddleGAN 的 Pix2Pix、CycleGAN 实现红外图像转化为可见光图像采用跨模态识别算法，实现黑暗环境下非在库人员的识别。后续可利用生成对抗网络 GAN 将红外热图作

为输入，生成可见光图像。然后再利用可见光图像的人脸识别算法进行处理。

表 1 不同目标检测模型识别红外人体姿态效果对比

| Network | MAP@0.5 | MAP@0.5:0.95 | FPS | Size |
|---------------|---------|--------------|-------|-------|
| PPYOLO | 95.40% | 84.40% | 38.06 | 183M |
| PP-YOLOE | 96.20% | 86.60% | 20.49 | 31.2M |
| PP-PicoDet | 95.30% | 85.70% | 69.08 | 8.35M |
| YOLOv3_paddle | 90.40% | 68.90% | 30.21 | 90.8M |
| YOLOX | 92.32% | 79.40% | 39.92 | 34.3M |
| YOLOv5 | 92.25% | 78.00% | 40.22 | 27.2M |
| YOLOv4 | 91.60% | 76.50% | 26.69 | 244M |
| YOLOv3 | 89.70% | 75.40% | 35.11 | 235M |
| SSD | 90.50% | 78.50% | 57.60 | 95.7M |
| Faster R-CNN | 92.70% | 74.90% | 13.82 | 108M |
| RetinaNet | 86.00% | 74.60% | 21.58 | 139M |
| EfficientDet | 86.40% | 73.20% | 19.14 | 15M |
| CenterNet | 88.60% | 76.90% | 59.83 | 124M |

表 2 热成像视频不同行为识别模型指标

| Network | Top-1 |
|----------------|--------|
| PP-TSM | 76.16% |
| PP-TSN | 75.06% |
| PP-TimeSformer | 79.44% |
| TSN | 71.70% |
| TSM | 71.06% |
| SlowFast | 74.35% |
| TimeSformer | 77.29% |
| VideoSwin | 82.40% |
| MoViNet | 66.62% |

5 结束语

本教学案例基于红外热成像领域结合百度飞桨框架进行创新应用，将红外热成像与深度学习相关知识紧密联系，有利于在数据采集环境恶劣的时候，高效准确的达到识别要求，拓宽计算机视觉技术的应用领域，激发学生对深度学习的学习兴趣。同时将图像分类、目标检测及视频识别，图像迁移四大网络模块综合应用到一个整体项目中，学生可以系统性地学习深度学习相关知识，锻炼学生动手能力和解决问题的能

力。在百度飞桨框架搭建的不同网络中引入级联思想实现多任务多目标的识别，培养学生发散思维及其创新整合能力。在后续的实验课程建设中，考虑让学生进一步实现实时性的操作。对自己的网络训练完成后，通过实验室的红外热成像仪器设备进一步完成实时性的人脸识别和人体姿态识别。

参考文献

- [1] 董如意,王通. 基于 Paddle 云平台的机器学习课程建设探索[J]. 电脑知识与技术,2021,17(34):234-235.
- [2] 马文彬,刘旭,魏建宇. 基于百度飞桨的人工智能导论课程实践教学初探[J]. 科学与信息化,2022(8):132-134.
- [3] Paszke A, Gross S, Massa F, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library[EB/OL].(2019-9-3).<https://arxiv.org/abs/1912.01703>.
- [4] Abadi M, Agarwal A, Barham P, et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems.[EB/OL].(2016-3-14).<https://arxiv.org/abs/1603.04467>.
- [5] 马艳军,于佃海,吴甜,等. 飞桨:源于产业实践的开源深度学习平台[J].数据与计算发展前沿, 2019, 1(1):11.
- [6] 陈玖玖,玄玉波,李兆玺,等. 基于百度飞桨 AI Studio 的机器学习教学新模式实践与探索[J]. 计算机教育, 2021(9):5.
- [7] 冯建周,余扬,刘磊. 基于飞桨框架的三阶递进式机器学习教学模式探索与实践[J]. 计算机教育, 2021(10):5.
- [8] 杨舒茜,丁伟. 基于红外热成像技术的物质溶解吸放热创新实验[J]. 教育与装备研究,2022(1):62-64.
- [9] 吴梅芬,王晓岗,许新华. 热成像技术在化学实验教学中的应用[J]. 实验技术与管理,2019,36(12):165-169.
- [10] 龙仲英,解凯彬,凌一洲. 热成像技术在生物学教学中的应用[J]. 生物学教学,2021,46(7):45-46.
- [11] S. Xu et al. PP-YOLOE: An evolved version of YOLO[EB/OL].(2022-04-02). <https://arxiv.org/abs/2203.16250>.
- [12] G. Yu et al. PP-PicoDet: A Better Real-Time Object Detector on Mobile Devices[EB/OL].(2021-11-01) .<https://arxiv.org/abs/2111.00902>.
- [13] Feichtenhofer C, Fan H, Malik J, et al. SlowFast Networks for Video Recognition[C]// 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE, 2019: 6201-6210