

DOI: 10.13382/j.jemi.B2205344

基于无人船视觉的水域人员类别识别算法*

程亮^{1,2} 吴兴辉³ 江云华² 苏雄² 吴佳晓² 周辉³ 丁美有³ 何贇泽³(1. 江苏海洋大学海洋工程学院 连云港 222005; 2. 珠海云洲智能科技有限公司 珠海 519085;
3. 湖南大学电气与信息工程学院 长沙 410006)

摘要:针对水域环境下人员识别,提出了一种基于水面无人船(unmanned surface ship,USV)视觉传感器的水域人员类别识别算法。依照数据采集与模型更新流程,将采集到的视频数据进行数据清洗与标记后,创建人员类别数据集39 959张图片,7个类别;实践了基于深度学习方法下主流目标检测网络YOLO v5,并针对水域环境场景特点,提出基于YOLO v5的人员类别识别算法;将人员类别识别算法部署到边缘计算平台,实现算法在无人船上的实时应用。算法在人员类别识别数据集上达到了平均精度86%,在无人船实测中实现了每秒处理38帧的人员类别识别实时性表现。

关键词:水面无人船;水域人员识别;YOLO v5;网络部署

中图分类号: TH98;TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Person category identification algorithm in water environment based on unmanned ship vision

Cheng Liang^{1,2} Wu Xinghui³ Jiang Yunhua² Su Xiong²Wu Jiexiao² Zhou Hui³ Ding Meiyu³ He Yunze³

(1. School of Ocean Engineering, Jiangsu Ocean University, Lianyungang 222005, China;

2. Zhuhai Yunzhou Intelligent Technology Co., Ltd., Zhuhai 519085, China;

3. College of Electrical and Information Engineering, Hunan University, Changsha 410006, China)

Abstract: To achieve person recognition in water environment, a person category identification algorithm based on vision sensors on unmanned surface ship (USV) is proposed. Firstly, base on the data acquisition and model update workflow, a person category dataset of 39 959 pictures and 7 categories is created after data cleaning and labeling on original videos. Secondly, YOLO v5, the mainstream object detection network in the field of deep learning method, is practiced, and an improved person category identification algorithm based on YOLO v5 is proposed according to the characteristics of water environment scenes. Thirdly, the algorithm is deployed to the edge computing platform to realize the real-time use of the algorithm on the unmanned ship. The algorithm achieves an average accuracy of 86% on our dataset and achieves real-time performance of processing 38 frames per second with accurate person recognition in the unmanned ship test.

Keywords: USV; person identification in water environment; YOLO v5; network deployment

0 引言

湖泊、河流、海洋等水域是人们出行、工作的重要活动区域,也是人类未来发展的重要空间。人类在水域环境下活动范围的增大和频率的增加,增强水域环境开发

能力、提升水域环境下人员类别识别能力是作为环境监控系统、智能救生系统等水域管理系统的基础。目前水域管理系统对水域人员的管理主要还依靠监控员在室外或监控室观察监控的,对于泳池等室内水域环境,已经有研究人员基于机器视觉方法进行简单的人员行为识别。北京工业大学的方之愉^[1]基于传统计算机视觉的方法建

收稿日期: 2022-04-04 Received Date: 2022-04-04

* 基金项目:湖南省自然科学基金重大项目、珠海云洲智能科技有限公司委托课题项目资助

立了基于滑动均值的背景建模算法和目标检测算法实现对溺水目标的实时跟踪。青岛大学乔羽等^[2]基于深度学习训练得到的 Mask R-CNN 网络实现对游泳、水中站立等行为的检测。但这些方法仅适用于背景简单的室内场景,检测精度一般,且算法需运行在服务器上。对于室外水域,由于水域环境背景复杂、离线端部署难等原因,尚未有基于机器视觉技术对水域环境人员进行类别识别的先例。

随着人工智能在机器视觉技术上的发展与应用,水域管理系统智能化、无人化会成为未来的发展趋势。近年来发展迅速的水面无人船已在水域环境测绘、水域生态环境保护、水质检测^[3-5]等领域成为重要的无人化作业工具,将基于深度学习的机器视觉技术与无人船上的光电传感器等视觉系统相结合,并将算法部署到搭载在无人船上的边缘计算设备,使无人船在进行水域巡航时运行人员识别算法实时识别水域人员类别,进而为后续水域人员安全管理、无人船智能避障、智能救援等功能提供支持,具有重要的应用价值与发展前景。

因此本文基于无人船搭载的摄像头等视觉系统,将目前国内的主流目标检测算法 YOLO v5 应用于水域人员类别识别问题,并基于水域场景进行多处改进,设计出搭载于无人船,用于离线端实时识别应用的水域人员类别识别算法。

1 基于无人船视觉的人员类别识别算法

基于深度学习的目标检测网络始于 Girshick 首次提出的采用基于区域的卷积神经网络^[6] (region based convolutional neural network, R-CNN),相比传统目标检测算法精度更高。目标检测网络在发展中分为单阶段目标检测与双阶段目标检测网络^[7]。双阶段目标检测算法第 1 阶段先产生目标候选框即目标位置,而在第 2 阶段对候选框做分类与回归。单阶段目标检测算法中,Redmon 等^[8]提出的 YOLO(you only look once)系列网络是目前应用广泛的经典网络之一。YOLO 算法针对双阶段目标检测算法运算速度慢的缺点,创造性地将目标定位与分类在一个阶段完成,开辟了单阶段目标检测新思路。算法通过端到端的网络设计,最终输出包含类别信息的空间位置向量,其想对于双阶段网络的检测速度优势使 YOLO 系统网络成功应用于如故障检测^[9]、红外场景目标检测^[10]、遥感目标检测^[11]、自动驾驶^[12]等机器视觉任务中。YOLO 系列在不断改进中已发展到 YOLO v5,是最先进的目标检测网络之一,其在推理上的速度优势也使其迅速应用于机器视觉任务中。

在训练集对图像进行数据增强处理能够有效地提高网络总体准确率^[13],YOLO v5 网络的输入端采用了

Mosaic 数据增强^[14],这种数据增强应用于模型训练阶段,不同于传统的旋转、裁剪等单幅图像增强方法,Mosaic 数据增强会随机抽取 4 张图片进行随机缩放、剪切与排布等方式将多张图片拼接成新图作为网络输入。Mosaic 数据增强既增加了数据的多样性,丰富了图片的背景,缓解模型训练过程中数据缺乏与场景单一问题,又因为小目标在网络中得到更多训练使得网络对小目标的识别能力增强。如图 1 所示为 Mosaic 数据预处理后的 4 张图片,每张增广后的图像由 4 张图片拼接而成。

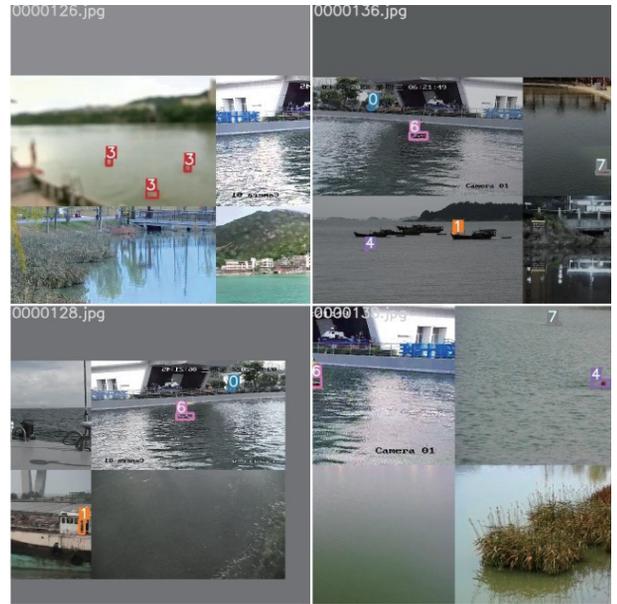


图 1 Mosaic 数据增强

Fig. 1 Mosaic data augmentation

YOLO v5 网络可分为骨干网络 (backbone)、中间层 (neck) 及预测层 (prediction) 3 大部分,具体结构如图 2 所示,其中“Conv”代表卷积层,“BN”代表批归一化层,“Leaky ReLU”表示激活函数,“CBL”代表一个卷积、一个批归一化、一个激活函数层,“Concat”代表矩阵的合并。

骨干部分主要采用了 CSPDarknet^[15] 和 SPP (spatial pyramid pooling) 模块^[16]。CSPDarknet 中的 CSP 模块结构如图 2 所示,在 YOLO v5 中 CSP 模块有 CSP1_X 结构和 CSP2_X 两种结构,网络结构图左下角展示了该模块结构,都是将输入分为两个分支后分别进行特征提取然后“Concat”。在骨干网络中 CSP 模块采用 CSP1_X 结构,其用于特征处理的主要是残差组件,其中“X”代表了使用了“X”个残差组件,残差组件连接之前所有层的输出特征来作为下一层的输入,增加残差结构可以增加层与层之间反向传播的梯度值,避免因模深度加深而带来的梯度消失,从而可以提取到更细粒度的特征并且不用担心网络退化。网络结构图右下角为 SPP 模块结构,SPP 模块使得卷积神经网络输入不受尺寸限制,专

注于空间信息,它有 4 个并行分支连接:3 个最大池操作(核大小为 5×5、9×9、13×13)和原输入的短接方式。SPP 模块通过合并输入的局部和全局特征来解决对象尺度变化过大的问题,在增大了网络感受野的同时更有效地分离上下文的特征。

网络的 Neck 部分使用了 CSP 模块中的 CSP2_X 结构,相对于 CSP1_X, CSP2_X 将残差组件换成了 2×X 个“CBL”。Neck 部分包含了 FPN (feature pyramid networks) 和 PAN(pixel aggregation network) 结构,FPN 是一个自顶向下的特征金字塔,将高层的强语义特征传递

下来,而 PAN 在 FPN 结构后,是一个自下向上的金字塔,将低层的定位特征传递上去,这样形成的金字塔既结合了语义信息又拥有定位信息。

在检测层, YOLO v5 采用了非极大值抑制(non maximum suppression, NMS) 用来删除高度冗余的检测。抑制的过程是一个迭代-遍历-消除的过程,按类别基于框置信度排序,选中置信度最高框后遍历同类别其余框,将与其重叠面积大于一定阈值框删除,而后从未处理的框中继续选取置信度最高框重复上述过程。

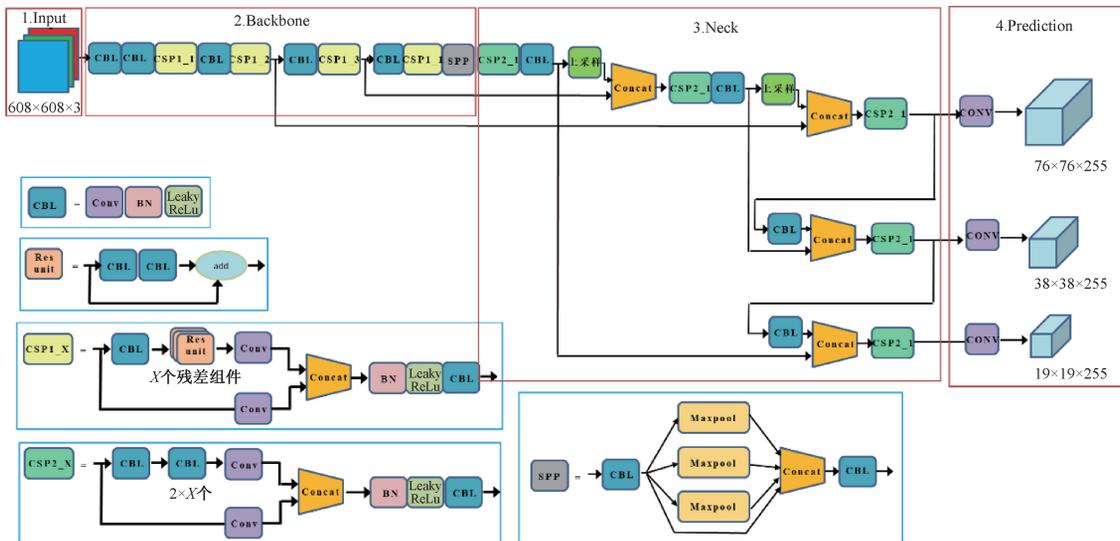


图 2 YOLO v5 结构图

Fig. 2 Structure of YOLO v5

具体到应用于水面无人船水域巡检场景,存在如下几个难点:天气较差光线不足情况下网络识别能力;小尺寸目标人员的识别;陌生场景类别误识别和漏检。这些应用场景的具体问题,给人员识别算法应用可行性带来相当大的挑战,因此本文提出了一种基于 YOLO v5 的改进算法。

1.1 采用直方图正规化应对背景较暗场景

对于数据集中一些背景较暗、整体灰度值偏低的图片,网络识别能力较差。本文采用直方图正规化的方法对数据集中整体灰度值较低的图片进行预处理,提高了网络识别能力。直方图正规化通过一个灰度映像函数,将原灰度直方图改造成所希望的直方图,强调某些感兴趣的特征,可将背景较暗图像变得清晰,扩大图像中不同目标特征间的差别,改善图像质量、丰富信息量,有助于提升网络识别效果。

假设输入图像为 I , 高为 H , 宽为 W , $I(r, c)$ 代表 I 的第 r 行第 c 列的灰度值, 将 I 中出现的最小灰度级记为 I_{\min} , 最大灰度级记为 I_{\max} , 即 $I(r, c) \in [I_{\min}, I_{\max}]$, 为使输出图像 O 的灰度级范围为 $[O_{\min}, O_{\max}]$, 对 $I(r, c)$ 和

$O(r, c)$ 做如下映射关系:

$$O(r, c) = \frac{O_{\max} - O_{\min}}{I_{\max} - I_{\min}}(I(r, c) - I_{\min}) + O_{\min} \quad (1)$$

式中: $0 \leq r \leq H, 0 \leq c \leq W, O(r, c)$ 代表 O 的第 r 行第 c 列的灰度值。

在本文直方图正规化程序中, 因为 $0 \leq \frac{I(r, c) - I_{\min}}{I_{\max} - I_{\min}} \leq 1$, 所以 $O(r, c) \in [O_{\min}, O_{\max}]$, 故令 O_{\min} 为 0, O_{\max} 为 255。

1.2 采用基于 patch 的图像处理应对小目标场景

水域巡航时, 在宽广水域如海边水域场景下进行大分辨率尺寸视频输入时算法对小目标人员存在部分漏检。为提升算法对该水域场景下小目标人员的检出率, 本文针对大分辨率尺寸视频输入时存在小目标场景设置检测选项, 采用将输入图像分成 $n \times n$ 个图像块(patch), 在对每个图像块四周进行一定比例的扩展后, 将每个图像块尺寸调整到原 YOLO v5 输入分辨率进行检测并映射回原图, 使用 NMS 对重叠区域的重叠目标的重复预测情况进行抑制。由于被检测目标在所占图像比例增加,

网络的目标检出率更高。综合考量精度与速度,一般 n 设置为 2 或 3,将一张图片转换为 4 或 9 个图片块,本文中 n 设为 2。

1.3 负样本训练 (NST)

在深度学习领域,深度学习模型遇到陌生对象或场景时(如训练集中不存在的类别或有意加入了特定干扰噪声的图片)模型的预测能力会降低,这导致在实际应用中会造成较为严重的误报现象。模型若将背景或陌生对象检测成数据集中的某类别,会影响无人船后续的功能实现。为了解决这一严重干扰识别的情况,降低在实际应用中的模型误报率,本文提出了负样本训练(negative samples training, NST)的方法。在数据集中加入空标签图像作为负面样本即不包含识别对象信息的图像作为部分训练集进行模型性能优化。实验结果证明在不影响识别速度的情况下,算法相比原 YOLO v5 算法的精度更高,算法通过降低对未知目标和场景的识别置信度,从而降低误报率。加入更多空标签,模型的准确性指标会提升明显,但对模型泛化性的影响难以估量,因此本文负样本根据实际误报情况酌情添加,在训练集的比例设置为 10%。

1.4 基于 TensroRT 模型实现离线端实时推理

在实际应用中,考虑到模型需搭载到水面无人船上,因此部署后的算法需对 1 920×1 080 pixels, 25 fps 的视频数据的处理输出频率在 10 Hz 以上。本文的部署平台采用英伟达 Jetson Xavier 开发者套件,该套件支持 NVIDIA JetPack 和 DeepStream SDK 等开发工具包以及 CUDA、cuDNN 和 TensorRT 等深度学习软件库。基于 Pytorch 深度学习框架的算法在服务器端基于数据集训练后,将得到的模型权重转换为基于 TensroRT 框架的模型并部署到 Xavier 上。在船载部署时,将边缘计算平台通过电源线、网线、连接线和无人船中控系统相连,而后摄像头采集的视频流数据传输给船上的 Xavier 嵌入式平台, Xavier 实时在线推理并反馈给无人船推理结果,如图 3 所示。

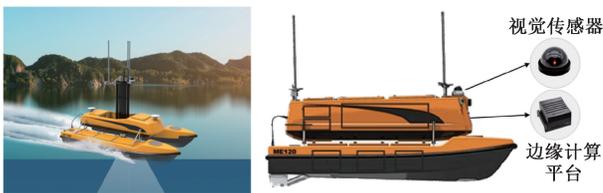


图 3 Xavier 嵌入式无人船系统

Fig. 3 Xavier embedded unmanned boat system

可应用于水域监控和人员救助场景,因此在本文数据集水域人员类别设置上分为了水中工作人员、岸上人员、船上人员、戏水人员、溺水人员、游泳人员、获救人员共 7 类,将其分别标记为 waterworking、person_shore、person_boat、waterplaying、drowning、swimming、waterhelping。

标准化的数据集对目标检测研究具有不可或缺的重要性,其数量和质量都将影响网络模型的性能。本文类别识别数据集来源于网络开源数据集、珠海云洲智能科技有限公司提供的真人下水模拟溺水场景视频以及水面无人船“海豚一号”水域作业视频。原始数据通过视频抽帧、基于结构相似度去重筛选、标注与审核等工作加入数据集,在模型不断迭代过程中数据集规模在逐渐增大。在原始数据集建立后,通过模型训练与测试后,将性能最好的模型进行部署并将实际使用过程中的漏检、误检等错误数据保存并加入数据集进行模型迭代。整体工作流程如图 4 所示。

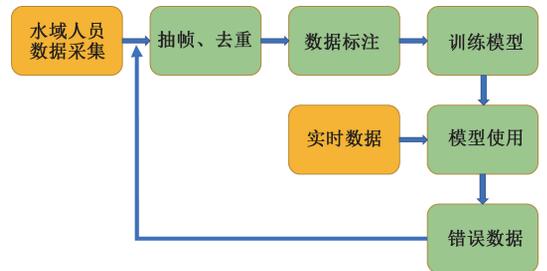


图 4 工作流程

Fig. 4 Work flow chart

最终本文类别识别数据集达到了 39 959 张图片,基于数量命名为 YZ40K,数据集的构成如表 1 所示,部分类别示例如图 5 所示,可以看出数据集中图片背景变化较大,所需检测人员的标签大小变化也很突出。总的来说这是个多目标检测任务,单张图片可能包含多类多个目标,图片数量为至少含有一个某类目标的图片之和,不同类别目标的图片数量与标签数量存在较大差异。

表 1 YZ40K 数据集构成

Table 1 YZ40K datasets composition

类别名	标签	图像数量	标签数量
水中工作人员	Waterworking	2 709	9 078
岸上人员	Person_shore	9 066	27 993
船上人员	Person_boat	7 904	13 517
戏水人员	Waterplaying	3 755	11 564
溺水人员	Drowning	6 262	6 538
游泳人员	Swimming	3 699	6 391
获救人员	Waterhelping	3 190	3 223

2 人员类别识别数据集

水域环境下的无人船水域巡检输出的人员类别识别

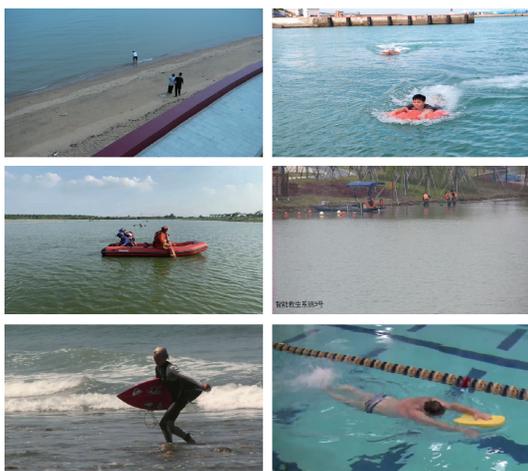


图 5 部分人员类别展示

Fig. 5 Display of some person categories

3 实验与结果分析

3.1 模型训练

1) 训练参数

模型的训练测试所用 CPU 处理器为 i9-10900X, GPU 处理器为 Nvidia RTX3090, 深度学习环境由 Python3.8、Opencv3.4.2、CUDA11.1、CUDNN8.0、PyTorch1.8、TensorRT7.0 等构成。YZ40K 数据集 39 959 张由多轮数据迭代得到, 初建数据集按照 9 : 1 的比例划分后, 后续新增误检场景数据按 0.95 : 0.05 比例加入, 最终得到训练集 37 737 张, 测试集 2 512 张, 同时将未处理过的实际应用场景视频作为验证集进行模型性能评估。根据模型深度, YOLO v5 网络包含 YOLO v5S、YOLO v5L、YOLO v5M、YOLO v5X 四个深度依次增大的网络模型, 本文算法基于 YOLO v5M 进行训练。训练时每轮模型的更新设置迭代世代 (epoch) 值 400, 同时采用早停法监控训练集的损失值, 当其长时间变化极小时中止网络训练以避免网络在训练集上的过拟合。模型批次大小设为 40, 初始学习率 0.001, 动量参数 0.9, 权重衰减参数 0.0005, 采用自适应矩阵估计动量法 (Adam) 进行损失优化。

2) 损失函数

本实验采用损失函数由边框回归损失、置信度损失和分类损失 3 部分组成。边框回归损失采用 GIoU loss 损失函数进行计算, 其中真值框为 A , 预测框为 B , 相较于 YOLO 系统原来的 IOU 损失函数, GIoU Loss 损失函数考虑到无重叠框时差距无法评估的情况, 加入了预测框和真实框的最小外接矩形即 C 作为惩罚项进行优化, 具体计算过程如式(1)所示。置信度损失与分类损失均采用

Focal Loss 损失函数^[17]进行计算。Focal Loss 损失函数将调节参数 α 和 γ 加到交叉熵损失函数之上, 可以调节不同样本影响权重的能力, 可以有效缓解目标检测数据集中正负样本比例失衡的难题, 计算过程如式(2)所示。

$$GIOU = IOU - \frac{\Delta}{C} = \frac{A \cap B}{A \cup B} - \frac{C - A \cup B}{C} \quad (2)$$

$$FocalLoss(p) = -a(1-p)^\gamma \ln(p) \quad (3)$$

式中: p 表示预测概率, a 和 γ 为预设调节参数。

3.2 测试指标

测试主要对网络训练结果的识别速度与识别能力以及网络是否支持部署进行评估。识别速度使用每秒识别图像数量 (FPS) 判定, 识别能力使用召回率和平均精度 mAP (mean average precision) 判定。FPS 越高说明网络检测速度越快, 召回率代表了某类目标所有检测正确的框与所有实际标记框为该类目标的数量占比, 其数值越高表明了模型目标的漏检能力越低, mAP 是所有类别 AP 的平均值, 其由各类别的准确率与召回率构成的 P-R 曲线计算得到, mAP 越高说明网络识别精度越高。本文根据 VOC 数据集定义^[18]公式, 将检出框与标签框的交并比 (intersection over union, IoU) 大于 50% 归为预测正确结果。模型能否部署主要由模型大小来衡量, 模型权重大小在 200 MB 以内可流畅部署运行。

3.3 测试结果

将训练所得最佳权重在测试集上进行测试, 其中网络检测置信度阈值设为 0.4, 预测置信度高于该阈值的目标框都将作为预测结果, 图 6 为网络在测试集上的测试结果展示。

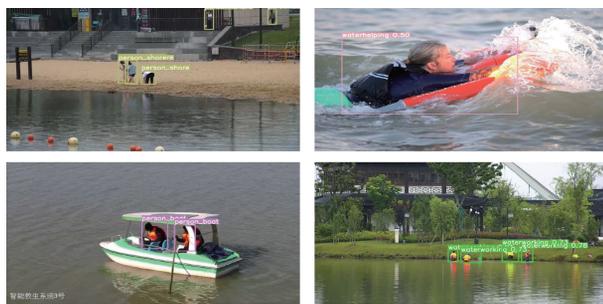


图 6 测试结果展示

Fig. 6 Display of detection results

对于数据集中整体灰度值较低图片, 通过算法的直方图正规化处理后, 图片的视觉效果得到改善, 目标特征更加明显。图 7 和 8 为应用该方法前后网络的预测结果对比, 可以看出处理后的图片清晰度更高, 人员的检出能力也更强。

对存在小目标的场景的测试证明了算法基于 patch 的处理有效提升小目标人员的检出水平。如图 9 所示为

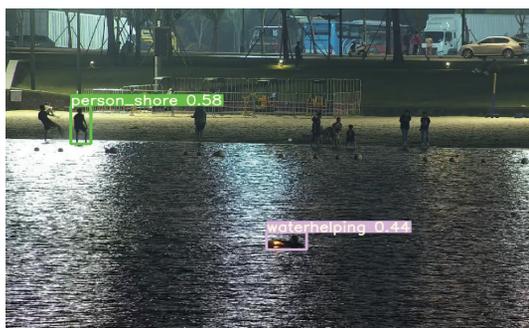


图 7 YOLO v5 网络预测图 1

Fig. 7 Detected picture 1 of YOLO v5



图 10 本文算法预测图 2

Fig. 10 Detected picture 2 of our algorithm

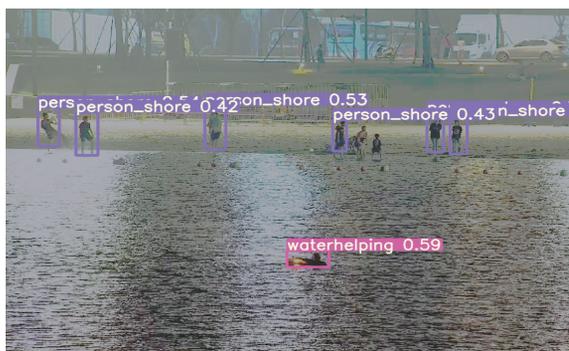


图 8 本文算法预测图 1

Fig. 8 Detected picture 1 of our algorithm

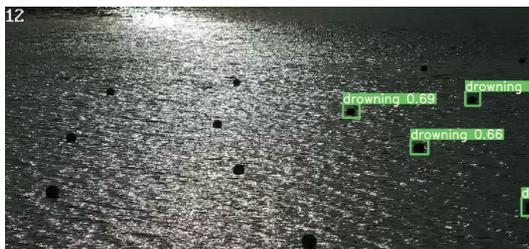


图 11 负样本训练前预测结果

Fig. 11 Detection result before the usage of blank labels training

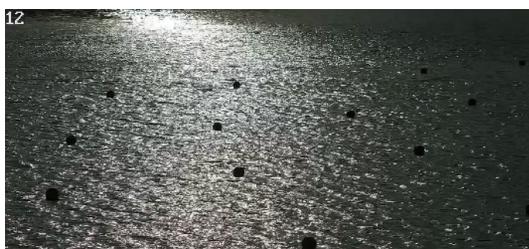


图 12 负样本训练后预测结果

Fig. 12 Detection result after the usage of blank labels training

YOLO v5 网络的预测结果,图 10 为本文算法预测结果,可以看出在维持原有检测结果外,有助于落水人员等小目标人员的检出。



图 9 YOLO v5 网络预测图 2

Fig. 9 Detected picture 2 of YOLO v5

负样本训练的方法有助于网络降低误检率,提升网络整体性能。如图 11 所示,最初网络在检测某水域下的圆球形浮标场景时,网络预测结果为溺水人员。在数据集加入该类不含人员目标的场景图像作为空标签负样本后,更新后的网络再没出现这种误检现象,如图 12 所示。

图 13 为模型各类别测试精度图,mAP 达到了 86%,整体来看网络表现出良好的适应性,各类别识别表

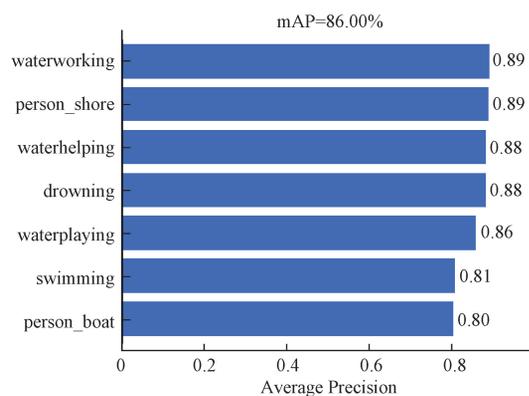


图 13 本文算法测试集 mAP

Fig. 13 The mAP of our algorithm on val dataset

现均衡,其中识别精度最高的水中工作人员和岸上的人

的达到了 89%, 识别效果最差的船上的人也有 80% 的精度。具体到各类别检出能力, 图 14 为模型的在测试集上的召回率, 可以看出模型对于水中工作人员检出性能最佳, 而船上的人召回率最低为 72%。图 15 为具体的检测结果, 可以看出网络的整体检测性能较好, 误检比例较低。

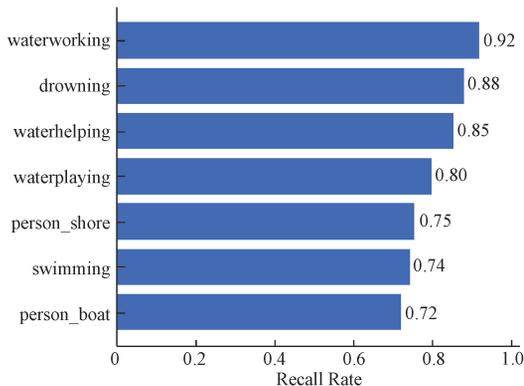


图 14 本文算法测试集各类别召回率

Fig. 14 Recall rate of our algorithm on val dataset

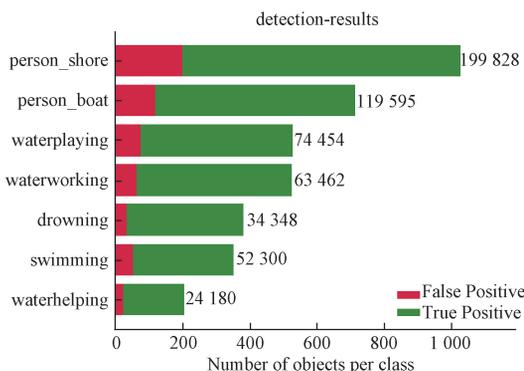


图 15 本文算法检测结果

Fig. 15 Detection results of our algorithm on val dataset

图 16 为模型的混淆矩阵。该图横坐标为预测类别, 纵轴坐标为标签类别。由于引入空标签作为负面样本, 故将空标签“none”纳入统计。对角线即为预测正确的。图 16 中将存在误检的检测框数量及占该类别比例显示出来, 可看出该网络存在戏水和游泳类别的混淆和部分的背景误识别(即本真值中不存在的区域却预测出目标)总的来说, 改进后的算法表现出良好的适用性, 对于动作差异较大的 7 个类别的识别精度维持在较高水准, 整体识别准确率达到了实用需求, 可以支持后续无人船智能避障或救援等需求实现。

3.4 模型部署

算法在服务器端训练得到最佳权重后经过 C++ 部署到 Jetson Xavier, 搭载在水面无人船后可以进行实时检测

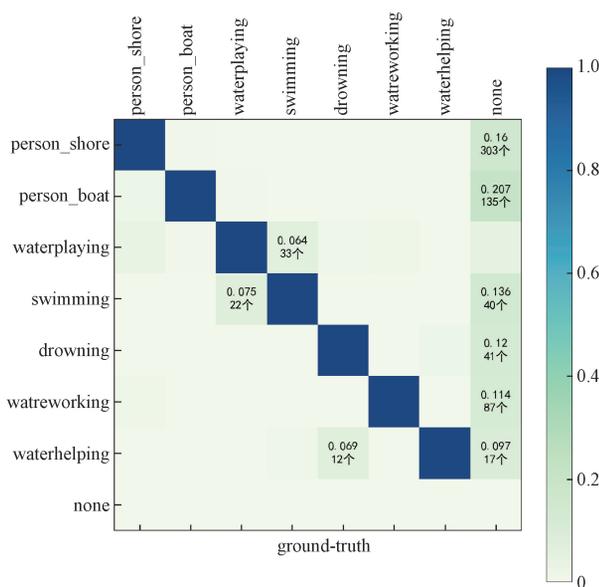


图 16 本文算法测试集混淆矩阵

Fig. 16 Confusion matrix of our algorithm on val dataset

并返回检测结果, 如图 17 所示, 无人船在湖泊和浅水区域航行时, 部署算法准确检测出岸边人员和水中人员, 整个过程运行良好。经过测试, 部署后的权重达到了 83% 的识别精度与 FPS38 的识别速度, 满足了实际应用场景的需求。

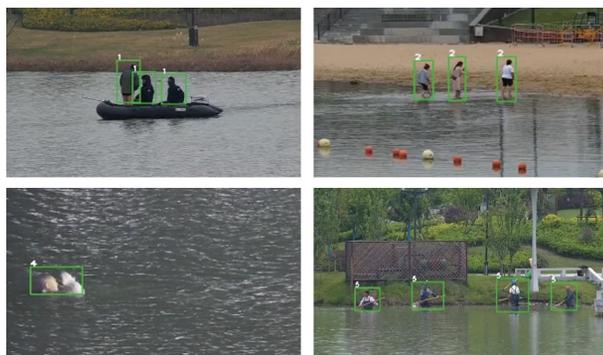


图 17 Xavier 端预测图

Fig. 17 Detected pictures on Xavier

4 结论

本文基于水面无人船视觉, 提出了一种可成功部署于无人船水域巡检的人员类别识别算法。算法基于 YOLOv5 网络的基础上进行改进, 对图像进行直方图正规化处理, 提升算法在恶劣天气较暗场景下的识别能力; 对图像进行基于 patch 的检测提升算法对中小目标人员的识别能力; 利用 NST 方法提升模型泛化性。从算法部署后的无人船测试结果可知, 本文方法具有稳定、快速与

准确等优点,在背景相对复杂的水面环境下有较好的鲁棒性,能够发挥出单阶段目标检测算法在速度上的最大优势。在后续研究工作中,会尝试轻量化网络结构、进行基于视频的三维卷积网络识别来提升识别精度与速度。

参考文献

- [1] 方之愉. 基于计算机视觉的泳池智能辅助救生系统研究[D]. 北京:北京工业大学,2018.
FANG ZH Y. Research on intelligent auxiliary life-saving system of swimming pool based on computer vision [D]. Beijing:Beijing University of Technology,2018.
- [2] 乔羽. 基于Mask R-CNN泳池中溺水行为检测系统的设计与实现[D]. 青岛:青岛大学,2019.
QIAO Y. Design and implementation of drowning behavior detection system in swimming pool based on mask R-CNN [D]. Qingdao:Qingdao University,2019.
- [3] 常继强,蒲进菁,庄振业,等. 无人船在海洋调查领域的应用分析[J]. 船舶工程, 2019, 41(1):6-10,78.
CHANG J Q, PU J J, ZHUANG ZH Y, et al. Application analysis of unmanned ship in marine survey [J]. Ship Engineering, 2019, 41 (1): 6-10,78.
- [4] 方中华,褚宏宪,冯京,等. 无人船艇在海洋地质调查中的应用及展望[J]. 海洋地质前沿, 2020,36(3):72-77.
FANG ZH H, CHU H X, FENG J, et al. Application and prospect of unmanned boat in marine geological survey [J]. Marine Geological Frontier, 2020,36(3):72-77.
- [5] 李伟,李天伟,张尚悦,等. 水面无人艇技术发展及展望[J]. 舰船电子工程, 2021, 41(4):1-3,12.
LI W, LI T W, ZHANG SH Y, et al. Development and prospect of surface unmanned craft technology [J] Ship Electronic Engineering, 2021, 41 (4): 1-3,12.
- [6] GUPTA S, GIRSHICK R, ARBEL E Z P, et al. Learning rich features from RGB-D images for object detection and segmentation [C]. European Conference on Computer Vision, 2014.
- [7] ZOU Z, SHI Z, GUO Y, et al. Object detection in 20 years: A survey [J]. ArXiv Preprint, 2019, arXiv:1905.05055.
- [8] REDMON J, FARHADI A, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. Computer Vision & Pattern Recognition, 2016.
- [9] 韩航迪,徐亦睿,孙博,等. 基于改进 Tiny-YOLOv3 网络的航天电子焊点缺陷主动红外检测研究[J]. 仪器仪表学报,2020,41(11):42-49.
HAN H D, XU Y R, SUN B, et al. Research on active infrared detection of aerospace electronic solder joint defects based on improved tiny YOLOv3 network [J]. Chinese Journal of Scientific Instrument, 2020,41 (11): 42-49.
- [10] 曹红燕,沈小林,刘长明,等. 改进的 YOLOv3 的红外目标检测算法[J]. 电子测量与仪器学报,2020,34(8):188-194.
CAO H Y, SHEN X L, LIU CH M, et al. Improved infrared target detection algorithm of YOLOv3 [J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(8): 188-194.
- [11] 化嫣然,张卓,龙赛,等. 基于改进 YOLO 算法的遥感图像目标检测[J]. 电子测量技术,2020,43(24):87-92.
HUA Y R, ZHANG ZH, LONG S, et al. Remote sensing image target detection based on improved YOLO algorithm [J]. Electronic Measurement Technology, 2020,43 (24): 87-92.
- [12] 汪辉,高尚兵,周君,等. 基于 YOLOv3 的多车道车流量统计及车辆跟踪方法[J]. 国外电子测量技术,2020,39(2):42-46.
WANG H, GAO SH B, ZHOU J, et al. Multi lane traffic flow statistics and vehicle tracking method based on YOLOv3 [J]. Foreign Electronic Measurement Technology, 2020,39 (2): 42-46.
- [13] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. ArXiv Preprint, 2018, arXiv:1804.02767.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H Y. YOLOv4: Optimal speed and accuracy of object detection [C]. ArXiv Preprint,2020,arXiv:2004.10934.
- [15] WAN C Y, LIAO H, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2020.
- [16] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 37(9):1904-1916.
- [17] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017: 2999-3007.
- [18] EVERINGHAM M, GOOL L V, WILLIAMS C K I, et al. The pascal visual object classes (VOC) challenge [J]. International Journal of Computer Vision, 2010, 88(2): 303-338.

作者简介



程亮, 2007 年于北京大学获得学士学位, 2008 年于英国帝国理工学院获得硕士学位, 2013 年于英国纽卡斯尔大学获得博士学位, 现为珠海云洲智能科技有限公司资深视觉算法工程师, 主要研究方向为无人船艇人工智能、视觉感知、多传感器的融合感知等技术的开发与应用。

E-mail: lance.cheng@yunzhou-tech.com

Cheng Liang received his B. Sc. degree in 2007 from Peking University, received his M. Sc. degree in 2008 from Imperial College London and received his Ph. D. degree in 2013 from Newcastle University. Now he is a Senior Engineer in Computer Vision in Zhuhai Yunzhou Intelligent Technology Co., Ltd. His main research interests include AI, computer vision, sensing fusion for USV.



吴兴辉, 2020 年于郑州大学获得学士学位, 现为湖南大学硕士研究生, 主要研究方向为图像处理、深度学习。

E-mail: hnuwxh@163.com

Wu Xinghui received his B. Sc. degree in 2020 from Zhengzhou University. Now he is a M. Sc. candidate in Hunan University. His main research interests include image processing 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, 现为湖南大学电气与信息工程学院仪器科学与技术系主任、教授, 主要研究方向为新能源设备健康监测、红外机器视觉与人工智能、智能传感与无损检测、电力电子和储能电池健康监测和预警、电力系统在线监测与故障诊断、基于物联网的健康监测。

测、红外机器视觉与人工智能、智能传感与无损检测、电力电子和储能电池健康监测和预警、电力系统在线监测与故障诊断、基于物联网的健康监测。

E-mail: yhe@vip.163.com

He Yunze (Corresponding author) received his B. Sc. degree in 2006 from Xi'an Jiaotong University, received his M. Sc. degree in 2008 from National University of Defense Technology, received his Ph. D. degree in 2012 from National University of Defense Technology. Now he is the director and professor of the Department of Instrument Science and technology, College of Electrical and Information Engineering, Hunan University. His main research interests include new energy equipment health monitoring, infrared machine vision and artificial intelligence, intelligent sensing and nondestructive testing, power electronics and energy storage battery health monitoring and early warning, online power system monitoring and fault diagnosis, and health monitoring based on the Internet of Things.