

## ◆特邀专栏◆

## 基于 YOLOv3 及迁移学习的河道船舶目标检测算法\*

李 惠<sup>1</sup>, 蒋 权<sup>2\*\*</sup>

(1. 广西民族大学电子信息学院, 广西南宁 530006; 2. 广西民族大学人工智能学院, 广西南宁 530006)

**摘要:**为加强对河道监控视频图像中散体物料采运船舶的监测和跟踪,从而辅助实现智能、高效的河道采砂监管和散体物料调度,基于 You Only Look Once version 3 (YOLOv3)算法及迁移学习提出一种河道散体物料船舶目标检测算法。首先使用 COCO 数据集训练初始的 YOLOv3 算法,得到模型的预训练权重;然后对从广西重要河道周围监控设备采集的采砂运砂船舶影像数据进行图像处理,得到高质量船舶数据集;最后以此数据集为驱动,利用迁移学习得到的预训练权重来训练针对河道采砂船等重点目标的 YOLOv3 检测模型。该模型采用 Darknet-53 作为主干网络,并融合了多尺度的特征图,从而实现对小、中、大等各类目标的检测。实验结果表明:该算法在测试集上的平均精度和检测速度分别达到 98.00% 和 17.78 fps,对提高河道采砂监管效能和实现散体物料智能调度具有现实意义。

**关键词:**YOLOv3; 迁移学习; 散料船舶; 目标检测; 采砂监管

中图分类号: TP183 文献标识码: A 文章编号: 1002-7378(2023)03-0331-09

DOI: 10.13657/j.cnki.gxkxyxb.20230829.013

建筑行业的快速发展使得河砂使用量和需求量大幅上升<sup>[1]</sup>。频发的非法采砂问题导致河砂过度开采,给河势稳定以及生态环境造成了严重影响,也给防洪安全埋下了隐患<sup>[2]</sup>。广西坐落于我国华南地区,面向东南亚,东临粤港澳,拥有我国西部地区唯一的出海口,内河航运企业众多。截至 2020 年底,广西约有航运企业 370 家<sup>[3]</sup>。随着平陆运河建设的推进,广西将会更快成为东盟国际大通道和国际陆海贸易新通道的重要枢纽<sup>[4]</sup>。然而,由于河道采砂管理工作中

存在采砂规划缺失,违法采砂监管不力以及管理法规操作性不强等问题,区内河段仍不时有违法偷采现象发生<sup>[5]</sup>,这对广西水运的发展管理与生态环境保护造成不良影响。丁继勇等<sup>[6]</sup>指出,非法采砂不仅不能获取公益性效益,还会对生态环境造成极大威胁,如加剧耕地盐碱化、加大边坡不稳定性、影响水位变化、破坏水生和河岸栖息地等。耿金荣<sup>[7]</sup>的研究表明,非法采砂会给河床、水上治安以及堤防安全带来严重危害,需引入高科技采砂技术等方式强化对河道的管

收稿日期: 2023-03-29

修回日期: 2023-06-26

\* 广西科技基地和人才专项(桂科 AD21220002)资助。

## 【第一作者简介】

李 惠(1995-),女,在读硕士研究生,主要从事计算机视觉研究,E-mail:lihuiwit@foxmail.com。

## 【\*\*通信作者】

蒋 权(1989-),男,博士,讲师,硕士研究生导师,主要从事智能机器人系统研究,E-mail:jiangquan@gxun.edu.cn。

## 【引用本文】

李惠,蒋权. 基于 YOLOv3 及迁移学习的河道船舶目标检测算法[J]. 广西科学院学报, 2023, 39(3): 331-339.

LI H, JIANG Q. A River Ship Target Detection Algorithm Based on YOLOv3 and Transfer Learning [J]. Journal of Guangxi Academy of Sciences, 2023, 39(3): 331-339.

理,以减少采砂工作对河道的影响,进而优化河道管理效果。另外,陈茂山等<sup>[8]</sup>提出要重点围绕“人、砂、船”等要素和“采、运、销”等环节进行监管。近年来,监管部门在港口、沿海以及江河沿岸布置了大量视频摄像头用来人工监控相关船舶状态。同时,基于视频监控的船舶目标检测算法在海上养殖、重要区域水下文化遗产保护、港口监控和海上设施保护等领域具有重要的用途<sup>[9]</sup>。要突破采砂监管中的痛点、难点,可以引入深度学习领域的图像识别、目标检测和追踪等新技术,结合视频监控监测目标船舶以提升监管效能<sup>[10]</sup>。研究人员证明了神经网络检测算法比传统的目标检测算法工作得更好<sup>[11,12]</sup>。虽然目前业界提出的一些基于 You Only Look Once version 3 (YOLOv3)的检测算法已经产生了良好的结果,但仍存在缺少目标和精度偏低<sup>[13,14]</sup>等问题。

当前,党和政府十分重视非法采砂的监管工作,为加大非法采砂打击力度,我国法律明确指出,无证采砂适用“非法采矿罪”定罪处罚<sup>[15]</sup>。尽管如此,主要依靠人力的非法采砂监管方式难以及时发现和保存非法采砂的证据,从而导致执法工作难以开展。为解决执法管理部门遇到的技术困难,帮助实现对违法采砂行为的更有力监管,优化河道散体物料运输调度,本文以广西重要河道监控区域摄像头中的采砂运砂船舶数据为驱动,基于 YOLOv3 算法并结合迁移学习开发了一种针对河道采砂船目标的智能检测算法,拟为监管人员快速分辨河道上的非法采砂行为提供智能化技术支撑。

## 1 相关工作

为了有效监控和规范采砂的作业行为,研究者采用不同的方法设计了监控系统。鲍凯等<sup>[16]</sup>设计了一种基于 GPRS 的采砂船舶监控系统,该系统运用了北斗导航定位技术,主要根据采砂船的位置和时间来判断采砂船采砂作业是否合法。付化胜<sup>[17]</sup>和索

然<sup>[18]</sup>基于无人机测绘技术分别提出一种采砂行为监测方法,这类方法通过低空航拍测量快速获取采砂区域影像,再制作成高精度影像和三维模型,从而确定疑似非法采砂点的位置、面积、堆砂量等信息。这种方法需要借助无人机获取数据,成本较高,难以常态化进行。

近年来,基于机器学习和深度学习的目标检测识别技术取得了飞速的发展<sup>[9]</sup>,并在医学图像分析、遥感图像分类、交通视频监控等诸多方面应用广泛<sup>[10]</sup>。随着目标检测技术的进一步发展和完善,人们把目标检测技术应用在船舶采砂方面。针对视频监控下的船舶检测识别主要分为视频预处理、兴趣区域提取、目标分割、特征提取和目标分类等多个阶段,学习过程较简单,研究成本较低,应用部署容易。通过多层卷积识别、提取、分类图像的各种特征,再不断进行训练和学习就可以得到深度学习模型。

## 2 船舶目标检测算法及评价指标

### 2.1 船舶目标检测算法的实现

在实现基于 YOLOv3 的船舶目标检测算法时,通过在训练模型时引入迁移学习来解决河道船舶数据集样本量小、规模有限的问题。在船舶目标检测算法实现过程中,首先采用 COCO 数据集训练初始的 YOLOv3 算法,从而得到预训练权重并保存到本地。然后通过加载该模型的权重来初始化主干特征提取网络的权值,利用自建的船舶数据集对 YOLOv3 算法中的网络模型进行微调和训练,进而训练出特征提取效果优秀的船舶检测模型。这样能使模型拥有大规模数据集的特征提取能力,从而加快模型收敛。该算法具体实现过程见图 1。

在训练船舶目标检测识别算法之前需要进行采集视频图像、预处理视频图像、标注船舶数据等操作。一般来说,视频图像预处理主要包括图像格式转化、视频稳像(防抖动处理)、去雾处理(阴雨天或雾天的

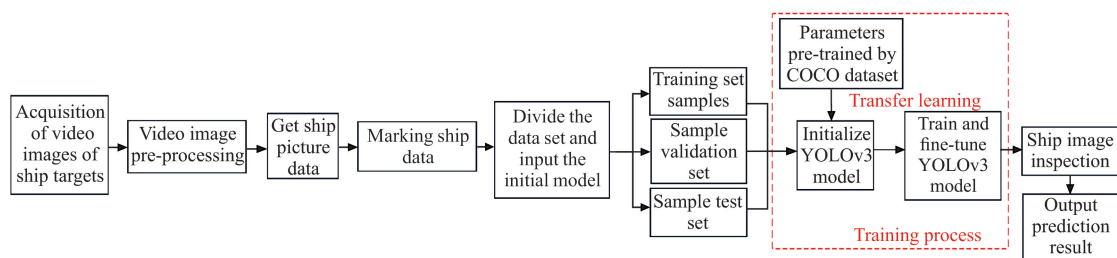


图 1 船舶目标检测算法实现流程

Fig. 1 The flowchart of the ship targets detection algorithm

视频清晰化)和图像滤波等<sup>[15]</sup>。在视频图像预处理阶段,对船舶视频进行加速、抽帧、图像格式转化等处理并保留船舶目标清晰且状态多样的高质量图片。获取船舶图像后对其进行合理标注,这样制作的数据集用于训练可使模型具有更好的鲁棒性。

图像特征是检测船舶目标的重要标识<sup>[9]</sup>,预处理得到的图片数据经由YOLOv3归一化操作后初始化为 $416 \times 416$ 大小的尺寸,平均划分为 $S \times S$ ( $S$ 代表目标检测所有步长的乘积)个单元格,每个单元格有3个锚框负责预测3个边界框。算法训练过程中通过特征提取网络来提取和学习船舶特征。

YOLOv3的网络共包含了107个层,由一个75层的Darknet-53网络和一个三尺度的YOLO层组成。随着网络的加深,特征的语义信息更加丰富<sup>[19]</sup>。其中的Darknet-53网络采用残差网络的连接思想,在网络结构很深的情况下仍能收敛<sup>[20]</sup>。YOLOv3利用3个不同尺度的特征图来检测目标,这样不仅可以加快模型的识别速度,还能显著提高小目标的识别准确率和分类准确率<sup>[21]</sup>。对船舶目标的检测结果则经过置信度和非极大抑制筛选后保留,并通过画外接矩形的形式在图像中表现出来,在此基础上预测目标的类别。

## 2.2 迁移学习

迁移学习是一种将已知领域的知识应用在目标领域的方法,将已经训练好的模型作为初始化模型,在此基础上对目标数据集进行训练,可以避免因数据集过小而出现拟合的现象,提高模型的训练速度和性能。该方法主要分为四大类别:基于实例的迁移学习、基于特征的迁移学习、基于模型的迁移学习和基于关系的迁移学习。本文选用基于模型的迁移学习,这类迁移学习方法可以共享原任务的模型或参数,适用于大多数迁移任务。

在迁移学习中,被学习的领域成为源域(source domain),待解决问题的领域成为目标域(target do-

表1 船舶图片类型及数量

Table 1 Type and number of ship pictures

船舶类型 Ship type	标签名称 Label name	标签数量 Number of labels	(训练集+验证集)数量 Number of (training sets+ validation sets)	测试集数量 Number of test sets
Dredger	Dredger	1 280	1 152	128
Sand ship	SandShip	3 730	3 357	373
Dredging platform	DredgingPlatform	750	675	75
Passenger ship	PassengerShip	550	495	55

main),公式如下:

$$D(s) = \{X_s, P(X_s)\}, \quad (1)$$

$$D(t) = \{X_t, P(X_t)\}, \quad (2)$$

其中, $s$ 表示源点, $t$ 表示目标点, $D(s)$ 为源域, $D(t)$ 为目标域, $X_s$ 是源域的特征空间, $X_t$ 是目标域的特征空间, $P(X_s)$ 是与 $X_s$ 对应的边际概率分布, $P(X_t)$ 是与 $X_t$ 对应的边际概率分布。

深度学习模型的检测效果与训练样本数密切相关,高质量的深度学习模型建立在大量训练样本的基础上。因此,本文将官方用于目标检测的COCO数据集作为源域,将手工标注的船舶数据集作为目标域,并基于迁移学习实现基于YOLOv3算法的网络模型的预训练、训练以及优化等,迁移学习过程如图2所示。

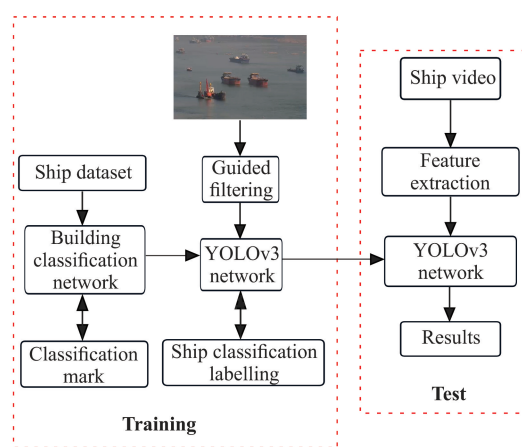


图2 基于预训练模型的迁移学习

Fig. 2 Transfer learning based on pre-trained model

## 2.3 船舶监控视频数据源

本文参照Pascal VOC数据集格式制作用于驱动船舶目标检测网络模型训练的河道船舶数据集<sup>[22]</sup>,整个数据集中的船舶类型、标签名称、标签数量以及划分的船舶数据集大小如表1所示。实验中的数据划分为训练集、验证集和测试集,并指定(训练集+验证集):测试集=9:1。

采砂作业频发的河道场景中,采砂船和运砂船较多,挖砂平台次之,客船数量最少。从表 1 可知,整理好的船舶图像数据集中,采砂船共有 1 280 艘,运砂船有 3 730 艘,挖砂平台有 750 个,客船有 550 艘。

## 2.4 模型训练

深度神经网络只有经过大量的样本训练,才能拥有较好的效果,在目标数据集很小的情况下,直接训练 YOLOv3 目标检测模型难以获得很好的检测效果<sup>[23]</sup>。迁移学习是机器学习中的一种方法,指将一个预训练的模型重新应用到另一个任务中,因此,本文在利用 YOLOv3 算法训练自建的船舶数据集时,引入了迁移学习的思想。对 COCO 数据集在原 YOLOv3 上预训练好的目标检测模型进行知识迁移,训练自建的船舶数据集,得到新的 YOLOv3 船舶检测模型。

在迁移学习 YOLOv3 模型的训练过程中,从 3 个方面、3 个尺度共 9 个维度构建损失函数,并通过反向迭代更新网络参数。其中,损失计算公式如公式(3)~(6)所示。

$$L_{P0} = L_{P0\_obj} + L_{P0\_location} + L_{P0\_cls}, \quad (3)$$

$$L_{P1} = L_{P1\_obj} + L_{P1\_location} + L_{P1\_cls}, \quad (4)$$

$$L_{P2} = L_{P2\_obj} + L_{P2\_location} + L_{P2\_cls}, \quad (5)$$

$$L_{total} = L_{P0} + L_{P1} + L_{P2}, \quad (6)$$

式中,下标  $P0$ 、 $P1$ 、 $P2$  分别代表  $52 \times 52$ 、 $26 \times 26$ 、 $13 \times 13$  的特征图;  $L_{P0}$ 、 $L_{P1}$ 、 $L_{P2}$  分别代表在  $52 \times 52$ 、 $26 \times 26$ 、 $13 \times 13$  这 3 个特征图上判断有无目标、定位目标位置及分类目标的总损失量;  $L_{P0\_obj}$ 、 $L_{P1\_obj}$ 、 $L_{P2\_obj}$  分别为不同尺寸特征图中用来判断是否包含物体的损失;  $L_{P0\_location}$ 、 $L_{P1\_location}$ 、 $L_{P2\_location}$  分别代表不同尺寸特征图上用以识别与确定目标位置的损失;  $L_{P0\_cls}$ 、 $L_{P1\_cls}$ 、 $L_{P2\_cls}$  分别表示不同尺寸特征图中用以预测目标所属类别的损失;  $L_{total}$  代表在 3 个不同特征图上识别和分类目标过程中所消耗的总损失。

实验中将船舶数据集的训练分为两个阶段:①冻结训练阶段,即冻结 YOLOv3 模型的主干特征提取网络,仅对主干特征提取网络 Darknet-53 以外的其他网络进行训练,包括 FPN、卷积层和 YOLO Head 预测层等。这些网络的参数通过迁移学习 COCO 权重文件进行初始化。冻结训练以 0.001 的学习率迭代 50 轮,使网络快速收敛,从而具备一定的船舶目标识别与分类的能力;②解冻训练阶段,此时不再冻结模型的主干网络,特征提取网络中的所有参数会发生

改变。解冻训练以 0.000 1 的学习率迭代 50 轮,优化网络结构,逐步降低损失值,使模型具有更好的检测性能。

网络训练过程中,损失函数下降情况如图 3 所示。从图 3 可以看出,通过迁移学习训练的模型损失下降速度很快,约训练 60 轮后趋于收敛且损失值趋于 0。由此可知,迁移学习有助于加速网络收敛并提高目标检测的准确率。

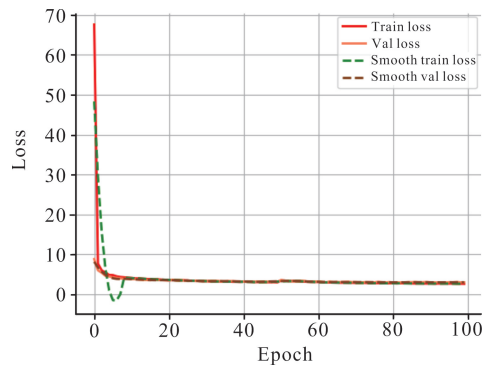


图 3 损失下降曲线

Fig. 3 Loss descent curve

## 2.5 船舶目标检测模型的评价指标

为了进一步评价本文模型的检测精度,选取精确率(Precision,  $P$ )、召回率(Recall,  $R$ )、精确率和召回率的调和平均数( $F1$ -score,  $F1$ )<sup>[24]</sup>、平均精度(Average Precision,  $AP$ )和平均精度值(mean Average Precision,  $mAP$ )指标来评价。精确率和召回率的表达式分别为公式(7)和公式(8)。

$$P = \frac{N_{tp}}{N_{tp} + N_{fp}}, \quad (7)$$

$$R = \frac{N_{tp}}{N_{tp} + N_{fn}}, \quad (8)$$

式中,  $N_{tp}$  为正确识别出的船舶数量,  $N_{fp}$  为预测框将其他物体错误预测成船舶的数量,  $N_{fn}$  为将船舶错误识别成其他物体的数量。

$F1$  常作为多分类问题的最终指标。由公式(9)计算类别  $k$  的  $F1$  分数,用  $F1_k$  表示。

$$F1_k = \frac{2 \times R_k \times P_k}{R_k + P_k}. \quad (9)$$

类别  $i$  的平均精度  $AP_i$  是类别  $i$  的精度-召回曲线  $P_i(R)$  所包围的区域,实际计算时取不同召回率下的最高精确值的平均值,表达式如公式(10)所示:

$$AP_i = \int_0^1 P_i(R) dR. \quad (10)$$

$mAP$  是所有类别平均精度( $AP$ )的平均值,用于衡量模型在所有类别上的训练效果。这里用

$mAP$  来计算船舶检测模型中在指定的交并比 (Intersection Over Union, IOU) 下每个船舶类别对应的平均精度的均值, 记作  $mAP_{IOU}$ , 其计算方法如公式 (11) 所示。

$$mAP_{IOU} = \frac{1}{N} \sum_i^N AP_i, \quad (11)$$

式中  $N$  表示类别的总数。根据公式 (7) 至 (11), 计算出优化后的 YOLOv3 模型目标分类和检测的准确性。

### 3 实验与结果分析

#### 3.1 实验平台与模型参数

所有实验均在 GPU 硬件加速平台上完成, 具体的实验平台配置及实验时使用的模型参数分别如表 2、表 3 所示。

表 2 实验平台配置

Table 2 Experimental platform configuration

名称 Name	配置 Configuration
Operating system	Windows 10
CPU	Intel(R) Core(TM) i9-10920X
GPU	NVIDIA GeForce RTX 3060 Ti
GPU acceleration environment	CUDA 11.2 + cuDNN 8.2
Deep learning framework	torch 1.7.1 + cu110

本实验设置的模型训练参数参考了刘博等<sup>[25]</sup>在改进 YOLOv3 算法中的冻结参数, 并在此基础上进行了微调和测试。如果 GPU 显存不足, 可以调整  $batch\_size$  的大小, 但受 BatchNorm 层的影响,  $batch\_size$  最小为 2。随着  $batch\_size$  的增大, 更高

表 4 模型性能评价

Table 4 Model performance evaluation

船舶类别 Ship label	精确率/% Precision/%	召回率/% Recall/%	调和分数/% F1/%	类别精度/% AP/%	类别平均精度/% $mAP$ /%	阈值 score_threshold
Dredger	95.69	86.72	91.00	97.48		
Sand ship	93.50	88.74	91.00	96.62	98.00	0.50
Dredging platform	96.05	97.33	97.00	98.43		
Passenger ship	98.21	100.00	99.00	99.48		

表 3 模型参数配置

Table 3 Model parameter configuration

参数 Parameter	配置 Configuration
Number of GPUs used	1
Number of images processed by GPU each time	1
Freeze_epoch	50
UnFreeze_epoch	50
Freeze learning rate (Freeze_lr)	0.001
UnFreeze learning rate (UnFreeze_batch_size)	0.000 1
Dimension of input	416 × 416 × 3

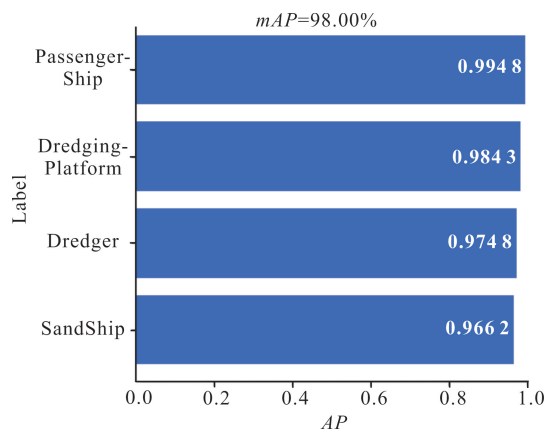
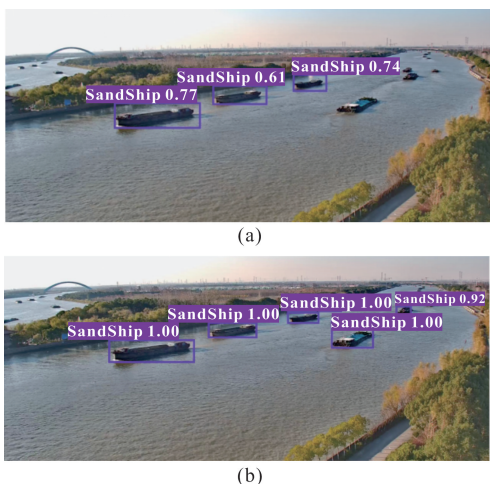
的学习率会带来更好的效果。因此, 在 GPU 显存允许的情况下, 设置冻结训练时的  $batch\_size$  (Freeze\_batch\_size) 为 8, 冻结训练时的学习率 Freeze\_lr 为 0.001, 冻结训练的轮数 Freeze\_Epoch 为 50; 设置解冻训练时的  $batch\_size$  (UnFreeze\_batch\_size) 为 4, 解冻训练时的学习率 UnFreeze\_lr 为 0.000 1, 解冻训练的轮数 UnFreeze\_Epoch 的值为 100。

#### 3.2 实验结果

本文采用目标检测常用的评价指标来评估模型的性能, 各类目标的指标结果见表 4。该模型的  $mAP$  统计结果如图 4 所示。

由图 4 可以得知, 各类的  $AP$  越高, 相应的  $mAP$  也越高, 说明模型的训练效果越好。

为了验证迁移学习的有效性和必要性, 图 5 对比了原 YOLOv3 模型 [图 5(a)] 和引入迁移学习的 YOLOv3 模型 [图 5(b)] 检测同一运砂船图像的效果。从图 5 可以看出, 基于迁移学习训练的模型能够检测出更多的目标且候选框分数更高。

图4 模型的  $mAP$  值Fig. 4  $mAP$  values of the model

(a)

(b)

图5 不同模型检测同一图像的对比结果

Fig. 5 Comparison results of detecting the same image by different models

网络训练结束后,随机选取 232 张测试集数据测试基于 YOLOv3 及迁移学习的模型性能。其中,用于测试的图片中标注的各类船舶数量统计结果如图 6 所示。

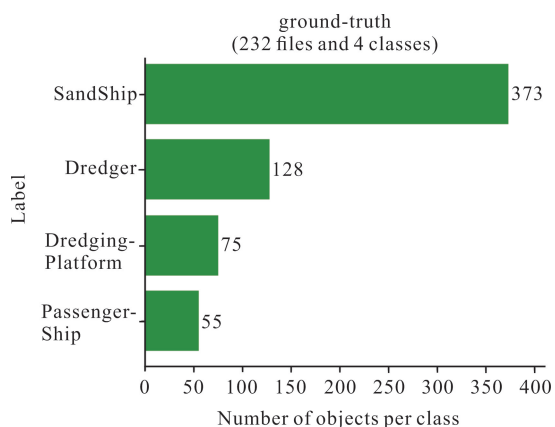
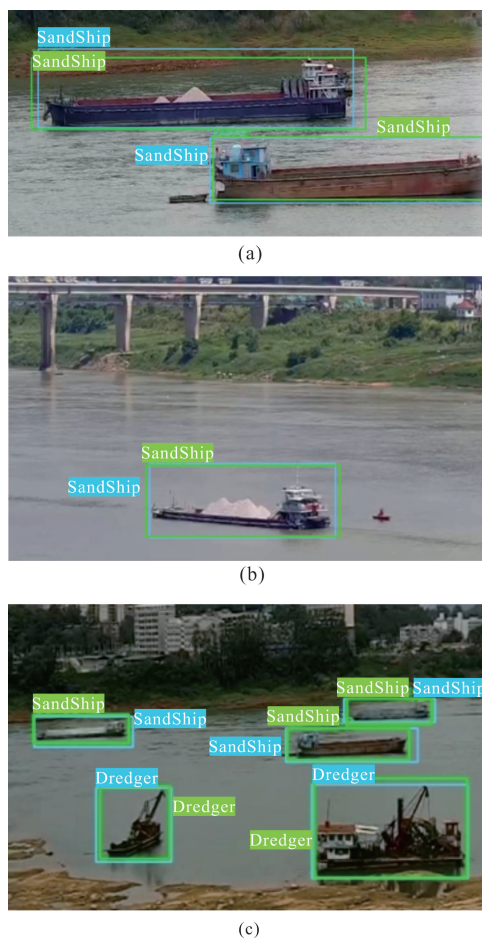


图6 用于测试的各类船舶数量

Fig. 6 Number of each type of ships used for testing

本文选用的置信度水平趋于 1,这是由于当训练次数足够多且不发生过拟合时,用该模型可以准确地检测出指定类别的目标,即召回率为 100%,如表 4 中的客船目标的召回率。因此,本实验中将原算法模型中的分类置信度调高至 0.75,使得输出的检测框概率值更大,从而提高了算法的可信度。只有当设置分类置信度为 0.75 且非极大抑制交并比阈值  $nms_{IOU} = 0.3$  时,预测框的识别结果才会被保留下来。通过对比预测框和真实框的重合度,以及计算准确度、召回率、 $mAP$  值等来评价模型的检测性能。其中,预测框和真实框的对比结果如图 7 所示。部分具有代表性的图片及视频检测结果见图 8。



(a)

(b)

(b)

(c)

图7 部分船舶的 ground-truth 和检测框的对比结果

Fig. 7 Comparison results of ground-truth and detection frames of some ships

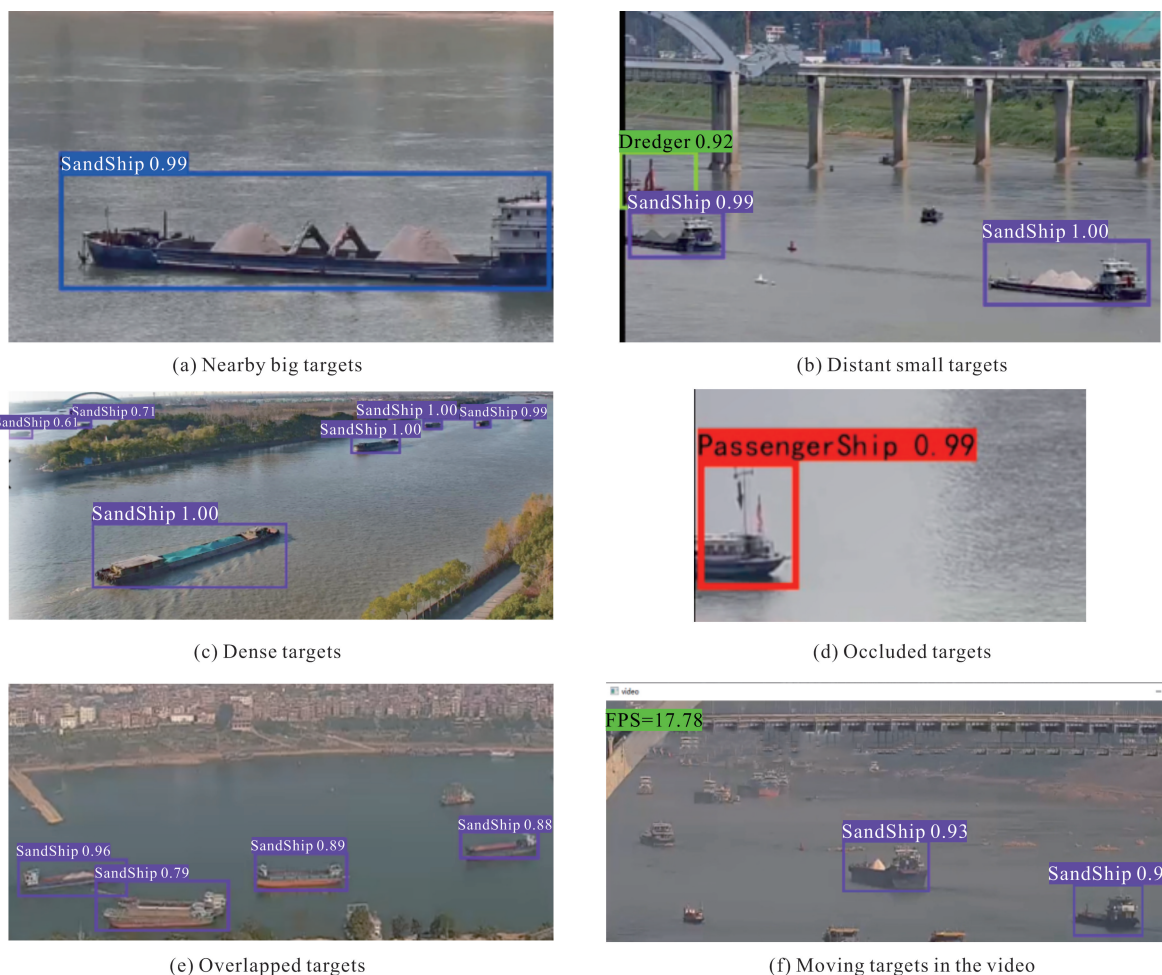


图8 部分船舶图片及视频检测结果

Fig. 8 Part of the ship pictures and video inspection results

结果表明,在分类置信度的限制下,该算法得到的目标候选框分数可以达到0.9以上,同时可以识别不同特征的运砂船(图8),这说明其具有较强的鲁棒性和泛化能力。该模型能识别不同距离下不同大小的目标,如图8(a)和图8(b)中的采砂船和运砂船。更重要的是,它还能检测出所有密集的小目标,如图7(c)中的每条运砂船;对半遮挡的目标也能正确检测,如图8(d)中的客船;对于重叠部分较多的目标,则被检测为一个目标[图8(e)]。由于本文在标注数据样本时,默认将重叠率较高的多个目标视为一个目标,所以这种结果的出现是合理的。

速度也是评估目标检测算法优劣性的一个重要参数,评估速度的常用指标是每秒帧率(Frame Per Second, FPS),即每秒内可以处理的图片数量。在视频监控场景中往往需要实时检测,这就对模型的检测速度提出了高要求。因此,在保证准确度的情况下,检测速度越快的模型越优秀。如图8(f)所示,该方法的FPS可以达到17.78,基本能够满足实际需求。

总的来说,该方法能够从不同角度较准确地识别不同大小和类型的目标,具有较好的泛化能力和较快的速度,能够满足实际内河船舶目标检测业务的需要。

#### 4 结论

本文首先将广西重要河道上的采砂运砂船舶影像数据进行预处理,构造高质量船舶数据集;然后在模型训练过程中引入迁移学习策略,将大型COCO图片数据集中已经预训练好的网络参数用于YOLOv3算法的模型训练;再利用自建的船舶数据集对模型进行训练和微调,进而得到河道采砂船等重点目标检测的算法模型。实验结果表明,该模型在速度及检测精度方面性能较好,并且不易受背景因素的干扰,通用性强。船舶识别的全类平均准确率达到98.00%,在GPU上的图片检测速度为0.0286 s/张,实时检测视频的FPS可达到17.78,能够满足视频监控的实时性要求。另外,由于训练所用船舶图像

数据集主要是从特定区域视频监控角度采集的,这样驱动得到的模型也能更加满足特定区域真实采砂监控场景的检测业务要求,为实现区域智慧视频监控以及非法采砂监管提供可靠的技术支撑。

本文在对比真实框和预测框的检测效果时,小部分船舶目标存在预测框错位,即交并比 IOU 较低的问题。这需要进一步解决识别物体位置精准性差的问题,可通过优化网络结构增强对船舶轮廓的识别与定位,并增加对船舶轮廓、船舶名号及船标等细节的信息检测。同时,基于检测的多目标跟踪是计算机视觉领域的研究热点<sup>[26]</sup>,今后将在检测的基础上继续深入研究以实现更优的多目标跟踪。

#### 参考文献

- [1] 孙晓伟.我国河道采砂管理研究[D].大连:大连海事大学,2016.
- [2] 吕奕霖.智慧河道采砂监管平台系统的设计与实现[D].郑州:华北水利水电大学,2019.
- [3] 张建新,黄月华,黄敬华.项目法人顺利组建 百色水利枢纽通航设施建设前期工作取得突破性进展[J].珠江水运,2020,510(14):22-23.
- [4] 马财林.西部陆海新通道中广西内河水运发展分析[J].中国水运,2021(7):9-11.
- [5] 郭超,王玉,姚仕明,等.长江干流河道采砂规划实施情况评估:以《长江上游干流宜宾以下河道采砂规划(2015-2019年)》为例[J].人民长江,2022,53(3):8-12,20.
- [6] 丁继勇,林欣,卢晓丹,等.河道采砂管理问题及其研究进展[J].水利水电科技进展,2021,41(4):81-88.
- [7] 耿金荣.河道采砂对河道的影响及管理对策[J].黑龙江水利科技,2020,48(5):123-125.
- [8] 陈茂山,吴强,王晓娟,等.河道采砂管理现状与立法建议[J].水利发展研究,2019,19(7):1-5.
- [9] 毕振波,张世友,杨花,等.基于浅层机器学习的视频监控船舶检测综述[J].系统仿真学报,2021,33(12):2792-2807.
- [10] 汤文华,陈灿斌,向舒华,等.基于深度学习的图像识别技术在非法采砂监管中的应用[J].中国农村水利水电,2021(5):108-112.
- [11] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1-9.
- [12] SCHMIDHUBER J. Deep learning in neural networks: an overview [J]. Neural Networks, 2015, 61: 85-117.
- [13] HASS F S, JOKAR A J. Deep learning for detecting and classifying ocean objects: application of YOLOv3 for iceberg-ship discrimination [J]. ISPRS International Journal of Geo-Information, 2020, 9(12): 758-772.
- [14] YANG J, LEONIDAS L, MUMTAZ F, et al. Ship detection and tracking in inland waterways using improved YOLOv3 and deep SORT [J]. Symmetry, 2021, 13(2): 308-326.
- [15] 付永冲,柯敏毅.采砂船舶管理系统的意义和设计[J].信息通信,2018(12):243-244.
- [16] 鲍凯,潘洪军,亓常松.基于GPRS的采砂船舶监控系统设计[J].人民长江,2017,48(18):20-22.
- [17] 付化胜.无人机测绘技术于河道采砂监管中的应用[J].现代信息科技,2020,4(20):42-46.
- [18] 索然.河道采砂监管中无人机测绘技术的应用[J].黑龙江水利科技,2021,49(11):156-158,202.
- [19] 盛明伟,李俊,秦洪德,等.基于改进YOLOv3的船舶目标检测算法[J].导航与控制,2021,20(2):95-109.
- [20] 刘博,王胜正,赵建森,等.基于Darknet网络和YOLOv3算法的船舶跟踪识别[J].计算机应用,2019,39(6):1663-1668.
- [21] 王飞,刘梦婷,刘雪芹,等.基于YOLOv3深度学习的海雾气象条件下海上船只实时检测[J].海洋科学,2020,44(8):197-204.
- [22] 陈连凯.基于深度学习的水上视频船舶目标检测方法研究[D].镇江:江苏科技大学,2020.
- [23] 杨耘,李龙威,高思岩,等.基于YOLOv3网络训练优化的高分辨率遥感影像目标检测[J].激光与光电子学进展,2021,58(16):155-161.
- [24] WANG Y T, YANG L, SONG X, et al. A multi-feature ensemble learning classification method for ship classification with space-based AIS data [J]. Applied Sciences, 2021, 11(21): 10336-10348.
- [25] 刘博,李艳武.基于改进的YOLOv3算法研究[J].现代计算机,2022,28(16):53-56.
- [26] XU J W, LU X B, ZHANG C. The research of multi-target tracking based on improved YOLOv3 [J]. Journal of Physics: Conference Series, 2021, 1976 (1): 012023.



## A River Ship Target Detection Algorithm Based on YOLOv3 and Transfer Learning

LI Hui<sup>1</sup>, JIANG Quan<sup>2\* \* \*</sup>

(1. School of Electronic Information, Guangxi Minzu University, Nanning, Guangxi, 530006, China; 2. School of Artificial Intelligence, Guangxi Minzu University, Nanning, Guangxi, 530006, China)

**Abstract:** In order to strengthen the monitoring and tracking of granular material transportation ships in river monitoring video images, so as to assist in the realization of intelligent and efficient river sand mining supervision and granular material scheduling, a river granular material ship target detection algorithm is proposed based on You Only Look Once version 3 (YOLOv3) algorithm and transfer learning. Firstly, the COCO dataset is used to train the initial YOLOv3 algorithm, and the pre-training weight of the model is obtained. Then, the image data of sand mining and sand transportation ships collected from the monitoring equipment around the important rivers in Guangxi are processed to obtain a high-quality ship dataset. Finally, driven by this dataset, the transfer learning pre-training weight is used to train the YOLOv3 detection model for key targets such as river sand mining ships. The model uses Darknet-53 as the backbone network and integrates multi-scale feature maps to achieve the detection of small, medium and large targets. The experimental results show that the average accuracy value and detection speed of the algorithm on the test set reach 98.00% and 17.78 fps, respectively, which is of practical significance for improving the supervision efficiency of river sand mining and realizing the intelligent scheduling of granular materials.

**Key words:** YOLOv3; transfer learning; granular material ship; target detection; sand mining supervision

责任编辑: 陆雁, 陈少凡

(上接第 330 页 Continued from page 330)

## Beijing Opera Binary Classification Simulation Based on RF-BiLSTM-Attention Music Classification Method

GONG Yicheng<sup>1,2\* \* \*</sup>, LIU Qing<sup>1,2</sup>, XIAO Haoyi<sup>1</sup>

(1. School of Science College, Wuhan University of Science and Technology, Wuhan, Hubei, 430065, China; 2. Hubei Province Key Laboratory of Systems Science in Metallurgical Process (Wuhan University of Science and Technology), Wuhan, Hubei, 430081, China)

**Abstract:** In order to popularize Beijing Opera, a music classification method of RF-BiLSTM-Attention was proposed based on Random Forest (RF) coupled Attention mechanism and Bidirectional Long Short Term Memory (BiLSTM) network, which was used to classify Beijing Opera and other types of music (hereinafter referred to as 'Beijing Opera binary classification'). Firstly, all spectral features of music were extracted, and important features were selected by RF. Then, the attention layer was embedded the hidden layer and the output layer of the BiLSTM network to classify, train and predict the data. Using 1 500 pieces of music from the popular music platform and GTZAN dataset for Beijing Opera binary classification experiments, the effect of RF on 9 models such as Recurrent Neural Network (RNN), Long Short Term Memory (LSTM) and BiLSTM was compared. The results showed that the classification accuracy of RF-BiLSTM-Attention was 89.00%, with a run time of 33.22 s. Compared with the best performing RF-BiLSTM model in the simple model, the classification accuracy was improved by 3.33%, and the run time was reduced by 40.54%. Compared with the original BiLSTM-Attention model, the classification accuracy was increased by 6.33%, and the run time was shortened by 96.89%. Compared with traditional audio classification work, this article considers the binary classification of Beijing Opera, which plays a good role in promoting Beijing Opera.

**Key words:** Beijing Opera; BiLSTM network; attention mechanism; random forest; binary classification

责任编辑: 梁晓