



## 基于改进YOLOV3-Tiny的海面舰船目标快速检测

李庆忠, 徐相玉

(中国海洋大学 工程学院, 山东 青岛 266100)

**摘要:** 为实现海面舰船目标的快速、准确检测, 提出一种改进的舰船目标检测算法。在网络结构方面根据舰船目标的特点, 对浅层信息进行强化重构以降低小目标的漏检率, 同时引入改进的残差网络增加网络深度和降低网络参数计算量, 并且采用金字塔网络进行多尺度特征融合, 以兼顾图像中大小舰船目标的检测性能。在网络训练中利用迁移学习策略进行网络模型的训练, 以克服舰船图像样本集有限的问题。在视频检测中利用帧间图像结构相似度进行选择网络前向计算, 以提高视频帧检测速率。实验结果表明, 该算法海面舰船目标检测的准确率达到92.4%, 较YOLOV3-Tiny提高7个百分点, 召回率达到88.6%, 且在CPU平台上舰船目标的检测速度达到12 frame/s。

**关键词:** 卷积神经网络; YOLO网络; 舰船目标检测; 迁移学习; 深度学习

开放科学(资源服务)标志码(OSID):



**中文引用格式:** 李庆忠, 徐相玉. 基于改进YOLOV3-Tiny的海面舰船目标快速检测[J]. 计算机工程, 2021, 47(10): 283-289, 297.

**英文引用格式:** LI Q Z, XU X Y. Fast target detection of surface ship based on improved YOLOV3-Tiny[J]. Computer Engineering, 2021, 47(10): 283-289, 297.

## Fast Target Detection of Surface Ship Based on Improved YOLOV3-Tiny

LI Qingzhong, XU Xiangyu

(College of Engineering, Ocean University of China, Qingdao, Shandong 266100, China)

**[Abstract]** In order to achieve fast and accurate detection of surface ship targets, this paper proposes a ship target detection algorithm based on improved YOLOV3-Tiny. Firstly, in network structure, the features of shallow layers of the network is enhanced and reconstructed according to the characteristics of ship targets to reduce the miss detection rate of small targets, and the improved residual network is introduced to improve the depth of the network while reducing the calculation of network parameters. Moreover, the pyramid network is used for multi-scale feature fusion to balance the detection capability between large ship targets and small ship targets in images. Secondly, in the network training, transfer learning strategy is employed to train the designed network model to alleviate the limitation of known ship image samples. Finally, in video detection, a video frame selection method for forward computation of the network model based on structure similarity of inter frames is proposed to improve the detection frame rate. The experimental results show that the proposed algorithm has precision rate up to 92.4%, with an increase of 7% compared with YOLOV3-Tiny, recall rate up to 84%, and detection frame rate up to 12 frames/s on CPU platform.

**[Key words]** Convolutional Neural Network (CNN); YOLO network; ship target detection; transfer learning; deep learning  
**DOI:** 10.19678/j.issn.1000-3428.0059305

### 0 概述

海事智能监测是维护海洋权益、加强海域监管、减少海事纠纷等的重要技术手段, 在军用和民用方面都具有重要的研究与应用价值。当摄像机安装在海洋浮标或无人海事执法船上时, 研究基于动态视频图像的海面舰船目标快速检测方法是实现大面积

海域智能视频监控的关键技术之一<sup>[1]</sup>。

在基于动态视频图像的海面目标自动检测方面, 目前常用的检测方法可分为3类。

第1类方法是基于边缘、纹理等特征的海面目标的检测方法。代表性研究包括: 黎经文等<sup>[2]</sup>提出的基于边缘梯度特征的舰船检测方法, 适合于港口舰船的快速检测; 李浩谊等<sup>[3]</sup>提出一种基于改进

**基金项目:** 国家重点研发计划(2017YFC1405202); 海洋公益性行业科研专项(201605002)。

**作者简介:** 李庆忠(1963—), 男, 教授、博士, 主研方向为图像处理、信号处理、模式识别; 徐相玉(通信作者), 硕士研究生。

**收稿日期:** 2020-08-19 **修回日期:** 2020-09-29 **E-mail:** xuxiangyu@stu.ouc.edu.cn

Scharr 边缘检测算子的船舶目标检测算法, 具有较好的抗噪声能力; YANG 等<sup>[4]</sup>利用图像子块的 DCT 域能量特征实现海天线的快速提取, 根据 DCT 纹理特征建立海面混合纹理模型, 并实现了海面背景与船舶目标快速分割。这类方法虽然具有较好的实时性, 但当海况变化剧烈时适应性下降。

第2类方法是模仿人眼视觉注意力选择机制, 根据建立的视觉注意力模型求取感兴趣目标的显著图并实现船舶目标的检测。代表性研究包括: 丁鹏等<sup>[5]</sup>提出一种基于多特征、多尺度视觉显著性的海面船舶检测方法, 在多海况下的实验结果表明该算法具有较高的检测精度和鲁棒性; SHI 等<sup>[6]</sup>在小波域分别提取图像的低频和低频特征, 利用改进的 Gabor 滤波器提取方向特征, 并在 HIS 空间提取颜色和矩特征, 融合以上各种特征得到显著图并实现船舶目标的分割; SHAO 等<sup>[7]</sup>提出一种基于卷积神经网络 (CNN) 的船舶目标检测算法, 利用 CNN 预测目标的类型和位置, 并辅助利用显著图校正目标的定位, 该方法具有较高的检测精度和速度。该类方法的优点是充分利用多种特征产生感兴趣目标的显著图, 但具有较高的计算复杂度。

第3类方法是基于深度学习的海面船舶目标检测算法。这是近年来国内外学者的研究热点, 具有特征提取能力强、识别精度高、实时性好等优点。代表性研究包括: ZHANG 等<sup>[8]</sup>提出的基于干扰因素判别器和船舶目标提取器集成目标分割方法, 利用 SqueezeNet 网络作为干扰因素判别器以判定输入图像中包含什么类型的干扰, 用改进的 DeepLabv3+ 深度网络进行船舶目标的分割, 该方法具有较高的分割精度及较好的抗雾能力; WANG 等<sup>[9]</sup>利用改进的 YOLOV3 (You Only Look Once) 实现了端对端的船舶目标快速检测, 在 GPU 1080Ti 硬件环境下, 检测精度达到 74.8%, 检测速度达到 29.8 frame/s; 马啸等<sup>[10]</sup>通过精简 YOLO 模型, 设计一个复杂度较小的 10 层卷积神经网络用于船舶目标的自动特征提取和分类识别, 改进后的网络虽然具有较高的检测精度, 但只适合大小为 112 像素×112 像素的输入图像, 在 CPU 硬件环境下, 识别一张船舶目标图像所用的平均时间为 0.265 s; 赵春晖等<sup>[11]</sup>提出基于改进 Fast-RCNN 算法的船舶目标检测与识别算法, 虽然检测精度较高, 但实时性较差。

总之, 基于深度学习的海面船舶目标检测方法具有很好的应用前景, 但目前该类算法仍然存在以下 2 方面的问题: 一是存在小目标检测准确性低的问题; 二是目前的算法大都适合于 GPU 服务器硬件运行环境, 对于安装在无人船或浮标上的嵌入式视频监控系统, 其实时性仍然无法满足实际需求。

为克服上述问题, 本文提出基于改进 YOLOV3-Tiny 的船舶目标检测方法。根据舰船目标的特点利用浅层感受野对浅层信息进行强化重构, 增加 I-ResNet 网络以及特征金字塔网络 FPN, 进行多尺寸特征融合以平衡船舶大小目标的检测能力。利用图像结构相似度进行视频关键帧选择性检测以提高视频检测的速度, 运用迁移学习方法进行网络模型的训练学习, 克服舰船训练样本集有限的问题, 最终在视频测试集上进行性能测试。

## 1 YOLO 网络基本原理

YOLO 是一种采用卷积神经网络 (CNN) 实现端到端目标检测的算法, 其主要有以下特点: 1) 运用回归的思想, 将目标检测看成是一个回归的问题, 能够实时预测多个目标的类别和目标边框的位置; 2) YOLO 采用滑动窗口的方式寻找目标, 与传统的基于候选区域方式不同, 其直接利用整幅图片训练网络模型<sup>[12]</sup>。

目前 YOLO 最常见的网络框架有 YOLOV1、YOLOV2、YOLOV3<sup>[13]</sup> 3 种。YOLOV3-Tiny 网络是 YOLOV3 网络的简化, 网络模型相对简单, 降低了对硬件的要求, 虽然提高了检测速度, 但检测精度有所下降。YOLOV3-Tiny 网络模型是由 13 个卷积层和 6 个最大池化层组成, 只使用了 2 个不同尺度的 YOLO 输出层, 分别是 YOLO16 和 YOLO23, 由 1×1 和 3×3 的卷积核提取特征, 输出层有 2 个预测尺度, 其大小分别为 13×13 和 26×26。YOLOV3-Tiny 网络模型的组成如图 1 所示。

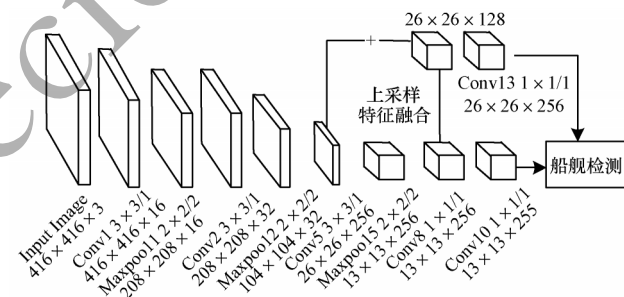


图1 YOLOV3-Tiny 网络模型框架

Fig.1 Framework of YOLOV3-Tiny network model

YOLOV3-Tiny 网络可以检测 80 多种不同对象, 虽然具有较快的检测速度, 但其目标检测精度不高, 且对复杂场景的适应能力较低。

## 2 基于改进 YOLOV3-Tiny 的船舶目标检测

为了实现安装在无人船或海洋浮标上的嵌入式智能视频监控系统的海面船舶目标快速检测, 根据船舶目标的特点, 在 YOLOV3-Tiny 的基础上, 本文提出适合海面船舶目标检测的改进型网络模型——I-YOLOV3-Tiny, 其结构组成如图 2 所示。

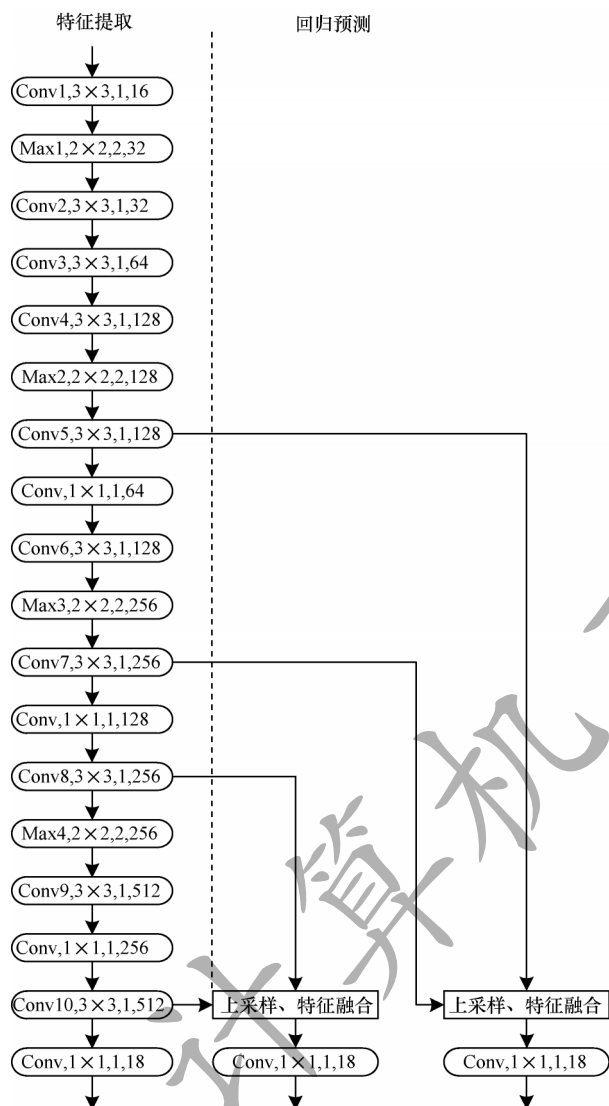


图2 I-YOLOV3-Tiny 网络结构

Fig.2 Structure of I-YOLOV3-Tiny network

由图2可见,改进的网络模型主要分为特征提取和回归预测2个部分。特征提取主要由卷积层和池化层完成,回归预测主要用于预测目标的类别概率和目标边界框坐标。网络结构的改进主要体现在3个方面:根据舰船目标的特点对浅层特征信息进行强化重构,以降低小目标的漏检率;增加改进残差网络(I-ResNet),既能增加网络的深度,又能减少网络参数的计算量;采用金字塔网络多尺度特征融合,可实现不同分辨率特征图的融合,以提高对大小舰船目标的综合检测性能。

### 2.1 浅层特征信息重构

舰船目标视频检测面临的问题是在保证大目标高精度检测的同时,易出现小目标具有较高的漏检率。为此,本文对舰船目标在YOLO网络下提取的特征进行了可视化研究,发现其特征提取部分会产生很多冗余的特征信息,并且随着网络层数的逐渐加深,小目标特征不再存在。显然,对于舰船小目标

的检测,既需要足够多的浅层特征图以提高其检测精度,又需要足够多的语义信息来区分目标和背景特征。

本文利用重构方法在网络的特征提取部分加强了对输入图像浅层特征信息的提取与重构利用。以分辨率为416像素×416像素大小的输入图片为例,在Darknet平台下,利用YOLOV3-Tiny网络模型,用已训练好的权重对舰船测试集进行测试,对卷积网络逐层提取特征图,分别将Conv2、Conv3、Conv4层的感受野进行可视化,如图3所示。

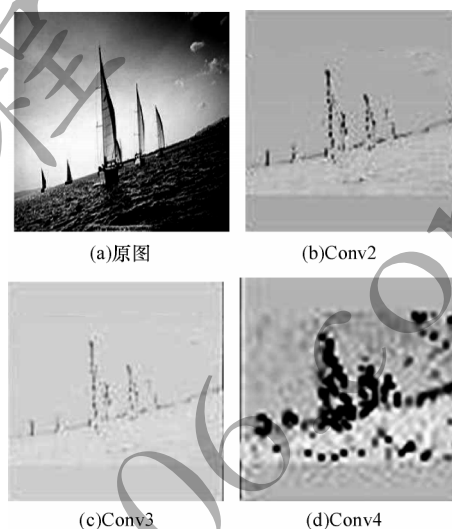


图3 YOLOV3-Tiny 部分卷积层可视化特征图

Fig.3 Feature maps of partial convolutional layers of YOLOV3-Tiny

由图3可知,Conv2层包含了更多的舰船小目标特征信息,如较多的边缘和纹理信息,Conv3层包含的小目标特征较少,而Conv4包含了太多复杂背景特征,对目标检测会产生较大的干扰,且其感受野最大。可见,浅的卷积层对小目标比较敏感,更能提取小目标的形状、纹理等特征,故适当增加浅层网络能更好地保留小目标的有效信息。

为了更有效地提取浅层特征,且充分利用上下层之间的语义信息,增强对Conv2层语义信息的提取效果,在图2中将YOLOV3-Tiny的Maxpooling2层改为Conv3层,增加卷积层Conv4层,卷积步长为1,不改变特征图尺寸,以便更好地提取Conv4层特征,增加特征通道维度,在Conv4层后增加最大池化层,用以改变特征图的尺寸。

### 2.2 改进的残差层

为实现在减少卷积层计算量的同时,增加网络深度以提高检测精度,本文借鉴Resnet提出的残差单元思想<sup>[14]</sup>,在网络结构中增加了3×3卷积层、1×1卷积层和3×3卷积层网络块,改进的部分网络结构块如图4所示。



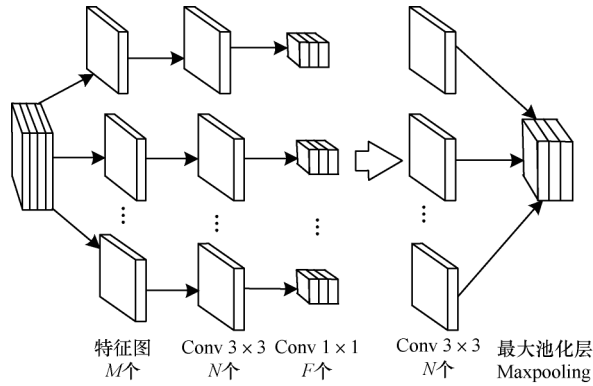


图4 改进的残差层网络块

Fig.4 Network block of improved residual layer

从图4可以看出,第1个3×3卷积层的作用是提取特征,1×1卷积层是一个映射层,可以将输入的特征图映射到较低维度的张量,因为当输入和输出的通道数很大时,卷积核参数会增加很多,而加入1×1卷积层后可以降低输入通道数,由此卷积核参数以及运算复杂度即降底。第2个3×3卷积层的步长为1,并增加了卷积核,故能够提取更多的特征。总之,该网络块结构不仅能增加网络的深度,提高检测精度,还能使网络计算参数减少,网络也更加轻量化,有利于实时性实现。图5所示为改进的网络框架。

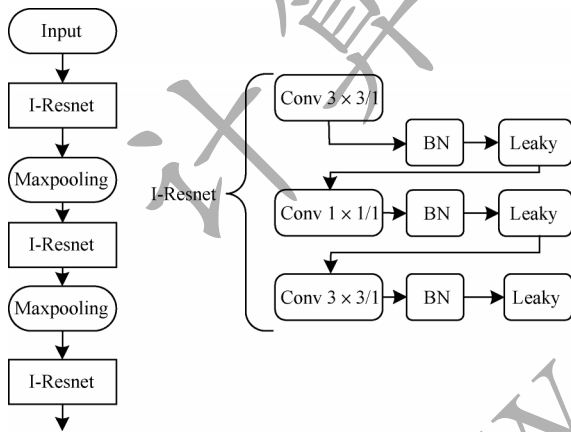


图5 改进的残差层网络框架

Fig.5 Network framework of improved residual layer

图5中的I-Resnet(Improved-Resnet)即改进的残差网络,即图2中的Conv5、Conv6。在图5中,网络框架中的BN(Batch Normalization)层,即批量标准化处理。如果样本不进行BN处理,每一层经过训练后,数据分布会不均匀,网络就需要重新学习新的数据分布规律,会造成网络计算量增大<sup>[15]</sup>。故增加BN层可以在一定程度上解决梯度消失和梯度爆炸问题,防止出现过拟合的情况,同时可以增加训练速度。

BN层之后增加激励函数,本文引入的激励函数是Leaky Relu,其表达式为:

$$f(x) = \begin{cases} x, & x \geq 0 \\ \alpha x, & x < 0 \end{cases} \quad (1)$$

其中: $\alpha$ 为系数,该值是很小的常数。当神经元处于激活状态时,允许非0的梯度存在,这样不会出现梯度消失问题,从而加快收敛速度。

### 2.3 基于特征金字塔网络的多尺度预测

网络低层的目标语义信息比较少,但是目标位置信息准确;而网络高层的目标语义信息丰富,但是目标位置信息粗略。为了在船舰目标检测中能同时以较高精度检测大目标和小目标,本文采用构造多尺度特征金字塔(FPN)的策略,通过网络连接的简单改变,在基本不增加原有模型计算量下,可大幅提升目标检测的性能<sup>[16]</sup>。图6是特征金字塔的原理图,左侧表示下采样,右侧表示上采样2倍,先进行降维处理,然后与上采样2倍后的卷积层相加,经过卷积操作,分别输出13×13和26×26的2个预测尺度特征。

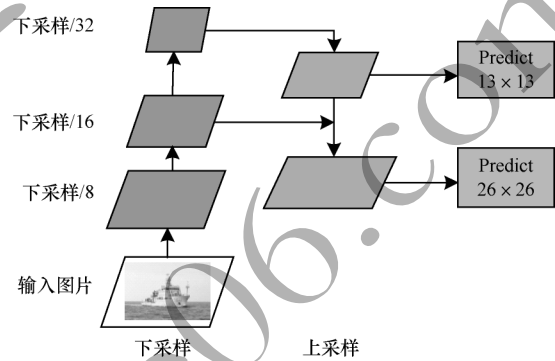


图6 特征金字塔原理示意图

Fig.6 Schematic diagram of feature pyramid principle

可见,多尺度特征金字塔可以将高层的语义信息与浅层的细节信息相结合,分别在不同层进行预测,由此在一定程度上可以平衡不同尺寸的目标,而且可以在不同层上输出对应的目标,不需要经过所有的层运算才输出对应的目标,起到一定的加速作用。因此,本文在图2所示的改进网络结构中,利用FPN设置了3个yolo层进行预测,分别输出12×12、24×24、48×48 3个预测尺度的特征,用于船舰目标的回归检测。并且在每个预测尺度上设有3个预设边界框,即每个网格预先定义3个不同尺寸、不同形状的边界框,然后从这些预先设定的边界框中找到与真实边界框误差最小的框,这样不仅能提高检测速率,而且还可以处理2个对象出现在同一个网格中的问题。Anchor box的大小和形状可以通过K均值聚类(K-Means)<sup>[17]</sup>方式得到,即对目标框的宽和高进行聚类。把最靠近它们的对象进行分类,通过不断学习和迭代,逐渐地更新各聚类中心的值,直至聚类中心不再变化,即准则函数值很小。K-Means算法步骤如下:

**步骤1** 根据所选样本 $y$ ,初始时随机划分 $k$ 个聚类中心。

**步骤2** 根据欧式距离公式,计算所有样本与 $k$ 个聚类中心的距离,样本对象与聚类中心欧式距离 $d$ 计算公式如下:

$$d(y, m_i) = \sqrt{\sum_{j=1}^n (y_j - m_{ij})^2} \quad (2)$$

其中: $m_i$ 是第 $i$ 个聚类中心; $n$ 为样本对象的维度, $y_i$ 、 $m_{ij}$ 是 $y$ 和 $m_i$ 的第 $j$ 个属性值。

**步骤3** 根据距离将样本分配到与其相距最近的聚类中心所在的簇中。

**步骤4** 更新 $k$ 个聚类中心,计算准则函数。准则函数的公式如下:

$$J_c = \sum_{i=1}^k \sum_{y \in \Gamma_i} \|y - m_i\|^2 \quad (3)$$

其中: $\Gamma_i$ 是第 $i$ 个聚类, $i=1, 2, \dots, k$ ;  $J_c$ 反映了用 $k$ 个聚类中心代表 $k$ 样本子集所带来的总的误差平方和。

**步骤5** 判断准则函数是否最小,若准则函数值小,则聚类结束,输出聚类结果;否则,则返回步骤2继续循环,直至准则函数值最小。

### 3 实验与结果分析

#### 3.1 实验数据集和运行环境

为了实现海面舰船目标的快速检测,首先需要构建网络学习的数据集。本文通过采集、拍摄、网上搜索等方式建立了舰船图像数据集,共有2000多幅图像,并将数据集分为训练集和测试集,其中75%作为训练图片,25%作为测试图片,并通过翻转、旋转、裁剪、对比度、色彩变换、高斯噪声等方法对数据集扩充。此外,为了降低误检测率和误识别率,在测试集中加入大量负样本,然后利用标注工具LabelImg手工标注数据集目标,并生成学习文件。数据集的构成如表1所示。

表1 舰船目标图像数据集

Table 1 Image dataset of ship targets

总数据集	训练集 比例/%	测试集 比例/%	测试集正样本 比例/%	测试集负样本 比例/%
2 000	75	25	10	90

本文基于深度学习框架Darknet训练深度学习模型,在64位的UBUNTU16.04计算机上进行评估,该计算机具有CPU Intel® i5的处理器,并在GPU服务器的CUDA9.0环境中进行GPU并行加速计算。

#### 3.2 基于迁移学习的I-YOLOV3-Tiny训练

对于构建的I-YOLOV3-Tiny网络模型的学习训练,由于手动标注的舰船数据集规模较小,直接用来训练I-YOLOV3-Tiny难以得到很好的训练效果,因此本文选用VOC数据集<sup>[18]</sup>,这是PASCAL Visual Object Classes挑战赛公开的数据集,有20类标注目标,共16000张图像。采用迁移学习方法对模型进行预训练,然后把迁移学习得到的模型参数迁移到舰船检测模型的二次训练中,具体的学习过程如图7所示。

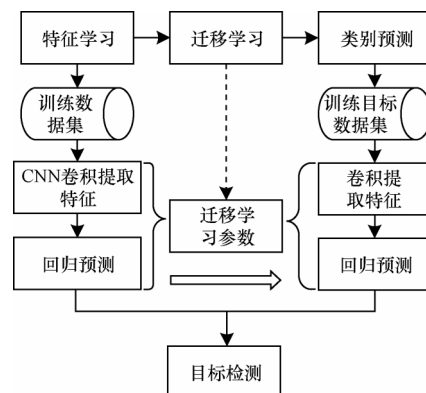


图7 基于迁移学习的I-YOLOV3-Tiny训练过程

Fig.7 Training process of I-YOLOV3-Tiny based on transfer learning

从图7可以看出,左侧为利用VOC数据集进行网络预训练的过程,右侧为利用迁移学习方法和小规模舰船图像训练集进行网络参数微调训练的过程。在网络参数微调的过程中,采用迁移学习和多尺寸训练2种方法,直至训练网络达到理想状态。

#### 3.3 视频检测速率

在实际的海面视频检测中,由于视频帧之间背景单一,视频中存在大量的冗余信息,若每帧图像都重复调用网络模型进行检测识别,则难以实现实时性检测。为了提高视频检测的速率,本文采用图像结构相似性(SSIM)指标来衡量相邻2帧图像的相似程度。当相邻图像帧结构相似性时,可以减少对网络模型的调用次数,从而提高视频帧检测的速率。SSIM主要从亮度、对比度、结构3个方面度量图像的相似性,SSIM算法的计算公式如下:

$$S_{SSIM}(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \quad (4)$$

其中:

$$l(x, y) = \frac{2\mu_x \mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (5)$$

$$c(x, y) = \frac{2\sigma_x \sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad (6)$$

$$s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x \sigma_y + c_3} \quad (7)$$

其中: $x, y$ 代表2幅对比图像; $l(x, y)$ 是亮度比较; $c(x, y)$ 是对比度比较; $s(x, y)$ 是结构比较; $\alpha, \beta, \gamma > 0$ 用来调整亮度、对比度、结构失真度的相对重要性; $\mu_x, \mu_y$ 分别为 $x, y$ 的均值; $\sigma_x, \sigma_y$ 分别为 $x, y$ 的标准差; $\sigma_{xy}$ 为 $x, y$ 的协方差; $c_1, c_2, c_3$ 是不为0的常数。一般设定 $\alpha = \beta = \gamma = 1, c_3 = c_2/2$ <sup>[19]</sup>。SSIM公式可以简化为:

$$S_{SSIM} = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (8)$$

SSIM值的范围为 $[0, 1]$ ,值越大说明2幅图像相似性越大。利用SSIM进行视频目标检测的流程为:以视频的第1帧图像作为参考帧,并进行网络模型的前向运算,并与视频的下一帧进行对比,计算2帧之间的结构相似度。设定结构相似度阈值,若2帧结构相似度大于规定阈值,则不进行前向推理计算,直接输出图像并且附加前一帧的检测结果;若2帧的结构相似度小

于等于规定的阈值,则对当前帧利用网络模型进行前向目标检测,并且将当前帧更新为新的参考帧。

### 3.4 实验结果

在测试集中,目标检测被分为正样本和负样本2类。精确度  $P$  (Precision) 用来衡量目标检测的精度;召回率  $R$  (Recall) 是检测到的真实目标个数与检测到所有目标个数的比值<sup>[20]</sup>;交并比 IoU (Intersection-over-Union) 表示产生的候选框与真实框的重叠率,相关度越高则值越大,完全重叠时值为1<sup>[21]</sup>。检测速率 (FPS) 是指每秒处理帧数, FPS 越大则实时性越高。各指标定义如下:

$$P = \frac{N_{tp}}{N_{tp} + N_{fp}} \quad (9)$$

$$R = \frac{N_{tp}}{N_{tp} + N_{fn}} \quad (10)$$

$$I_{IoU} = \frac{N_{tp}}{N_{tp} + N_{fp} + N_{fn}} \quad (11)$$

其中:  $N_{tp}$  是实际为正样本被判定为正样本的个数;  $N_{fp}$  是实际为负样本被判定为正样本的个数;  $N_{fn}$  是实际为正样本被判定为负样本的个数。

表2是I-YOLOV3-Tiny是否使用迁移学习和多尺寸训练方法的评价结果。由表2可以看出,使用迁移学习和多尺寸训练方法能使目标检测效果更好,并且训练达到收敛的时间也大大缩短。表3是本文算法I-YOLOV3-Tiny与另外4种代表算法检测性能的对比实验结果。

表2 训练结果对比

Table 2 Comparison of training results

训练策略	$P/\%$	$R/\%$	IoU/%	训练时间/h
无	68.6	79.6	62.4	15
有	92.4	88.6	74.1	8

表3 5种算法性能对比

Table 3 Performance comparison of five algorithms %

网络模型	$P$	$R$	IoU
Fast-R-CNN	66.9	65.0	60.9
YOLOV2	77.6	72.0	68.7
YOLOV3	88.7	85.0	77.6
YOLOV3-Tiny	85.4	91.6	71.1
I-YOLOV3-Tiny	92.4	88.6	74.0

由表3可见, Fast-R-CNN<sup>[22]</sup>是使用卷积神经网络直接产生候选框,其实时性和检测精度都低于YOLO网络模型。与YOLOV3-Tiny相比,本文算法的精确度提高7%,交并比IoU值提高3%左右,主要原因是网络特征提取部分加强了对浅层信息的提取,采取的改进残差网络加深了网络,提高了对深层卷积特征的提取。此外,以上检测指标的提高,也与采用的迁移学习策略有关。

图8给出了3组代表性检测结果,每一排为一组。由图8(c)可见,本文算法提高了对小目标的检测能力,在不同背景下,对于大小船舰目标都具有较高的检测性能。由图8(a)、图8(b)可见,YOLOV3、YOLOV3-Tiny容易出现船舰大目标或小目标的漏检问题,且IoU较低。

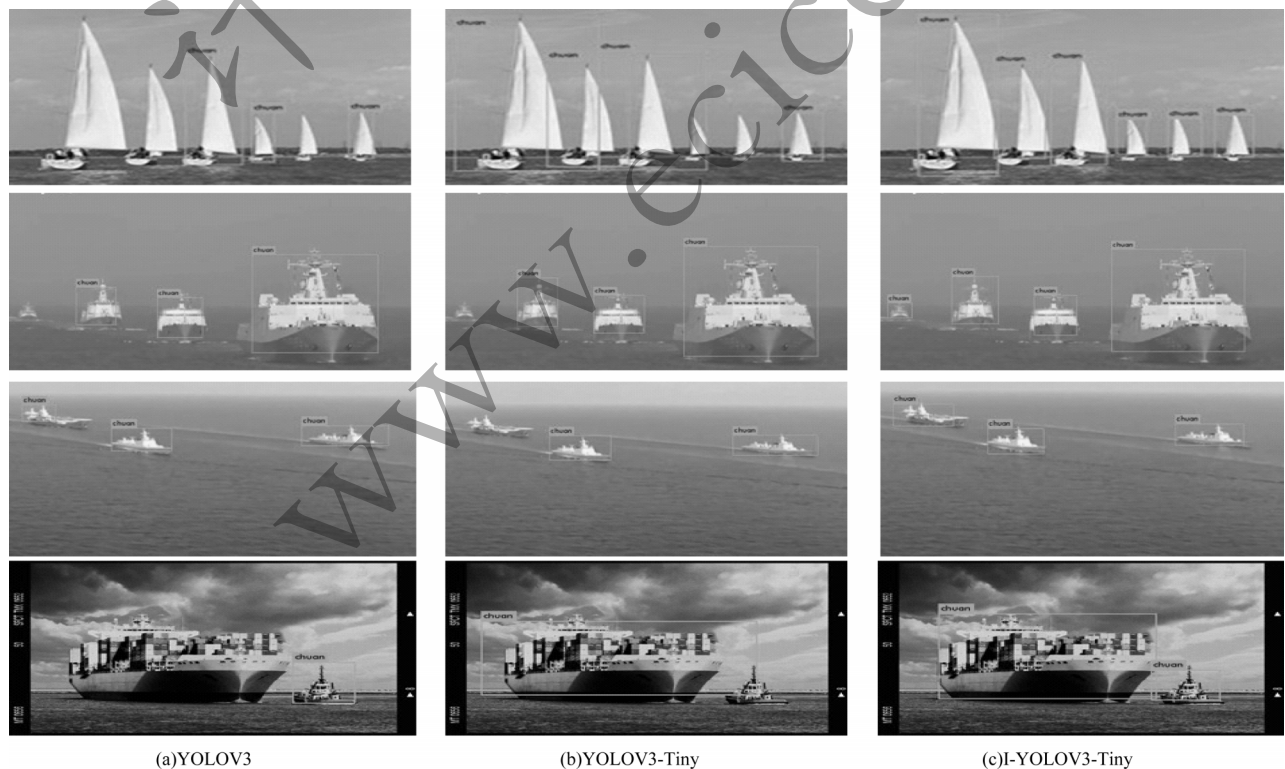


图8 船舰目标检测性能对比

Fig.8 Performance comparison of ship target detection



在GPU平台上,网络模型对舰船目标的检测速度可达每秒几十帧,但是对于海面移动平台的视频处理识别系统,只能采用轻量级设备,不方便配置高耗能的GPU。因此,本文在Intel i5 CPU平台上,分别统计了各种网络模型对舰船视频的处理速度。表4显示的是未采用SSIM与采用SSIM进行视频检测的FPS对比结果。

表4 CPU平台舰船视频平均检测速率

Table 4 Average detection rate of ship video on

网络模型	CPU platform (frame·s <sup>-1</sup> )	
	检测速度	SSIM
YOLOV3	0.5	2
YOLOV3-Tiny	3.0	14
I-YOLOv3-Tiny	2.0	12

从表4可以看出,在CPU平台下采用SSIM进行选择计算检测,本文提出的I-YOLOv3-Tiny其检测速率可以达到12 frame/s,基本满足快速视频检测的需求。其平均检测速率速比传统的YOLOV3要高,而比YOLOV3-Tiny稍低,主要原因是改进的网络比YOLOV3-Tiny网络结构稍深。总之,基于SSIM视频检测算法明显提高了视频检测的速率。

#### 4 结束语

本文基于DarkNet平台,利用深度学习的方法,提出一种基于改进YOLOV3-Tiny网络模型的海面舰船目标检测算法。利用浅层感受野对浅层信息进行强化重构,根据浅层信息提取目标特征,提高对小目标的检测精度。通过增加I-ResNet网络,不仅能增加网络的深度,而且减少网络参数的计算量,提高检测的速率。采用的金字塔网络多尺度特征融合,提高对大小舰船目标的检测性能,利用帧间图像结构相似度进行选择网络前向计算,以显著提高视频帧检测速率,并运用迁移学习方法和多尺寸训练模式,提高网络的学习能力与对不同尺寸目标检测的泛化能力。实验结果表明,该算法可以快速有效地检测海面舰船目标。下一步将研究深度卷积网络结构的精简方法,在满足检测精度的前提下提高算法的实时性。

#### 参考文献

- [1] WANG C B, ZHANG X Y, CONG L Z, et al. Research on intelligent collision avoidance decision-making of unmanned ship in unknown environments[J]. *Evolving Systems*, 2019, 10(4): 649-658.
- [2] 黎经文, 厉小润, 赵辽英. 基于边缘线分析与聚合通道特征的港口舰船检测[J]. *光学学报*, 2019, 39(8): 225-234. LI J Y, LI X R, ZHAO L Y. Port ship detection based on edge line analysis and aggregation channel characteristics [J]. *Acta Optica Sinica*, 2019, 39(8): 225-234. (in Chinese)
- [3] 李浩谊, 马春庭. 基于改进的Scharr算法的海上舰船图像边缘检测[J]. *舰船电子工程*, 2019, 39(3): 103-106. LI H Y, MA C T. Image edge detection of marine ships based on improved Scharr algorithm[J]. *Ship Electronic Engineering*, 2019, 39(3): 103-106. (in Chinese)
- [4] ZHANG Y, LI Q Z, ZANG F G. Ship detection for visual maritime surveillance from non-stationary platforms[J]. *Ocean Engineering*, 2017, 141(1): 53-63.
- [5] 丁鹏, 张叶, 贾平, 等. 基于视觉显著性的海面舰船检测技术[J]. *电子学报*, 2018, 46(1): 127-134. DING P, ZHANG Y, JIA P. Detection technology of sea surface ships based on visual saliency[J]. *Chinese Journal of Electronics*, 2018, 46(1): 127-134. (in Chinese)
- [6] SHI G M, SUO J D. Ship target detection based on visual attention[C]//*Proceedings of International Conference on Signal Processing, Communications and Computing*. Qingdao, China: [s. n.], 2018: 1-4.
- [7] SHAO Z F, WANG L G, WANG Z Y, et al. Saliency-aware convolution neural Network for ship detection in surveillance video[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2019, 30(3): 1-15.
- [8] ZHANG W, HE X J, LI W Y, et al. An integrated ship segmentation method based on discriminator and extractor [J]. *Image and Vision Computing*, 2019, 93(1): 1-12.
- [9] WANG Y C, NING X Y, LENG B H, et al. Ship detection based on deep Learning[C]//*Proceedings of International Conference on Mechatronics and Automation*. Tianjin, China: [s. n.], 2019: 275-279.
- [10] 马啸, 邵利民, 金鑫, 等. 改进的YOLO模型及其在舰船目标识别中的应用[J]. *电讯技术*, 2019, 59(8): 869-874. MA X, SHAO L M, JIN X, et al. Improved YOLO model and its application in ship target recognition [J]. *Telecommunications Technology*, 2019, 59(8): 869-874. (in Chinese)
- [11] 赵春晖, 周瑶. 基于改进Faster R-CNN算法的舰船目标检测与识别[J]. *沈阳大学学报(自然科学版)*, 2018, 30(5): 366-371, 380. ZHAO C H, ZHOU Y. Ship target detection and recognition based on improved Faster R-CNN algorithm[J]. *Journal of Shenyang University (Natural Science Edition)*, 2018, 30(5): 366-371, 380. (in Chinese)
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[EB/OL]. [2020-07-10]. <https://arxiv.org/abs/1506.02640>.
- [13] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. [2020-07-10]. <https://arxiv.org/abs/1804.02767>.
- [14] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE Press, 2016: 770-778.
- [15] HUNG K Y, ZHANG Z K, JIANG J M. Real-time image super-resolution using recursive depthwise separable convolution network[J]. *IEEE Access*, 2019, 7: 99804-99816.

(下转第297页)

(上接第 289 页)

- [16] 彭秀平,全其胜,林洪彬,等. 一种面向散乱点云语义分割的深度残差-特征金字塔网络框架[J]. 自动化学报, 2019, 39(1): 1-10.  
PENG X P, TONG Q S, LIN H B, et al. A deep residual-feature pyramid network framework for semantic segmentation of scattered point clouds[J]. Acta Automatica Sinica, 2019, 39(1): 1-10. (in Chinese)
- [17] TANG J L, WANG D. Weed identification based on K-means feature learning combined with convolutional neural network[J]. Computers and Electronics in Agriculture, 2017, 135(1): 63-70.
- [18] EVERINGHAM M, GOOL L, WILLIAMS C K I, et al. The pascal visual object classes challenge[J]. International Journal of Computer Vision, 2010, 88(2): 303-338.
- [19] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [20] 周志华. 机器学习[M]. 北京:清华大学出版社, 2016.  
ZHOU Z H. Machine learning [M]. Beijing: Tsinghua University Press, 2016. (in Chinese)
- [21] DOLLAR P, WOJEK C, SCHIELE B, et al. Pedestrian detection: a benchmark [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA: IEEE Press, 2009: 304-311.
- [22] 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 and Machine Intelligence, 2017, 39(6): 1137-1149.

编辑 索书志