Journal of Computer Applications
计算机应用,2020,40(9):2561 - 2570
ISSN 1001-9081
CODEN JYIIDU
2020-09-10
http://www.joca.cn

文章编号:1001-9081(2020)09-2561-10
DOI:10.11772/j.issn.1001-9081.2020010097

# 复杂场景下基于增强YOLOv3的船舶目标检测

聂　鑫[1,2],刘　文[1,2*],吴　巍[3]
(1.武汉理工大学 航运学院,武汉 430063;　2.内河航运技术湖北省重点实验室(武汉理工大学),武汉 430063;
3.武汉理工大学 信息工程学院,武汉 430070)
(*通信作者电子邮箱wenliu@whut.edu.cn)

**摘　要**:为提升水上交通安全监管的智能化水平,进一步提高基于深度学习的船舶目标检测算法的定位精度和检测准确率,在传统YOLOv3算法基础上,提出用于船舶目标检测的增强YOLOv3算法。首先,在网络预测层引入预测框不确定性回归,以预测边界框的不确定性信息;然后,使用负对数似然函数和改进的二值交叉熵函数重新设计损失函数;其次,针对船舶形状使用$K$均值聚类算法重新设计先验锚框尺寸并平均分配到对应预测尺度;在网络训练阶段,使用数据增强策略扩充训练样本数量;最后,使用加入高斯软阈值函数的非极大值抑制(NMS)算法对预测框进行后处理。对各种改进方法和不同目标检测算法在真实海事视频监控数据集上进行对比实验。实验结果显示,与传统YOLOv3算法相比,带有预测框不确定性信息的YOLOv3算法的假正样本(FP)数量降低了35.42%,真正样本(TP)数量提高了1.83%,所以提高了准确率;增强YOLOv3算法在船舶图像上的平均准确率均值(mAP)达到87.74%,与传统YOLOv3算法和Faster R-CNN算法相比分别提高了24.12%和23.53%;所提算法的每秒钟检测图像数量达到30.70张,满足实时检测的要求。实验结果表明,所提算法在雾天和低照度等不良天气条件与复杂通航背景下,均能实现船舶目标的高精度稳定实时检测。

**关键词**:水路运输;目标检测;YOLOv3;船舶;深度学习
**中图分类号**:TP391.41　　**文献标志码**:A

## Ship detection based on enhanced YOLOv3 under complex environments

NIE Xin[1,2], LIU Wen[1,2*], WU Wei[3]
(1. *School of Navigation, Wuhan University of Technology, Wuhan Hubei* 430063, *China*;
2. *Hubei Key Laboratory of Inland Navigation Technology (Wuhan University of Technology), Wuhan Hubei* 430063, *China*;
3. *School of Information Engineering, Wuhan University of Technology, Wuhan Hubei* 430070, *China*)

www.joca.cn

**Abstract:** In order to improve the intelligence level of waterway traffic safety supervision, and further improve the positioning precision and detection accuracy in the ship detection algorithms based on deep learning, based on the traditional YOLOv3, an enhanced YOLOv3 algorithm for ship detection was proposed. First, the prediction box uncertain regression was introduced in the network prediction layer in order to predict the uncertainty information of bounding box. Second, the negative logarithm likelihood function and improved binary cross entropy function were used to redesign the loss function. Then, the *K*-means clustering algorithm was used to redesign the scales of prior anchor boxes according to the shape of ship, and prior anchor boxes were evenly distributed to the corresponding prediction scales. During training phase, the data augmentation strategy was used to expand the number of training samples. Finally, the Non-Maximum Suppression (NMS) algorithm with Gaussian soft threshold function was used to post-process the prediction boxes. The comparison experiments of various improved methods and different object detection algorithms were conducted on real maritime video surveillance dataset. Experimental results show that, compared to the traditional YOLOv3 algorithm, the YOLOv3 algorithm with prediction box uncertainty information has the number of False Positives (FP) reduced by 35.42%, and the number of True Positives (TP) increased by 1.83%, thus improving the accuracy. The mean Average Precision (mAP) of the enhanced YOLOv3 algorithm on ship images reaches 87.74%, which is improved by 24.12% and 23.53% respectively compared to those of the traditional YOLOv3 algorithm and Faster R-CNN algorithm. The proposed algorithm has the number of images detected per second reaches 30.70, meeting the requirement of real-time detection. Experimental results indicate that the proposed algorithm can achieve high-precision, robust and real-time detection of ships under adverse weather and conditions such as fog weather and low-light condition as well as the complex navigation backgrounds.

**Key words:** waterway transportation; object detection; YOLOv3 (You Only Look Once v3); ship; deep learning

**收稿日期**:2020-02-04;**修回日期**:2020-04-29;**录用日期**:2020-05-08。
**基金项目**:国家自然科学基金资助项目(516091295);中央高校基本科研业务费专项(2019-zy-215)。
**作者简介**:聂鑫(1995—),男,河南巩义人,硕士研究生,CCF会员,主要研究方向:计算机视觉;　刘文(1987—),男,湖北孝感人,副教授,博士,CCF会员,主要研究方向:计算机视觉、海事信息感知与处理;　吴巍(1971—),男,湖北黄石人,副教授,博士,主要研究方向:信息处理、图像处理、模式识别。

## 0 引言

随着水上交通运输业的蓬勃发展,水上交通安全形势面临严峻考验。船舶作为水路运输的重要载体,准确识别船舶类型和检测船舶位置对感知水路交通态势、保证船舶航行安全和水上违法行为预警等具有重要意义。实现船舶目标的高效准确检测是后续进行船舶行为识别与轨迹跟踪等高级视觉任务的基础。因通航环境复杂、水面反光和易受不良天气影响等因素,与道路车辆检测相比,水上船舶目标检测更具挑战性。

传统目标检测方法主要有帧间差分法和背景建模法等。帧间差分法依赖于从视频图像前后帧之间的差分运算获取目标轮廓,无法提取目标的完整区域,由于船舶运动速度较慢,难以选取合适的时间间隔,不能保证船舶目标检测的准确性和稳定性。背景建模法通过将当前帧与背景模型进行比较,通过阈值法区分像素是否属于前景目标。由于水面波纹和船舶尾迹等因素干扰,Prasad 等[1]采用 23 种基于背景建模的传统目标检测算法对水上船舶目标进行实验,均无法准确检测出船舶目标。另外,传统目标检测方法需对目标特征进行人工建模,限制了检测模型泛化能力的提升。因此,亟须一种新的方法提高船舶目标检测能力,并实现目标船舶所属船舶类型的准确识别。

2012 年兴起的卷积神经网络具备强大的目标特征表达和建模能力,通过监督学习的方式,逐层、自动地学习目标从低级到高级的特征表示,实现对物体层次化的抽象和描述,成为目前目标检测任务的首选。基于神经网络的目标检测算法主要分为以 R-CNN(Region-Convolutional Neural Network)系列[2-4]为代表的两次目标检测网络和以 SSD(Single Shot multibox Detector)[5]、YOLO(You Only Look Once)系列[6-8]为代表的单次目标检测网络。两次检测网络首先通过区域建议的方式产生大量候选区域,再将得到的候选区域送入神经网络,预测目标的类别概率和位置信息。单次检测网络把目标检测作为回归问题处理,将目标的位置预测和类别概率预测整合到一个神经网络中,实现端到端的训练,在实际应用中提高了检测速度,但准确率低于两次检测网络。

深度卷积网络已成功应用于水上船舶目标检测领域。王新立等[9]和陈从平等[10]使用 Faster R-CNN 网络分别对港口与内河船舶进行了自动检测,相较于传统方法大幅提高了检测准确率。王言鹏等[11]使用 SSD 网络克服了波浪造成的误检测,并利用迁移学习技术提高网络的泛化能力。赵蓬晖等[12]通过改进特征提取网络提高了 SSD 网络检测船舶目标的准确率。于洋等[13]通过在 YOLOv2 中加入直通层对船舶目标进行检测,与原始 YOLOv2 算法相比准确率有所提高。刘博等[14]在 YOLOv3 算法中加入惩罚机制实现了对船舶目标的识别与跟踪。Kim 等[15]通过将 Faster R-CNN 网络与贝叶斯概率结合实现对船舶视频图像的准确检测。因缺少用于船舶检测的数据集,实验所用数据主要来自从网络获取的低分辨率图像,造成无法准确识别船舶类型的问题,缺乏实际应用的价值。

本文选择兼具检测精度与速度的 YOLOv3 算法作为基础,针对水上船舶目标检测的具体应用场景进行改进,采用真实高分辨率监控视频图像进行实验,验证了所提算法在不同天气条件与复杂通航背景下的有效性。

## 1 YOLOv3 目标检测算法

### 1.1 YOLOv3 目标检测原理

YOLO 系列目标检测网络是单次目标检测网络中最具代表性的网络结构,YOLOv3 是 YOLO 系列的最新改进网络,因为在检测精度上可以与两次目标检测网络相媲美,同时可以达到实时检测速度,所以成为目前最受欢迎的目标检测算法之一。考虑到船舶目标检测在实际应用中需要同时兼顾检测的精度和速度,所以本文以 YOLOv3 为主体,结合船舶目标检测的应用场景进行改进,完成对船舶的位置定位和类型识别。

YOLOv3 网络结构如图 1 所示。

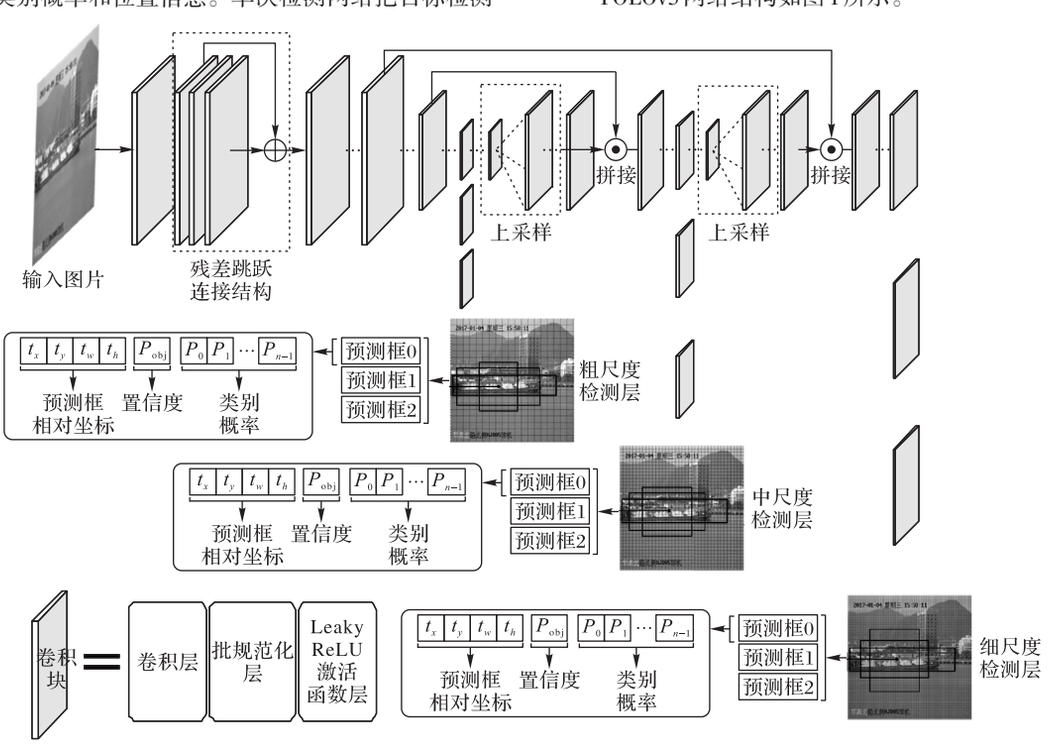



图 1 YOLOv3 算法框架

Fig. 1 YOLOv3 algorithm framework

www.joca.cn

在提取特征之前,输入图像首先被调整为统一的尺寸,然后输入图像被送入 Darknet-53 特征提取网络提取特征。Darknet-53 特征提取网络是包含 53 个卷积层的深度卷积神经网络,可以用来提取输入图像从低级到高级的特征表示,在进行目标检测任务时需要移除最后一层全连接层,所以可以对输入图像做 52 次卷积操作,得到不同深度的特征图。卷积层连续交替使用 3 × 3 和 1 × 1 两种不同尺寸的卷积核,通过 3 × 3 的卷积核使得到的特征图数量增加一倍,便于网络提取更多的目标特征,使用 1 × 1 的卷积核压缩特征图数量,从而减少模型参数,降低网络计算的复杂度。为了可以在训练时使用比较大的学习率以减少训练时间,提高收敛速度,达到稳定训练的目的,在每一个卷积层后使用批规范化层,将输入数据做均值为 0、方差为 1 的归一化处理。在批规范化层之后使用泄露线性修正单元(Leaky Rectified Linear Unit, Leaky ReLU)激活函数进行非线性操作,从而使网络可以应用在非线性模型中。卷积层、批规范化层和激活函数层构成卷积块结构。由于深层的卷积神经网络在训练时容易出现梯度消失的问题,导致在训练过程中检测准确率达到饱和,所以引入残差跳跃连接结构,在加深网络层数的同时解决了网络退化问题。

YOLOv3 算法在预测阶段使用了多尺度预测的策略,通过 Darknet-53 特征提取网络得到的特征图经过几次卷积操作得到粗尺度特征图用于检测大尺度目标,粗尺度特征图经过上采样操作并与 Darknet-53 网络中间层特征图拼接得到中尺度特征图用于检测中尺度目标,中尺度特征图再经过上采样操作并与 Darknet-53 网络更浅层的特征图拼接得到细尺度特征图用于小尺度目标检测,这种类似特征金字塔网络的结构保证了算法对不同尺度的目标都有比较好的预测效果。不同于两次目标检测网络使用的区域提名方法,YOLOv3 网络将特征图划分成 $S \times S$ 的网格单元,如果目标的中心点坐标落在某个网格单元内部,则该网格单元负责这个目标的预测。同时,YOLOv3 算法借鉴 Faster R-CNN 的锚框机制,在每个尺度上预先设置 3 个大小、宽高比不同的先验锚框提高目标检测的准确率。在每一个预测尺度,检测层对于每一个网格单元预测 3 个预测框信息,每个预测框信息包括网络学习到的预测框中心点相对于网格单元左上角坐标的相对坐标 $(t_x, t_y)$、预测框宽、高相对于先验锚框宽、高的缩放尺度 $(t_w, t_h)$,网格单元内有无目标的置信度 $P_{\text{obj}}$ 以及对应 $n$ 个目标类别的类别概率 $P_0, P_1, \cdots, P_{n-1}$。最后,将 3 个预测尺度的预测信息组合并通过非极大值抑制算法得到最终的预测结果。

### 1.2　YOLOv3 存在的问题

虽然经过多次改进的 YOLOv3 网络已经在检测精度和检测速度之间取得了比较好的平衡,且因实现简单,成为许多目标检测任务的首选算法,但是作为一种单次目标检测网络,仍然存在着定位误差大和前景、背景复杂度不平衡的问题。

为了追求检测速度,YOLOv3 算法将目标定位和分类集成在一个卷积神经网络中,同时预测目标的位置坐标和类别信息。然而,由于神经网络中深度特征图包含更高级和抽象的特征信息,适合用于目标分类,但由于丢失了较多的空间信息,对目标定位的效果较差。而浅层特征图的特征更加具体,包含空间信息较多,适合坐标定位但对目标分类的效果并不理想。虽然 YOLOv3 尝试使用深度特征图和浅层特征图拼接来融合不同级别的特征信息,但与两次目标检测算法相比仍然存在着目标定位不准确的问题。

此外,YOLOv3 算法在 3 个预测尺度上一共要预测 10 000 个以上可能的预测框,其中只有一少部分预测框中包含有目标,而大多数预测框中只包含有图像背景信息,导致前景、背景数量极度的不平衡。YOLOv3 网络需要在上万个可能的预测框中搜索到包含有目标的最适合的一个预测框,给网络训练带来很大困难,网络更多地聚焦在容易分类的背景框,造成对前景的分类能力下降。

## 2　增强 YOLOv3 船舶目标检测算法

针对 YOLOv3 网络定位误差大和前景、背景复杂度不平衡的问题,本部分通过加入对预测框的不确定性回归和重新设计损失函数,提出增强 YOLOv3 算法,并以海事监控视频中的船舶目标作为检测对象进行实验。

### 2.1　预测框不确定性回归

由于 YOLOv3 算法同时对目标的位置和类别信息进行回归,所以自然存在高分类准确率和低定位准确率,而且不同于对预测框内是否存在目标的置信度和类别概率的回归都是概率值,可以指示分类的准确程度,YOLOv3 算法对目标位置信息的预测是确定的值,即预测框的中心点相对于负责预测该目标的网格单元左上角坐标的相对坐标 $(t_x, t_y)$ 和预测框的宽、高相对于先验锚框宽、高的缩放尺度 $(t_w, t_h)$,所以无法知道网络对目标位置预测的准确程度。虽然置信度 $P_{\text{obj}}$ 可以反映该预测框内是否有目标,但无法反映出该预测框定位的准确度,即目标有多大部分在预测框内。Choi 等[16]在自动驾驶中通过使用定位的不确定性提高了车辆和行人的检测精度。受 Choi 等的启发,通过在 YOLOv3 算法的预测框输出信息中加入表示预测框准确程度的指标,在网络训练过程中指导网络学习预测更加准确的预测框,从而降低 YOLOv3 算法的定位误差。

因为每一个目标只有一个人工标注的真实边界框(Ground Truth, GT),所以可以使用单高斯模型分别对预测框中心点的相对坐标和归一化尺寸的概率分布进行建模。对于给定的输入 $x$,输出 $y$ 的单高斯模型如式(1)所示:

$$p(y|x) = N\left(y;\ \mu(x), \Sigma(x)\right) \tag{1}$$

其中:$\mu(x)$ 表示均值,$\Sigma(x)$ 表示方差。用均值 $\mu$ 表示预测框的相对位置和尺寸信息,用方差 $\Sigma$ 反映预测框定位的准确程度。所以预测框的每一个坐标和尺寸信息可以分别被建模为均值和方差,即网络学习到的预测框信息可以表示为 $\dot{\mu}_{t_x}$、$\dot{\Sigma}_{t_x}$、$\dot{\mu}_{t_y}$、$\dot{\Sigma}_{t_y}$、$\dot{\mu}_{t_w}$、$\dot{\Sigma}_{t_w}$、$\dot{\mu}_{t_h}$、$\dot{\Sigma}_{t_h}$。沿用 YOLOv3 使用预测框相对于对应网格单元左上角坐标的偏移值和相对于先验锚框的缩放尺度表示预测框的位置和尺寸的方法,需要对网络学习到的预测框相关信息使用 sigmoid 函数进行预处理:

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \in (0, 1) \tag{2}$$

$$\begin{cases} \mu_{t_x} = \sigma(\dot{\mu}_{t_x}), & \Sigma_{t_x} = \sigma(\dot{\Sigma}_{t_x}) \\ \mu_{t_y} = \sigma(\dot{\mu}_{t_y}), & \Sigma_{t_y} = \sigma(\dot{\Sigma}_{t_y}) \\ \mu_{t_w} = \dot{\mu}_{t_w}, & \Sigma_{t_w} = \sigma(\dot{\Sigma}_{t_w}) \\ \mu_{t_h} = \dot{\mu}_{t_h}, & \Sigma_{t_h} = \sigma(\dot{\Sigma}_{t_h}) \end{cases} \tag{3}$$

其中:均值 $\mu_{t_x}$ 和 $\mu_{t_y}$ 表示预测框中心点相对于负责该预测框回归的网格单元左上角坐标的偏移值,使用 sigmoid 函数处理后的取值范围在 (0, 1),控制目标中心的偏移位于对应的网格单元之内,确保不会过偏移;$\mu_{t_w}$ 和 $\mu_{t_h}$ 表示预测框宽、高相对于先验锚框宽、高的缩放尺度;$\Sigma_{t_x}$、$\Sigma_{t_y}$、$\Sigma_{t_w}$、$\Sigma_{t_h}$ 分别表示对应信息的

www.joca.cn

不确定性，取值为(0, 1)，值越小，说明对预测框的回归越准确。

加入预测框不确定性回归后网络对每一个预测框输出 8 个位置和尺寸信息，1 个有无目标的置信度信息和多个类别信息，改进后的输出信息如图 2 所示。由于预测框不确定性回归只在网络最后的检测层起作用，不会对网络的特征提取造成影响，所以不会额外增加网络的计算量。

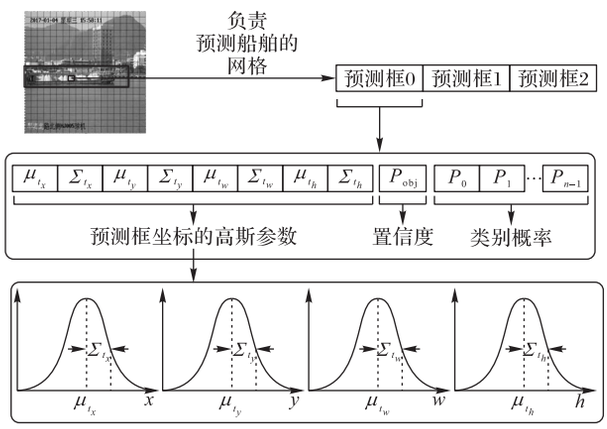



图 2 带有预测框不确定性回归的输出信息

Fig. 2 Output information with prediction box uncertain regression

通过网络学习到的预测框信息的偏移值可以计算出预测框的坐标和尺寸信息，如图 3 所示，网络最后输出的预测框的坐标、尺寸信息可以由式(4)计算：

$$\begin{cases} b_x = \mu_{t_x} + c_x \\ b_y = \mu_{t_y} + c_y \\ b_w = p_w \mathrm{e}^{\mu_{t_w}} \\ b_h = p_h \mathrm{e}^{\mu_{t_h}} \end{cases} \tag{4}$$

其中：$(b_x, b_y)$表示输出的预测框中心点坐标，$b_w$和$b_h$分别表示输出的预测框的宽和高；$p_w$和$p_h$分别表示先验锚框映射到特征图的宽和高；$(c_x, c_y)$表示网格单元左上角的坐标，在预测层的特征图中，每一个网格单元的边长均为 1。

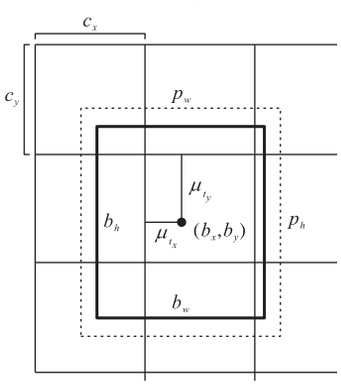


图 3 预测框的坐标信息

Fig. 3 Coordinate information of prediction box

## 2.2 损失函数改进

损失函数是评价网络训练情况的重要指标，通过对网络预测结果与真实样本标记产生的误差反向传播指导网络调整参数学习，YOLOv3 算法主要使用了两种损失函数，对预测框信息使用了均方误差损失函数，对有无目标的置信度信息和类别概率信息使用二值交叉熵损失函数。

由于预测信息中加入了对预测框的不确定性回归，所以需要对位置损失函数进行重新设计。使用负对数似然损失代替均方误差损失如式(5)所示：

$$\begin{cases} L_x = -\delta \ln\left(N\left(x \middle| \mu_{t_x}(\hat{x}), \Sigma_{t_x}(\hat{x})\right) + \varepsilon\right) \\ L_y = -\delta \ln\left(N\left(y \middle| \mu_{t_y}(\hat{y}), \Sigma_{t_y}(\hat{y})\right) + \varepsilon\right) \\ L_w = -\delta \ln\left(N\left(w \middle| \mu_{t_w}(\hat{w}), \Sigma_{t_w}(\hat{w})\right) + \varepsilon\right) \\ L_h = -\delta \ln\left(N\left(h \middle| \mu_{t_h}(\hat{h}), \Sigma_{t_h}(\hat{h})\right) + \varepsilon\right) \end{cases} \tag{5}$$

其中：$L_x$、$L_y$、$L_w$、$L_h$分别表示预测框的中心坐标位置损失和尺寸损失；$\delta$表示是否为最合适的预测框：是为 1，否则为 0；$\varepsilon$取一个极小的值；$x$、$y$表示目标真实边界框相对于网格单元左上角坐标的偏移值，$w$、$h$表示目标真实边界框宽、高相对于先验锚框宽、高的缩放尺度；$\mu_{t_x}(\hat{x})$、$\mu_{t_y}(\hat{y})$表示预测框相对于网格单元左上角坐标的偏移值，$\mu_{t_w}(\hat{w})$、$\mu_{t_h}(\hat{h})$表示预测框宽、高相对于先验锚框宽、高的缩放尺度。其中，真实值可由式(6)计算：

$$\begin{cases} x = G_x - c_x \\ y = G_y - c_y \\ w = \ln(G_w / p_w) \\ h = \ln(G_h / p_h) \end{cases} \tag{6}$$

其中：$G_x$、$G_y$、$G_w$、$G_h$分别表示目标真实边界框映射到特征图上的中心点坐标和宽、高。

不同于 YOLOv3 中使用的均方误差损失函数无法反映噪声数据对网络训练的影响，负对数似然损失函数可以对网络训练过程中不一致的数据添加不确定性惩罚，从而使网络从稳定的数据中学习预测框的准确位置，提高网络对预测框坐标位置估计的准确程度。

YOLOv3 算法存在的第二个问题是前景、背景复杂度不平衡，如图 4 所示，在船舶目标检测中，待检测的船舶是前景，其他部分是背景，简单样本的特点是距离近、尺寸大，与背景差异较大，容易区分；复杂样本的特点是拍摄距离远、尺寸小，与背景极为相似，不容易识别。而且在上万个预测框中，最后只有一个预测框是最适合的，这种“万里挑一”的做法给网络训练带来极大的困难，网络将过多的焦点专注于学习样本数量多的几类目标，而忽略了对样本数量较少的目标的特征的学习，影响损失函数的梯度更新。



(a) 简单样本

(b) 复杂样本

图 4 复杂度不平衡的样本

Fig. 4 Samples with imbalanced complexity

YOLOv3 算法在置信度损失和类别概率损失中使用的传统二值交叉熵函数为$L_{\mathrm{bce}} = -\ln(p_t)$，式中，$p_t$表示概率。借鉴文献[17]中对密集物体的检测方法，在上式中加入动态放缩因子，使损失函数能够降低简单样本的损失，集中训练数据中

www.joca.cn

的复杂样本。改进后的损失函数如式(7)所示:

$$L_{\text{focal}} = -\alpha_t(1-p_t)^{\gamma}\ln(p_t) \tag{7}$$

其中:$(1-p_t)^{\gamma}$表示动态放缩因子;$\gamma$是可调超参数,用来控制放缩比例,解决前景、背景的不平衡,本文取2;$\alpha_t$表示平衡变量,取值为(0,1),用来解决船舶样本复杂度的不平衡,本文取0.3。

图像中的前景属于正样本,背景属于负样本。正样本中又分为简单样本和复杂样本,对于简单样本,容易分类,输出概率值大,复杂样本不容易分类,输出概率之小。对于简单样本,$(1-p_t)^{\gamma}$的值会更小,得到的损失函数也更小,而复杂样本得到的损失相对变更大;对于负样本,预测概率值小的样本损失比预测概率值大的损失要小得多,可以使损失函数更多地聚焦在复杂样本的学习上,减少简单样本的影响。由于正样本占比例较小,而且负样本相对比较容易进行分类,所以使用一个比较小的平衡变量对正样本进行平衡。

在目标检测中使用边界框真值与预测框的交并比(Intersection over Union, *IoU*)表示预测的准确程度,但是当两个框没有重合时会导致梯度为0,造成损失函数无法优化。为了解决这个问题,受文献[18]的启发,对于任意的两个检测框,定义广义交并比(Generalized Intersection over Union, *GIoU*)如式(8)所示:

$$GIoU = IoU - \frac{|C\backslash(A\cup B)|}{|C|} \tag{8}$$

其中:$A$、$B$表示任意两个检测框;$C$表示两个检测框的最小闭合凸面;$(A\cup B)$表示两个检测框重叠部分的面积。则$GIoU$的损失函数可以表示为$L_{\text{giou}} = 1 - GIoU$。

所以,改进后的总体损失函数为:

$$\begin{aligned}L_{\text{total}} &= L_{\text{coor}} + L_{\text{giou}} + L_{\text{conf}} + L_{\text{prob}} = \\ &L_x + L_y + L_w + L_h + l_{\text{bbox_scale}} \times (1 - GIoU) + \\ &L_{\text{focal}}(\mathbf{1}^{\text{obj}}, \hat{1}^{\text{obj}}) \times L_{\text{bce}}\left(\mathbf{1}^{\text{obj}}, \hat{1}^{\text{obj}}\right) + L_{\text{bce}}\left(P(c), \hat{P}(c)\right)\end{aligned} \tag{9}$$

其中:$L_{\text{coor}}$、$L_{\text{conf}}$、$L_{\text{prob}}$分别表示坐标损失、置信度损失和类别概率损失;$l_{\text{bbox_scale}} = 2 - \frac{wh}{WH}$表示预测框尺度损失,$w$、$h$和$W$、$H$分别表示预测框和输入图片的宽和高;$\mathbf{1}^{\text{obj}}$表示网格中是否存在船舶目标,存在为1,不存在为0;$P(c)$表示船舶所属的类别概率。

## 3　网络训练

### 3.1　数据集

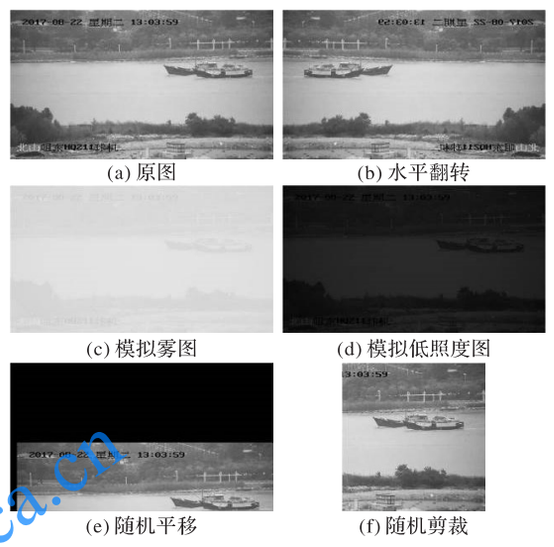
数据集在基于深度学习的目标检测中起着至关重要的作用,卷积神经网络从庞大的训练数据中学习目标特征。由于监督学习需要人工标记标签,是一项费时费力的工作,所以获得训练数据集极其困难。SeaShips[19]是第一个公开的用于船舶目标检测的数据集,包含7 000张6个类别的船舶图像。将数据集分为:训练集1 750张、验证集1 750张,测试集3 500张,分别对3个子数据集中每个类别的船舶数量(GT)进行统计如表1所示,3个子数据集的目标数量分布相似,保证了网络具有良好的泛化能力。

文献[20]证明雾天、低照度条件会严重影响船舶目标检测的准确率,为了增强网络的泛化能力,缓解训练过程中的过拟合现象,提高算法在恶劣天气条件下的检测能力,在训练过程中使用数据增强技术扩充样本数量。除了使用水平翻转、随机平移和随机剪切外,针对水上常见的雾天和低照度环境,但难以获取大量恶劣天气环境下的船舶图片的事实,使用基于大气散射模型和Retinex模型的模拟雾图、模拟低照度图技术。样本增强后的效果如图5所示。

表1　不同类别船舶目标数量统计

Tab. 1　Statistics of numbers of different ship categories

| 船舶类别 | 训练集 | 验证集 | 测试集 |
|---|---|---|---|
| 客船 | 100 | 126 | 248 |
| 集装箱船 | 245 | 219 | 744 |
| 普通货船 | 393 | 368 | 995 |
| 散装货船 | 489 | 468 | 437 |
| 渔船 | 510 | 492 | 1 188 |
| 矿砂船 | 542 | 586 | 1 071 |
| 总数 | 2 279 | 2 259 | 4 683 |



(a) 原图　(b) 水平翻转　(c) 模拟雾图　(d) 模拟低照度图　(e) 随机平移　(f) 随机剪裁

图5　数据增强

Fig. 5　Data augmentation

### 3.2　先验锚框设计和结果后处理

由于船舶目标呈细长状,也就是边界框的宽比高要大很多,而且不同类型船舶之间尺寸差异很大。YOLOv3原始设定的先验锚框尺寸主要面向自然界中通用目标检测,不能满足船舶目标检测的需要,所以需要重新对先验锚框进行设计。不同于依赖人的先验知识设定的锚框尺寸,通过使用*K*-means聚类算法对数据集中所有已标注的目标边界框进行聚类,产生不同数量的先验锚框,可以使锚框与目标边界框更加匹配,从而提高检测精度。不同聚类数目对应的平均交并比如图6所示。

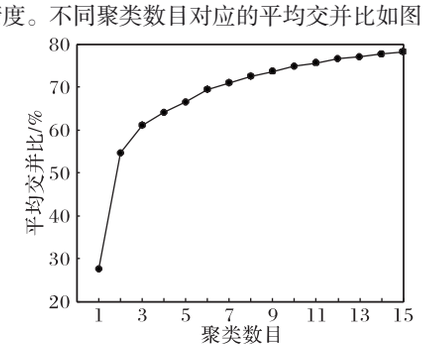



图6　不同聚类数目与平均交并比

Fig. 6　Different number of clusters and mean intersection over union

考虑到计算效率与准确率的平衡,同时为了使产生的先验锚框平均分配到3个预测尺度,选择产生9个先验锚框。最终确定的先验锚框尺寸经归一化后如表2所示。将YOLOv3原始设计的先验锚框尺寸与本文重新设计的先验锚框尺寸进行比较如图7所示,可以看到两者的差异很大,说明重新设计

www.joca.cn

先验锚框的必要性。

**表 2 先验锚框尺寸**

Tab. 2 Prior anchor box scales

| 特征图尺度 | 锚框尺寸 | | |
|---|---|---|---|
| | 锚框 1 | 锚框 2 | 锚框 3 |
| 细尺度 | (0.072 4,0.033 3) | (0.142 7,0.058 3) | (0.193 2,0.095 4) |
| 中尺度 | (0.268 2,0.161 1) | (0.274 0,0.064 8) | (0.352 1,0.110 2) |
| 粗尺度 | (0.447 9,0.210 2) | (0.583 3,0.137 0) | (0.776 6,0.254 6) |

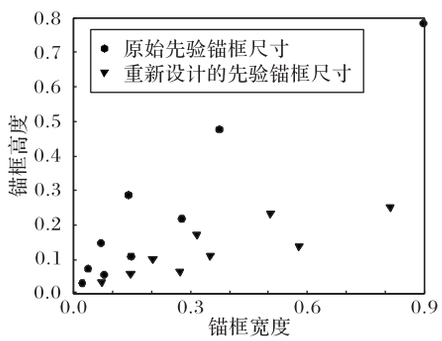



图 7 不同锚框尺寸比较

Fig. 7 Comparison of different prior anchor box scales

在预测的后处理阶段，YOLOv3 使用非极大值抑制(Non-Maximum Suppression，NMS)算法消除冗余的预测框，降低误检率。由于非极大值抑制算法会将和得分最高的预测框的交并比大于某个阈值的其他预测框直接置零，所以在有重合目标的情况下会导致有目标检测失败。为了改进非极大值抑制算法硬阈值的缺点，设置高斯软阈值函数如式(10)所示：

$$s_i = s_i e^{-\frac{\left(1-IoU\left(M,b_i\right)\right)^2}{\sigma}},\ \forall b_i \notin D \tag{10}$$

其中：$b_i$ 表示网络预测出的一系列预测框；$s_i$ 表示预测框得分；$M$ 表示得分最高的预测框；$D$ 表示经过非极大值抑制算法处理后保留的一系列预测框的集合，初始时将 $D$ 设为空集；$IoU$ 表示交并比；$\sigma$ 取值为(0,1)，本文取 0.3。高斯软阈值函数可以在得分最高的预测框和其他预测框的交并比越大时降低其他预测框的置信度，而不是粗暴地将其他预测框置零，从而缓解了目标重合时的检测失败问题。

## 4 实验结果分析

### 4.1 实验条件与评价指标

本文实验使用的操作系统为 ubuntu18.04，处理器型号为 Intel Core i7-8700K，显卡型号为 NVIDIA GeForce GTX 1080 Ti，采用 NVIDIA CUDA9.0 加速工具箱。

本文中除了特别说明以外，在网络训练阶段使用的部分实验参数如表 3 所示。

为了评价网络的性能和说明船舶目标检测网络的有效性，选取下列评价指标。

1) 查准率(Precision，P)和召回率(Recall，R)：查准率指网络检测到的正样本数量占检测到的所有样本数量的比率；召回率指网络检测到的正样本数量占标记的真实样本数量的比率。查准率和召回率的计算公式为：

$$\begin{cases} Precision = TP/(TP + FP) \\ Recall = TP/(TP + FN) \end{cases} \tag{11}$$

其中：真正样本(True Positive，$TP$)表示检测到的目标类别与真实目标类别一致的样本数量；假正样本(False Positive，$FP$)为检测到的目标类别与真实目标类别不一致的样本数量；假负样本(False Negative，$FN$)为真实目标存在但未被网络检测出来的样本数量。

**表 3 实验参数设置**

Tab. 3 Setting of experimental parameters

| 参数 | 值 |
|---|---|
| 输入数据的批大小 | 6 |
| 调整输入图像尺寸 | 416 × 416 |
| 采样步长 | 32 |
| 迭代轮数 | 100 |
| 初始学习率 | 0.000 1 |
| 结束学习率 | 0.000 001 |
| 动量参数 | 0.999 5 |

2) 平均准确率(Average Precision，AP)和平均准确率均值(mean Average Precision，mAP)：一个理想的目标检测网络应该在召回率增长的同时查准率保持在很高的水平，但现实情况是召回率的提高往往需要损失查准率的值，所以通常情况下用查准率-召回率(Precision-Recall，P-R)曲线来显示目标检测器在准确率和召回率之间的平衡。对于每一个船舶类别，该类别的平均准确率定义为 P-R 曲线下方的面积；平均准确率均值是所有类别的平均准确率的均值。AP 和 mAP 的计算公式为：

$$\begin{cases} AP = \int_0^1 P(R)\,\mathrm{d}R \\ mAP = \sum_{i=1}^{N} AP_i \Big/ N \end{cases} \tag{12}$$

其中 $N$ 表示所有目标类别的数量。

3) 帧率(Frames Per Second，FPS)：目标检测网络每秒钟能够检测的图片数量，用该指标评价目标检测网络的检测速度。

### 4.2 改进方法对比实验

为了获得更高性能的训练模型，在网络训练阶段会使用多种有利于网络训练的方法，这些方法对神经网络的影响效果如表 4 所示。其中，“√”表示使用了对应的方法。由于原始的数据集样本数量太少，所以直接使用 YOLOv3 进行训练的准确率比较低，同时由于原始的先验锚框尺寸与船舶目标形状差异较大，得到的预测框的交并比值也比较低，这说明对船舶目标的定位不够准确。使用了合成恶劣天气数据增强方法之后，有效扩充了训练样本的数量，从而提高了网络的鲁棒性和泛化能力，使算法的准确率提升了 5.24 个百分点。在网络训练时使用重新设计的先验锚框，与船舶的形状更加匹配，使网络的平均交并比提高了 6.93 个百分点，说明网络输出的预测框与人工标注的真实边界框重合程度更高，对船舶的定位更加准确，而且在与其他方法组合使用时，网络的平均交并比还会进一步提升。在网络的后处理阶段，在传统非极大值抑制算法中加入了高斯软阈函数，主要是为了解决重合目标的漏检测问题。从表 5 可以看出，使用了改进后的非极大值抑制算法可以提高算法的召回率，这意味着有一些传统 YOLOv3 算法无法检测出来的船舶目标被检测出来，其中散装货船、矿砂船和渔船召回率提升较多，主要是在这三种类型的船舶相互重合的场景多。因为有一些重合的船舶被正确检测，也让整个网络的检测精度有了略微的提高。使用了 3 种训练方法的组合对 YOLOv3 算法进行训练，得到的模型比原始模型的检测精度提升了 5.85 个百分点，平均交并比提升了 6.86 个百分点，说明了在网络训练阶段使用合成恶劣天气数据增强、根据检测对象重新设计先验锚

www.joca.cn

框和加入高斯软阈值函数的非极大值抑制算法的组合训练方法能够有效提升网络的性能。

**表 4　不同改进方法对算法性能的提升**

Tab. 4　Performance improvement of algorithms by using different improvement methods

| 方法 | 预测框不确定性回归 | 改进二值交叉熵损失函数 | 重新设计先验锚框 | 改进非极大值抑制算法 | 合成恶劣天气数据增强 | mAP/% | 平均 IoU/% |
|---|---|---|---|---|---|---|---|
| YOLOv3 | | | | | | 70.69 | 45.51 |
| 改进算法 1 | | | √ | | √ | 75.93 | 52.34 |
| 改进算法 2 | | | √ | √ | √ | 76.43 | 52.37 |
| 改进算法 3 | √ | | √ | √ | √ | 83.24 | 57.57 |
| 改进算法 4 | | √ | √ | √ | √ | 80.36 | 53.46 |
| 增强 YOLOv3 | √ | √ | √ | √ | √ | 85.62 | 58.07 |

**表 5　使用不同非极大值抑制算法的船舶召回率　单位:%**

Tab. 5　Recall of ships using different NMS algorithms　unit:%

| 船舶类别 | 传统 NMS | 改进 NMS | 变化值 |
|---|---|---|---|
| 普通货船 | 83.47 | 83.87 | +0.40 |
| 散装货船 | 79.37 | 80.30 | +0.93 |
| 矿砂船 | 90.05 | 91.16 | +1.11 |
| 客船 | 60.08 | 60.48 | +0.40 |
| 渔船 | 78.70 | 81.06 | +2.36 |
| 集装箱船 | 87.87 | 88.33 | +0.46 |
| 平均召回率 | 81.89 | 83.07 | +1.18 |

在使用网络组合训练方法的基础上,在网络预测层加入对预测框的不确定性回归使网络整体性能提升了 6.81 个百分点,平均交并比提高 5.2 个百分点,这证明了加入预测框不确定性回归减小了 YOLOv3 算法的定位误差。如果只使用改进的二值交叉熵损失函数进行网络训练,也会使 YOLOv3 算法的检测精度提高 3.93 个百分点,平均交并比有小幅提升,说明使用改进后的损失函数可以有效提高算法的性能。

最后,使用所提改进方法的组合模型进行网络训练,本文将该模型称为增强 YOLOv3,实验结果显示增强 YOLOv3 相较于使用单独改进方法的模型,算法性能均有提升,主要原因是每一项改进方法都针对传统 YOLOv3 算法的某一问题进行改进,而使用组合的改进方法可以有效解决传统 YOLOv3 算法在船舶目标检测任务中存在的问题,取得较好的检测效果。实验结果表明,与传统 YOLOv3 算法相比,增强 YOLOv3 算法的检测精度提升了 21.12%,平均交并比提升了 27.60%。

### 4.3　预测框不确定性回归及损失函数改进实验分析

为了验证加入预测框不确定性回归后算法对目标定位误差是否有改进,实验在 YOLOv3 的最后预测层加入预测框不确定性信息,对预测框坐标的损失函数使用负对数似然损失函数替换均方误差损失函数,而对置信度和类别概率仍然使用传统的二值交叉熵损失函数,所得模型称为带有预测框不确定性信息的 YOLOv3(YOLOv3 with Prediction Box Uncertainty Information, YOLOv3-PBUI)。传统 YOLOv3 算法在最后的每个尺度的预测层输出的张量维度为 $S \times S \times (5+n) \times 3$,其中 $S \times S$ 是对特征图划分的网格单元的数量,5 指的是 5 个预测信息,3 表示每个网格单元预测的 3 个预测框,$n$ 表示共有 $n$ 个目标类别。加入预测框不确定性回归信息后,YOLOv3-PBUI 每个尺度的预测层输出的张量维度变为 $S \times S \times (9+n) \times 3$,本文中共有 6 个船舶类别,所以 $n=6$,输出的张量维度为 $S \times S \times 45$。

分别对传统 YOLOv3 算法和 YOLOv3-PBUI 算法进行训练,得到训练好的模型。使用测试集进行测试,设置 IoU 阈值为 0.5,得到的 TP 值和 FP 值如表 6 所示,图 8 显示了部分样本的检测结果。由图 8 可以看到,对于一些非常小的船舶目标,传统 YOLOv3 算法无法准确检测,而改进后的算法可以准确对这些小船舶目标进行检测,从而增加了算法的 TP 值,说明改进后的算法可以对更多的正样本准确检测;对于部分船舶目标,虽然传统 YOLOv3 算法可以检测出该船舶,但是错误识别了该船舶的类型,从而导致了 FP 值很高,而改进后的 YOLOv3-PBUI 算法可以同时准确定位船舶位置和识别船舶类型,比传统 YOLOv3 算法更加鲁棒。同时,因为改进后的 YOLOv3-PBUI 算法可以更好地学习目标的位置信息,所以可以消除对船舶的错误定位。从表 6 可以看出,加入预测框不确定性回归之后的 YOLOv3-PBUI 算法 FP 值降低了 35.42%,而 TP 值也提高了 1.83%,这是改进模型检测精度提高的主要原因。对于海事视频监控来讲,FP 值降低说明能够准确识别船舶类型,这对重点关注某一类型的船舶或防止船舶碰撞是有帮助的,而 TP 值增加说明可以准确检测出更多的船舶数量,降低漏检的可能性,这对统计船舶交通流量和监控船舶行为十分必要。

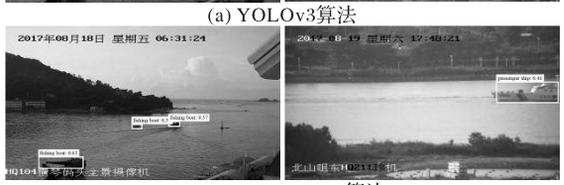
**表 6　FP 和 TP 数值评价**

Tab. 6　Numerical evaluation of FP and TP

| 算法 | FP 数量 | TP 数量 | GT 数量 |
|---|---|---|---|
| YOLOv3 | 2 250 | 3 835 | 4 683 |
| YOLOv3-PBUI | 1 453 | 3 905 | 4 683 |
| 变化率/% | −35.42 | +1.83 | 0 |



(a) YOLOv3算法



(b) YOLOv3-PBUI算法

图 8　YOLOv3 和 YOLOv3-PBUI 的检测结果

Fig. 8　Detection results of YOLOv3 and YOLOv3-PBUI

为了验证改进后的二值交叉熵损失函数是否有助于解决传统 YOLOv3 算法存在的类别不平衡问题,实验用改进的二值交叉熵损失函数替换了传统二值交叉熵损失函数,得到的实验结果如表 7 所示。在训练集中,客船的样本数量是最少的,不足渔船和矿砂船样本数量的 1/5,不足所有训练样本的 5%,存在着样本类别的不平衡问题,同时训练样本中有很多图像中只有一个或几个极小的船舶目标,而传统 YOLOv3 算法在 3 个尺度的预测层共产生 10 647 个预测框,存在着严重的前景、背景不平衡问题,但是实验结果显示改进后的模型对客船这一类船舶的检测

www.joca.cn

准确率提高最多，提高了 8.08 个百分点，而对样本数量相对少的船舶类别的检测准确率都有不同程度的提高，说明改进后的损失函数可以使网络专注于这些比较困难的样本的学习，提高了网络的性能。而测试样本中大多数小船舶目标都是渔船，而且渔船的测试样本数量是最多的，虽然改进后的网络对渔船的检测性能有略微的降低，但是从整体上看改进的二值交叉熵损失函数对解决类别不平衡问题还是有效果的，对渔船的性能损失可以通过使用其他改进方法进行弥补。

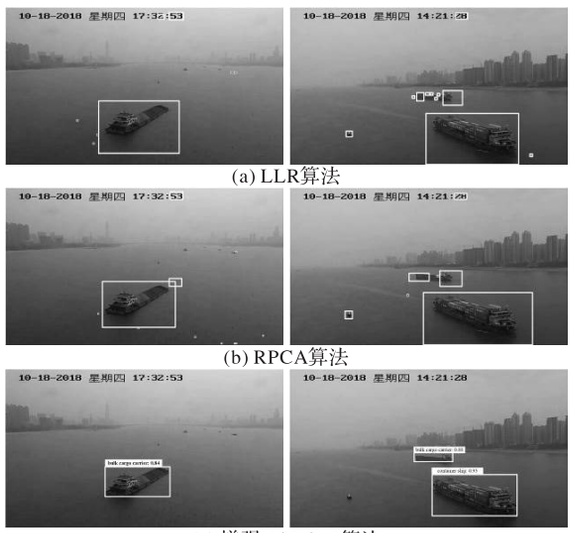
表 7 使用不同损失函数的船舶平均精度 单位：%

Tab. 7 mAP of ships using different loss functions unit：%

| 船舶类别 | 普通损失函数 | 改进损失函数 | 变化值 |
|---|---|---|---|
| 普通货船 | 78.29 | 83.83 | +5.54 |
| 散装货船 | 79.07 | 83.64 | +4.57 |
| 矿砂船 | 78.40 | 81.38 | +2.98 |
| 客船 | 57.98 | 66.06 | +8.08 |
| 渔船 | 77.16 | 77.04 | −0.12 |
| 集装箱船 | 87.71 | 90.23 | +2.52 |
| mAP | 76.43 | 80.36 | +3.93 |

### 4.4 不同算法实验结果分析

为了进一步说明本文所提方法的优越性，增强 YOLOv3 算法将与两种典型的基于背景建模的目标检测方法进行比较分析，即低秩复原（Low Rank Recovery，LLR）算法[21]和鲁棒主成分分析（Robust Principal Component Analysis，RPCA）算法[22]。由于传统的船舶目标检测往往作为船舶目标跟踪的前序步骤，所以需要使用视频数据进行实验。本文选择了两段采集于武汉二七长江大桥的监控视频，分别使用三种检测算法进行检测，得到的结果如图 9 所示。由于传统目标检测算法使用人工设计的特征表达，特征设计往往比较简单，所以无法对不同船舶的类别进行识别。水面动态背景会导致传统检测算法将水面检测为前景，造成误检测非常多。背景建模方法获得的前景往往很难是一个闭合的整体，所以容易导致对同一个船舶的多个检测，使检测的准确率下降。而基于深度学习的增强 YOLOv3 算法在网络训练中只学习前景的特征而忽略了背景，所以有效解决了传统目标检测算法存在的问题，而且可以识别出不同类型的船舶，更符合智能视频监控的需要。




(a) LLR算法

(b) RPCA算法

(c) 增强YOLOv3算法

图 9 增强 YOLOv3 算法与传统方法检测结果比较

Fig. 9 Detection result comparison of the enhanced YOLOv3 algorithm and traditional methods

本文还与目标检测中常用的几种深度学习算法进行了对比实验，包括单次目标检测算法 YOLOv2、SSD 和两次目标检测算法 Faster R-CNN。实验使用同一训练集对每种算法进行训练，网络训练的相关参数都使用对应论文中建议的参数，在网络训练过程中输入图像的尺寸也使用论文中建议的输入尺寸，其中 Faster R-CNN 算法限制输入图像的最长边为 1 000 像素值，并对输入图像等比例缩放。通过官方评估算法使用同一测试集对算法性能进行评估，各检测算法的性能如表 8 所示，由于 SSD 算法的官方评估算法没有提供对单种类别的平均精度，所以表 8 中没有列出相关结果。具有代表性的 4 种检测算法的检测效果如图 10 所示。

表 8 不同目标检测网络性能比较

Tab. 8 Performance comparison of different object detection networks

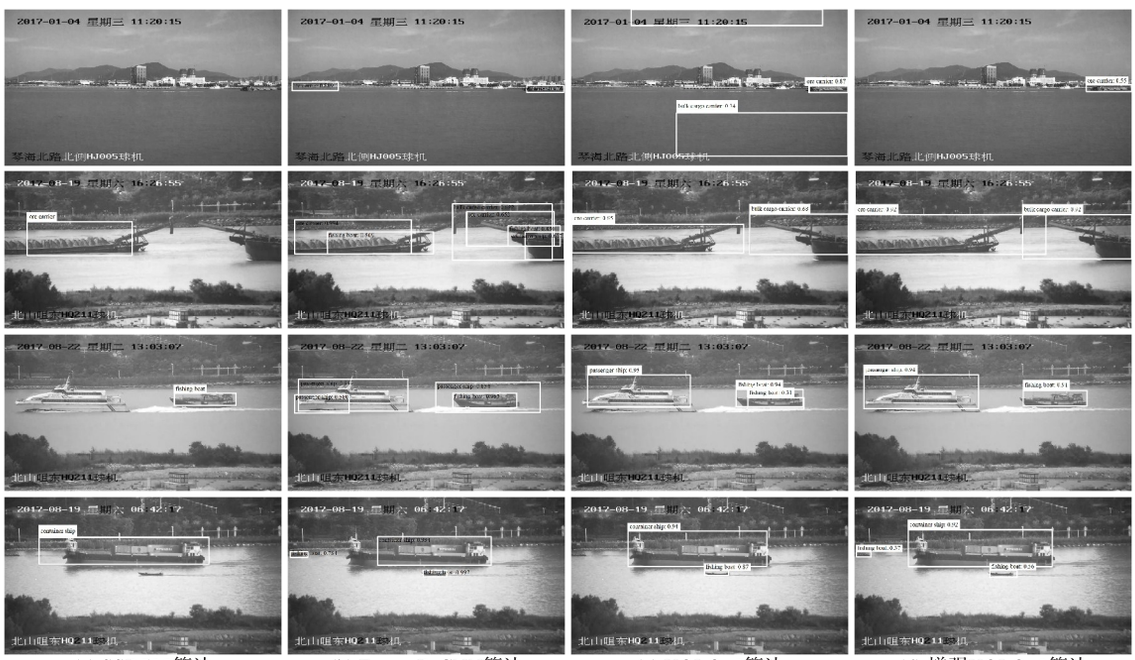
| 检测算法 | AP/% | | | | | | mAP/% | FPS | 输入尺寸 |
|---|---|---|---|---|---|---|---|---|---|
| | 普通货船 | 散装货船 | 矿砂船 | 客船 | 渔船 | 集装箱船 | | | |
| YOLOv2 | 70.31 | 66.85 | 58.45 | 48.00 | 40.76 | 74.77 | 59.85 | 23.97 | 416 × 416 |
| SSD300 | — | — | — | — | — | — | 60.90 | 23.81 | 300 × 300 |
| SSD512 | — | — | — | — | — | — | 66.71 | 18.26 | 512 × 512 |
| Faster R-CNN | 70.19 | 66.12 | 80.83 | 54.66 | 71.61 | 82.75 | 71.03 | 4.78 | max1 000 |
| YOLOv3 | 74.00 | 70.24 | 70.89 | 53.16 | 72.53 | 83.34 | 70.69 | **43.53** | 416 × 416 |
| 增强 YOLOv3 | **87.95** | **85.21** | **90.61** | **82.09** | **89.51** | **91.07** | **87.74** | 30.70 | 608 × 608 |

实验结果表明，本文所提出的增强 YOLOv3 算法表现出了较好的性能，对每一种船舶类别的检测效果相对平均，而另外 4 种目标检测算法对于样本数量最少的客船检测效果，相比样本数量较多的船舶类别的检测效果都有极大的差距，这说明增强 YOLOv3 算法很好地解决了样本类别不平衡的问题。图 10 的检测效果显示增强 YOLOv3 算法能够准确对图像中的船舶目标进行定位和种类识别，而其他 3 种算法漏检测、对同一个船舶的多重检测和错误识别船舶类型的现象严重，这说明增强 YOLOv3 算法对船舶目标的定位更加准确，原因是增强 YOLOv3 算法增加了对预测框的不确定性回归，可以使网络更好地学习目标的位置信息，加入高斯软阈值函数后的非极大值抑制算法也比传统非极大值抑制算法更加稳定。作为两次目标检测算法，Faster R-CNN 受到检测速度的制约，无法达到实时检测的要求，而 YOLOv3 算法在所有算法中表现出了最快的检测速度，这也是单次目标检测算法最大的优点。增强 YOLOv3 算法因为需要预测更多的信息，所以检测速度略有下降，但仍然快于其他检测算法，可以满足实时检测的需要。增强 YOLOv3 算法使用了分辨率更高的输入图像，这是因为分辨率高的图像包含更多的细节特征，有助于网络的学习，可以在一定程度上获得检测性能的提升。

图 11 展示了在不同场景下增强 YOLOv3 算法的检测效果，在低照度条件和雾天环境下都能准确检测出船舶目标，这得益于在网络训练过程中使用了合成恶劣天气图像的数据增强方法，提高了网络在各种天气条件下的稳定性和泛化能力。增强

www.joca.cn

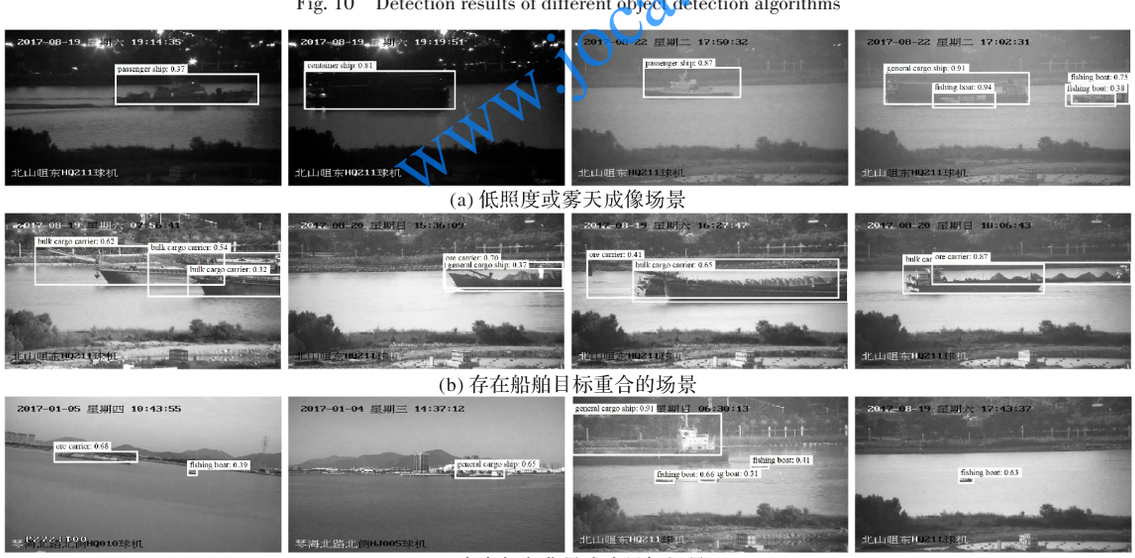
YOLOv3算法在船舶目标有大部分重叠的情况下和船舶目标与通航背景极为相似的情况下也都可以准确对船舶进行检测,适用于各种复杂场景下的船舶目标检测。通过多项实验表明,本文所提出的增强YOLOv3算法检测精度可以达到87.74%,远超其他目标检测算法,对船舶航行中遇到的多种复杂场景都有良好的检测的结果,是兼具检测精度和检测速度的船舶目标检测算法。



(a) SSD512算法　(b) Faster R-CNN算法　(c) YOLOv3算法　(d) 增强YOLOv3算法

图 10　不同目标检测算法的检测结果

Fig. 10　Detection results of different object detection algorithms



(a) 低照度或雾天成像场景



(b) 存在船舶目标重合的场景



(c) 存在复杂背景或小目标场景

图 11　增强 YOLOv3 算法在不同场景下的船舶目标检测结果

Fig. 11　Ship detection results of the enhanced YOLOv3 algorithm under different scenes

## 5　结语

传统 YOLOv3 目标检测网络无法准确检测复杂场景下的船舶目标。本文通过在 YOLOv3 网络的预测层加入对预测框的不确定性回归并重新改进了损失函数,使用组合网络训练方法提出了适合船舶目标检测的增强 YOLOv3 算法。实验结果表明预测框不确定性回归有效降低了对船舶目标的定位误差,改进的二值交叉熵损失函数解决了样本类别和前景、背景不平衡的问题,有助于算法性能的提升。在网络训练阶段使用重新设计的先验锚框提高了预测框位置的准确度,使用数据增强策略提高算法在各种天气条件和复杂通航背景下的检测能力,增强网络泛化能力和鲁棒性,加入高斯软阈值函数的非极大值抑制算法有效解决了船舶重合时容易漏检测的问题。通过与其他算法的对比实验验证了增强 YOLOv3 算法有效平衡了检测的精度和速度,适合复杂场景下的船舶目标检测。

**参考文献** (References)

[1] PRASAD D K, PRASATH C K, RAJAND D, et al. Object detection in a maritime environment: performance evaluation of background subtraction methods [J]. IEEE Transactions on

www.jcca.cn

Intelligent Transportation Systems, 2019, 20(5): 1787-1802.

[2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 580-587.

[3] GIRSHICK R. Fast R-CNN [C]// Proceedings of the 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2015: 1440-1448.

[4] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [C]// Proceedings of the 28th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2015: 91-99.

[5] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]// Proceedings of the 2016 European Conference on Computer Vision, LNCS 9905. Cham: Springer, 2016: 21-37.

[6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 779-788.

[7] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 6517-6525.

[8] REDMON J, FARHADI A. YOLOv3: an incremental improvement [EB/OL]. [2019-12-15]. https://arxiv. org/pdf/1804. 02767. pdf.

[9] 王新立,江福才,宁方鑫,等. 基于改进卷积神经网络的船舶目标检测[J]. 中国航海, 2018, 41(2): 41-45, 51. (WANG X L, JIANG F C, NING F X, et al. Ship detection with improved convolutional neural network [J]. Navigation of China, 2018, 41(2): 41-45, 51.)

[10] 陈从平,吴喆,吴杞,等. 动态背景下航道船舶目标检测方法[J]. 三峡大学学报(自然科学版), 2019, 41(1): 97-102. (CHEN C P, WU Z, WU Q, et al. Detection method of ship target in waterway under dynamic background [J]. Journal of China Three Gorges University (Natural Sciences), 2019, 41(1): 97-102.)

[11] 王言鹏,杨飏,姚远. 用于内河船舶目标检测的单次多框检测器算法[J]. 哈尔滨工程大学学报, 2019, 40(7): 1258-1262. (WANG Y P, YANG Y, YAO Y. Single shot multibox detector for ships detection in inland waterway [J]. Journal of Harbin Engineering University, 2019, 40(7): 1258-1262.)

[12] 赵蓬辉,孟春宁,常胜江. 基于改进 VGG 网络的单阶段船舶检测算法[J]. 光电子·激光, 2019, 30(7): 719-730. (ZHAO P H, MENG C N, CHANG S J. Single stage ship detection algorithm based on improved VGG network [J]. Journal of Optoelectronics·Laser, 2019, 30(7): 719-730.)

[13] 于洋,李世杰,陈亮,等. 基于改进 YOLO v2 的船舶目标检测方法[J]. 计算机科学, 2019, 46(8): 332-336. (YU Y, LI S J, CHEN L, et al. Ship target detection based on improved YOLO v2 [J]. Computer Science, 2019, 46(8): 332-336.)

[14] 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. (LIU B, WANG S Z, ZHAO J S, et al. Ship tracking and recognition based on Darknet network and YOLOv3 algorithm [J]. Journal of Computer Applications, 2019, 39(6): 1663-1668.)

[15] KIM K, HONG S, CHOI B, et al. Probabilistic ship detection and classification using deep learning [J]. Applied Sciences, 2018, 8(6): No. 936.

[16] CHOI J, CHUN D, KIM H, et al. Gaussian YOLOv3: an accurate and fast object detector using localization uncertainty for autonomous driving [C]// Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 502-511.

[17] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]// Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 2999-3007.

[18] REZATOFIGHI H, TSOI N, GWAK J, et al. Generalized intersection over union: a metric and a loss for bounding box regression [C]// Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 658-666.

[19] SHAO Z, WU W, WANG Z, et al. SeaShips: a large-scale precisely annotated dataset for ship detection [J]. IEEE Transactions on Multimedia, 2018, 20(10): 2593-2604.

[20] NIE X, YANG M, LIU R W. Deep neural network-based robust ship detection under different weather conditions [C]// Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference. Piscataway: IEEE, 2019: 47-52.

[21] SHU X, PORIKLI F, AHUJA N. Robust orthonormal subspace learning: efficient recovery of corrupted low-rank matrices [C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 3874-3881.

[22] KANG Z, PENG C, CHENG Q. Robust PCA via nonconvex rank approximation [C]// Proceedings of the 2015 IEEE International Conference on Data Mining. Piscataway: IEEE, 2015: 211-220.

This work is partially supported by the National Natural Science Foundation of China (516091295), the Fundamental Research Funds for the Central Universities (2019-zy-215).

**NIE Xin**, born in 1995, M. S. candidate. His research interests include computer vision.

**LIU Wen**, born in 1987, Ph. D., associate professor. His research interests include computer vision, maritime information perception and processing.

**WU Wei**, born in 1971, Ph. D., associate professor. His research interests include information processing, image processing, pattern recognition.

www.joca.cn