

梯度纹理直方图与多层感知器船舶快速检测

董超^{1,2}, 冯俊健³, 田联房³, 郑兵^{1,2}

- (1. 国家海洋局南海调查技术中心, 广东 广州 510300;
2. 国家海洋局南海维权技术与应用重点实验室, 广东 广州 510300;
3. 华南理工大学 自动化科学与工程学院, 广东 广州 510641)

摘要: 在动态的复杂环境中, 受背景建模失效影响, 传统船舶目标检测方法的精度较低。针对该问题, 提出一种基于梯度纹理直方图特征与多层感知器的船舶快速检测算法。该算法利用多层感知器将目标的梯度与纹理的直方图进行特征融合, 为船舶目标构建特征空间。首先, 基于二值梯度的特征训练船舶候选区模型, 以快速生成具有高召回率的少量船舶候选窗口, 并在每个候选窗口提取梯度纹理直方图特征; 其次, 设计一个多层感知器作为船舶分类器, 对提取到的梯度纹理直方图特征进行判别。实验结果表明, 该算法在多个海上场景中船舶检测平均精确率达 90.0%, 平均执行时间为 20.4 ms/frame, 有效实现海上船舶精确与快速的检测。

关键词: 快速船舶检测; 梯度纹理直方图; 多层感知器

中图分类号: TP391.4 **文献标志码:** A **DOI:** 10.3788/IRLA201948.1026004

Rapid ship detection based on gradient texture features and multilayer perceptron

Dong Chao^{1,2}, Feng Junjian³, Tian Lianfang³, Zheng Bing^{1,2}

- (1. South China Sea Marine Survey and Technology Center, State Oceanic Administration, Guangzhou 510300, China;
2. Key Laboratory of Technology for Safeguarding of Marine Rights and Interests and Application, State Oceanic Administration, Guangzhou 510300, China;
3. School of Automation Science and Engineering, South China University of Technology, Guangzhou 510641, China)

Abstract: Aiming at the issues of low ship detection rate caused by the failure of background modeling in the dynamic complex environment of traditional ship detection methods, a rapid ship detection algorithm based on gradient texture histogram features and multilayer perceptron was proposed. The feature fusion between gradient and texture histogram of the target was performed using multilayer perceptron, constructing the feature space for ship targets. Firstly, the region proposal model based on binarized normed gradient feature was trained to quickly generate a small number of ship candidate windows with high recall rate and then the gradient texture histogram features were extracted from each

收稿日期: 2019-06-05; 修订日期: 2019-07-15

基金项目: 国家自然科学基金-浙江两化融合联合基金(U180920009); 全球变化与海气相互作用专项;
海洋公益性行业科研专项(201305028)

作者简介: 董超(1982-), 男, 副研究员, 博士后, 主要从事无人艇研制与应用、图像处理及模式识别方面的研究。

Email: dongchaoxj888@126.com

candidate window. Secondly, a multilayer perceptron was designed as a ship classifier to distinguish the gradient texture histogram features. Experimental results show that the proposed algorithm has an average precision of 90.0% and an average time of 20.4 ms/frame in multiple maritime scenes, which effectively realizes rapid ship detection in maritime scenes.

Key words: rapid ship detection; gradient texture histogram; multilayer perceptron

0 引言

目前,基于遥感图像^[1-2]或者红外图像^[3]的船舶检测对船只目标的外观描述信息量较少,容易存在误检。而基于光电平台的船舶检测是一种比较直观、有效的海上监控方法,它既能有助于对航道或者码头上船舶的管理,还能在监控区域内及时发现可疑船舶目标。从光电平台的监控区域内获取的图像或视频中利用船舶检测算法自动发现船舶目标并返回有效船舶数据,能极大地提高管理人员的工作效率。然而,基于光电平台采集到的海上视频数据会受到海上的各种环境影响,从而导致船舶检测存在以下困难:(1) 光电平台的晃动、海上天气和光照等变化导致视频图像中的船舶目标不清晰;(2) 海浪、海岸、建筑、云等复杂背景影响船舶检测的精度;(3) 船舶目标外观的多样性导致难以精确检测。当前,基于光电平台的船舶检测方法主要以背景建模为基础延伸出多种方法^[4]。张仲瑜等人^[5]利用边缘和灰度等多特征从海天背景下提取船舶目标,该方法在复杂背景中会受到干扰。进一步的研究发现,在广阔的海域中,远距离的船舶目标一般是出现在海天线区域,因此海天线提取结合背景建模的方法被广泛应用于船舶检测从而提高检测准确率。Zhang 等人^[5]采用海天线检测、背景建模和背景相减的方式进行船舶的检测。这种方法提高了船舶的检测准确率,但是缺乏对目标判别的能力,容易受到其他目标如鸟鱼等动物的干扰,而且,海天线的引入使应用场景受到限制。Li 等人^[6]结合视觉显著性模型,提出了基于统计先验知识的显著区域检测。但是该显著性区域的计算容易受到其他背景如海浪等干扰,并且实时性不能满足海上船舶检测的需要。

在其他目标检测领域,Wang 等人^[7]提出联合 Histograms of Oriented Gradients (HOG) 和 Local Binary Pattern (LBP) 的特征实现部分遮挡场景中的

行人检测。程全等人^[8]利用目标的运动特性通过分块投影进行运动目标检测,但该方法只适应于目标平移运动。陆福星等人^[9]利用背景自适应结合多特征实现弱小目标的检测,虽然该方法利用序列图像消除大部分假目标,但准确率依然不能满足实际场景中的船舶检测。吴天舒等人^[10]基于改进的 SSD 网络实现轻量级的目标检测,虽然该方法能在 ARM 设备上运行,但依然需要大量数据对网络进行训练。以上所提及的这些方法都不能直接应用于海上船舶检测,原因如下:(1) 目前已标注的海上船舶数据集的规模受到限制;(2) 在船舶检测中,实时性和准确率都有较高要求。

因此,针对上述方法在复杂海上场景中实现船舶检测存在的局限性,文中提出基于梯度纹理特征和多层感知器实现快速的船舶检测。在特征提取方面,为了应对海上场景中光照变化和船舶外表多样性带来的表征困难,采用对光照具有鲁棒性的梯度信息和纹理信息得到的直方图统计特征向量表征船舶。在分类器设计方面,作为一种神经网络模型,多层感知器具备良好的函数逼近能力^[11]。通过实验调试,设计一个轻量级的多层感知器模型,使其在船舶检测中对候选船舶样本实现快速和高精度的分类。

1 算法概述

相对于从背景建模的方式进行船舶检测,文中针对被检测目标的特征,为船舶建立特征模型,旨在使文中算法能适用于更广泛的海上场景。因此,船舶检测过程转化为分类问题,即在输入的视频帧中搜索与船舶特征模型相似的区域,算法流程图如图 1 所示。首先,利用基于梯度的候选框提取方法生成少量具有高召回率的候选框,减少后续的计算量,并降低对特定背景因素(如海天线和静态背景条件)的依赖。接着,从批量候选框内图像中提取得到的梯度纹理特征,送入已训练好的多层感知器分类模型进行判别。最终,输出船舶目标位置信息。

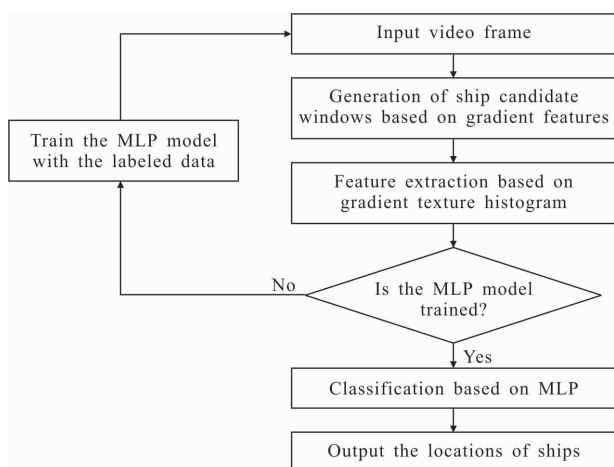


图 1 基于梯度纹理直方图和多层感知器的船舶快速检测算法流程图

Fig.1 Flowchart of rapid ship detection based on gradient texture histogram and multilayer perceptron

2 基于梯度纹理特征和多层感知器的船舶快速检测

2.1 基于梯度的船舶候选框提取

为了克服背景建模的船舶检测方法，例如海天线检测对海天线或静态背景等背景因素的依赖性，很多基于目标候选框的方法被提出^[12]。其中，滑动窗口搜索是一种快速生成候选窗口的方法，在单尺度条件下需要生成大量的候选窗口来保证目标检测的高召回率。海上场景中由于视场较大，船舶目标呈现多尺度的特性，因此多尺度的滑动窗口搜索会造成沉重的计算成本。Cheng 等人^[13]提出基于二值化梯度特征的候选框模型，极大提高目标候选框的生成速度。该模型以滑动窗口的方式产生不同尺寸的候选样本 p_i ，将每个候选框 p_i 调整为对目标抽象描述的尺寸(8×8)后计算梯度特征 g_p 。梯度特征 g_p 的每个值等于 $\min(|g_x|+|g_y|, 255)$ 。其中， g_x 和 g_y 分别为水平方向梯度和垂直方向梯度。每个候选框由公式(1)来评价得分，得分越高，候选框包含船舶目标的可能性越高。利用已标注的船舶数据，把与船舶真实位置交并比大于等于 0.5 的候选样本记为标签 $y_i=1$ ，少于 0.5 的候选样本记为标签 $y_i=-1$ 。因此，公式(2)可以利用坐标下降算法^[14]进行优化，从而得到船舶候选框模板 w_p 如图 2 所示。

$$f_s(p) = w_p^T g_p \quad (1)$$

$$w_p = \operatorname{argmin}_{w_p} \|w_p\|_1 + \sum_{i=1}^n \{\max[0, 1 - y_i f_s(p_i)]\}^2 \quad (2)$$

式中： p 为候选框； w_p 为船舶候选框模板； g_p 为梯度特征； $\|\cdot\|_1$ 代表一范数，目的是将 w_p 尽可能稀疏化； n 为候选样本总数； p_i 为第 i 个候选样本和 y_i 为对应的标签。

为了加快目标候选框的生成速度，对梯度特征 g_p 和船舶候选框模板 w_p 以二进制形式表示，从而能使计算效率极大提高。因此，梯度特征 g_p 以二进制中前 N_g 位的二值化梯度特征 b_i 进行近似，如公式(3)。使用算法 1，得到船舶候选框模板 w_p 的二值化结果，形式如公式(4)。最终，公式(1)可重写如公式(5)所示。

$$g_p \approx \sum_{i=1}^{N_g} 2^{8-i} b_i \quad (3)$$

$$w_p \approx \sum_{i=1}^{N_w} \alpha_i a_i \quad (4)$$

$$f_s(p) \approx \sum_{i=1}^{N_w} \alpha_i \sum_{j=1}^{N_g} 2^{8-j} a_i^T b_j \quad (5)$$

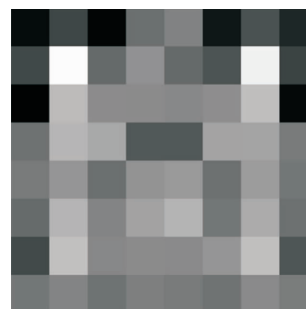


图 2 训练好的基于梯度特征的船舶候选框模型

Fig.2 Trained proposal model of ship based on gradient features

算法 1 二值化候选框模板 w_p 。

输入：已训练的 w_p ，

(1) 初始化残差 $e = w_p$

(2) 执行(3)、(4)、(5)迭代 N_w 次

(3) $a_i = \operatorname{sign}(e)$

(4) $\alpha_i = \langle a_i, e \rangle / \|a_i\|^2$

(5) $e \leftarrow e - \alpha_i a_i$

输出： $\{\alpha_i\}_{i=1}^{N_w}, \{a_i\}_{i=1}^{N_w}$

2.2 梯度纹理直方图特征的提取

丰富的特征表示可以提高目标检测的准确度，同时降低虚警率。文中梯度纹理直方图特征主要由 HOG 特征^[7]和 Extended Local Binary Pattern (ELBP) 特征^[15]组成。该特征既能描述目标的整体形状，又能描述目标的局部纹理，形成能表征目标且具备光照

鲁棒性的整体-局部特征向量,如公式(6)所示:

$$x=[\phi_H(p), \phi_E(p)] \quad (6)$$

式中: p 表示候选框; $\phi_H(p)$ 表示从候选框 p 中提取得到的HOG特征向量; $\phi_E(p)$ 表示从候选框 p 中提取得到的ELBP特征向量。

2.2.1 $\phi_H(p)$ 特征提取

基于HOG特征的 $\phi_H(p)$ 对光照具有鲁棒性,对描述海上船舶目标轮廓具有优势,如图3(c1)所示。以下介绍 $\phi_H(p)$ 的特征提取。

(1) 梯度聚合统计

使用卷积核 $[-1, 0, 1]$,从水平方向和垂直方向对候选框内图像进行卷积运算,得到梯度图。把梯度图划分成 $[x/k] \times [y/k]$ 个胞块,其中 $[\cdot]$ 代表取整操作。利用公式(7),对胞块中每个像素点的梯度方向 $\theta(x, y)$ 和梯度幅值 $r(x, y)$ 进行映射与统计。

$$C(i, j)_b = \sum_{(x, y) \in p} w_k(x, y) r(x, y) l_b[\theta(x, y)] \quad (7)$$

$$l_b(\theta) = \begin{cases} 1 & \text{if } b = \left(\left\lfloor \frac{m\theta}{\pi} \right\rfloor \bmod m \right) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

式中: $C(i, j)_b$ 表示从梯度图聚合统计后得到的胞块; $(i, j) = ([x/k], [y/k])$, $b \in \{1, \dots, m\}$ 为梯度方向统计区间, m 设为9。 $w_k(x, y)$ 是一个周期为 k 的高斯权重, $(k/2, k/2)$ 为权重中心, k 设为8。

(2) 块归一化

不同胞块中光照条件有差异,因此产生梯度幅值的统计量数量级存在差异。对每个 2×2 相邻的胞块进行归一化,交叠为1,即每个胞块参与邻域里的归一化。设 2×2 胞块形成的特征向量为 $v \in \mathcal{R}^n$,其中 $n = 2 \times 2 \times m$ 。归一化采用L2范数如公式(9), ε 为一个微小常量,防止除数为0。

$$v = v / \sqrt{\|v\|_2^2 + \varepsilon^2} \quad (9)$$

归一化后,则得到了维度为 $(([x/k]-1) \times ([y/k]-1) \times 2 \times b)$ 的 $\phi_H(p)$ 特征。

2.2.2 $\phi_E(p)$ 特征提取

基于ELBP特征的 $\phi_E(p)$ 具备光照不变性和局部旋转不变性,体现目标表面的结构属性,并结合梯度信息的特征形成对目标从整体轮廓到局部结构的特征向量。计算 $\phi_E(p)$,需要先计算候选框 p 的纹理特征图,如图3(b2)所示。接着,根据船舶目标的特性,将纹理特征图划分为 $M \times N$ 个区域,文中将 $M \times N$ 设为 3×3 ,即把船舶目标垂直方向和水平方向分别划分3个区域,并对每个区域内的纹理特征值进行统计,形成具备统计特性的纹理直方图特征,如图3(c2)所示。同时,将各直方图特征串联,得到描述目标的纹理特征向量 $\phi_E(p)$ 。纹理特征图每个特征值

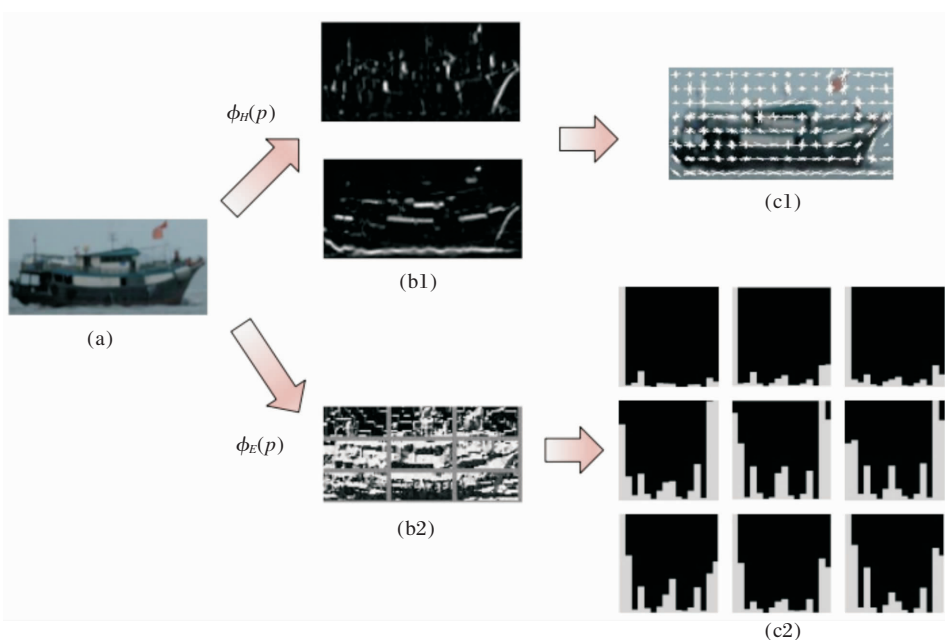


图 3 纹理直方图特征提取过程:(a)候选框图像;(b1)水平方向与垂直方向梯度图;(b2)纹理特征图;(c1)梯度特征;(c2)纹理特征直方图

Fig.3 Extraction of texture histogram features: (a) Image in proposal; (b1) gradient maps in horizontal and vertical directions;

(b2) texture feature map; (c1) gradient features; (c2) histogram of texture features

的计算如公式(10)所示。以坐标 (x,y) 为中心定义一个半径为 r 的圆形窗口,在该窗口邻域内使用双线性插值等间隔采样 n 个像素点。用获得的采样点像素值 $g(x,y)$,与中心点像素值 $g(x,y)$ 进行比较,并进行编码,如公式(11)、(12)所示。通过该方式,使纹理特征对光照具备鲁棒性。此外,通过循环右移获得编码的最小值,具备局部旋转不变性。

$$T_{\min}(x,y)=\min\{ROR_j[T(x,y)]\} \quad (10)$$

$$T(x,y)=\sum_{i=1}^n 2^{i-1}s[g(x,y)-g(x,y)] \quad (11)$$

$$s(a)=\begin{cases} 1 & \text{if } a \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

式中: $T_{\min}(x,y)$ 为像素点坐标 (x,y) 的纹理特征值; $ROR_j(\cdot)$ 表示循环右移 j 位; $T(x,y)$ 为纹理编码值; $g(x,y)$ 为坐标 (x,y) 的像素值; $g(x,y)$ 为以坐标 (x,y) 为中心的 i 个采样点像素值。

2.3 基于多层感知器的分类器设计

神经网络是由具有适应性的简单单元组成的互连网络,它能模拟生物神经系统对真实世界物体作出的交互反应。根据神经网络的结构特点可分为前馈型网络、反馈型网络和竞争学习网络。多层感知器,由多个感知器分层组合而来,作为一种典型的前馈型网络,能通过多层网络的映射学习到任意复杂的非线性函数^[11],从分类角度来看,多层感知器可以看作是一种通用的非线性分类器设计方法,文中按公式(13)设计多层感知器的分类模型。

$$y=f(x|\theta) \quad (13)$$

$$\theta=\{w^{(i)},b^{(i)}|i=1,\dots,m\} \quad (14)$$

$$w^{(i)}=[w_1^{(i)} \dots w_j^{(i)}] \quad (15)$$

$$b^{(i)}=\begin{bmatrix} b_1^{(i)} \\ \dots \\ b_j^{(i)} \end{bmatrix} \quad (16)$$

式中: y 为输出向量,向量中每个值代表对应类别的得分值; x 表示输入特征向量; θ 为多层感知器中的可学习参数集合,其中 $w^{(i)}$ 和 $b^{(i)}$ 分别表示第 i 层网络的权重参数矩阵和偏置参数矩阵; $w_j^{(i)}$ 为第 i 层网络第 j 个感知器权重参数向量; $b_j^{(i)}$ 为对应的偏置参数; m 为网络层数。

多层感知器可看作由多个感知器神经元组合而成,如图4所示。对于第 i 层网络的输入输出模型可

用公式(17)进行表示。因此,公式(13)可展开为多层神经元模型的输入输出映射迭代,如公式(19)所示。

$$f^{(i)}(x^{(i)})=x^{(i+1)}=\sigma(x^{(i)T}x^{(i)}+b^{(i)}) \quad (17)$$

$$\sigma(z)=\frac{1}{1+e^{-z}} \quad (18)$$

$$y=f^{(m)}\{\dots f^{(i)}\dots [f^{(0)}(x^{(0)})]\} \quad (19)$$

式中: $f^{(i)}(\cdot)$ 表示第 i 层网络的输入输出映射; $\sigma(\cdot)$ 代表激活函数,采用 Sigmoid 函数,如公式(18)所示。

文中利用公式(19)多层神经元模型学习海上船舶与背景样本的类别映射关系,得到作为判别性依据的船舶分类模型。因此,该模型的优化目标为最小化船舶样本与背景样本的错分类率。通过反向传播算法^[16]最小化平方误差,得到如公式(20)所示的多层感知器参数。

$$\theta=\arg\min_{\theta} \frac{1}{2} \sum_{k=1}^n \|y_k^{GT} - y_k\|^2 \quad (20)$$

式中: y_k 为第 k 个训练样本通过公式(19)映射得到的预测标签; y_k^{GT} 为训练样本的真实标签; n 代表训练样本总数。

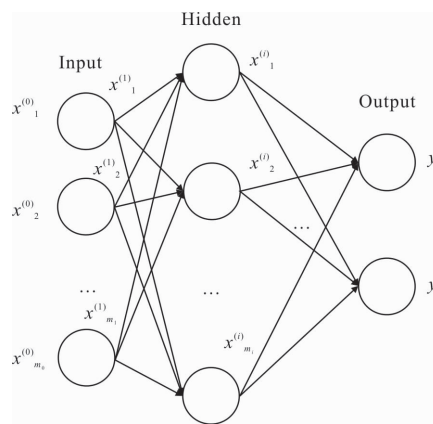


图4 多层感知器网络模型。每个圆圈代表一个感知器模型,通过感知器层层非线性映射,达到强大的非线性函数表达能力

Fig.4 Network model of multilayer perceptron. Each circle denotes a perceptron model, which achieves a powerful representation of nonlinear functions through a non-linear mapping of the perceptron layers

3 实验结果与分析

文中收集了来自南海、三亚等地和一些公开的船舶视频数据,并用矩形框标注了船舶的位置信息,制作成实验所用的船舶数据集。通过对船舶视频数

据的采样, 实验船舶数据集由 1 500 张船舶图片组成, 其中 1 000 张图片作为训练集, 500 张图片用作测试集, 包含了目标紧邻与局部、视角变化、海浪和尺度变化等不同场景。

文中以平均执行时间(Time)和平均精度(Average precision, AP) 作为衡量船舶检测实时性和鲁棒性的指标, 其中平均执行时间测量算法处理每帧图片所需的时间, 平均精度以交并比大于等于 0.5 的情况下 Precision-Recall 曲线下的包围面积作为量化指标, 从而测量算法得到的目标框包围目标的精确率。

此节包含四个对比实验, 包括:(1) 船舶候选框实验对比;(2) 多层感知器 (MLP) 网络结构实验对比;(3) 不同特征实验对比;(4) 不同检测算法实验对比, 从船舶候选框、船舶分类器、船舶特征表征和算法整体性能等四个方面分析文中算法在船舶检测中的实时性和鲁棒性。

实验平台配置为: Intel(R) Core(TM) i7-6820HQ CPU 2.70 Hz, 16 GB 内存, Visual Studio 2015 C++ 和 OpenCV 2.4.13 视觉库。实验中图片处理尺寸为 240×320(height×width)。

3.1 船舶候选框实验对比

海上场景视野广阔, 船舶目标通常表现出多种尺度。利用滑动框窗口对目标进行采样会产生冗余计算。因此, 在场景中预先提取目标候选框有助于提高目标的搜索效率。文中利用基于二值梯度特征的候选框模型^[13]训练船舶候选框模型(记为BINGship), 可快速提取具有高召回率的船舶候选位置。结合表 1 与图 5 分析可得, 在符合交并比(IoU)大于等于 0.5 的情况下, 船舶候选框模型 BINGship 以平均每张图片 50 个候选框实现 96.1% 的召回率, 执行时间(9 ms)最短。而在候选窗口数大于 150 后, Selective Search^[17]的召回率不再增加, 固定在 85.2%, Objectness^[18]虽然召回率在持续增长, 但由于执行时间太长(6 500 ms), 不予考虑。

表 1 不同候选框模型对船舶目标候选的性能对比

Tab.1 Performance comparison of different proposal models in ship target proposal

Methods	Time/ms	Recall
BINGship ^[13]	9	0.961
Objectness ^[18]	6 500	0.918
Selective search ^[17]	170	0.852

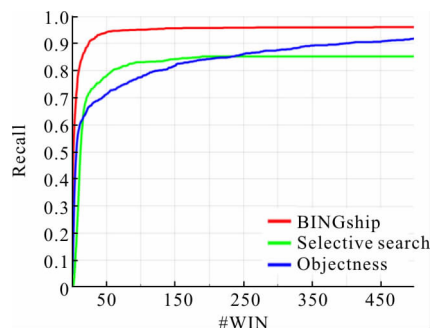


图 5 在交并比大于等于 0.5 的情况下, 不同候选框生成模型的检测窗口数与召回率关系曲线

Fig.5 Tradeoff between #WIN and Recall in different proposal generation models with the IoU over 0.5

3.2 多层感知器网络结构实验对比

多层感知器具有强大的非线性映射能力, 能够逼近复杂函数^[11], 但是具体的网络模型(多少层网络, 每层多少感知器)需要经过实验调试。因此, 对多层感知器的结构进行多组实验, 综合平均时间(Time)平均精度(AP)和网络模型大小, 选择最优模型结构。

文中把目标检测简化为目标分类问题, 故场景中只有船舶和背景两类目标。采用 HOG-LBP 特征作为多层感知器的输入, 先从最简单的 $n-2$ 网络结构(n 为特征输入维度)进行实验, 并逐步增加网络的层数。如表 2 所示, 多层感知器 MLP₃ 具有 4 层网络

表 2 不同网络结构的多层感知器在船舶检测中的性能对比(其中, 网络结构中的 n 代表输入的特征维度, 以“-”作为每层感知器的连接符)

Tab.2 Performance of multilayer perceptrons with different network structures in ship detection (Among them, n in the network structure indicates the feature dimension of the input, and "-" is used as the connector of each layer of the perceptron)

Models	Structure	Precision	Recall	AP	Time/ms
MLP ₁	$n-2$	0.945	0.715	0.703	18.4
MLP ₂	$n-256-2$	0.986	0.957	0.902	28.0
MLP ₃	$n-256-256-2$	0.947	0.924	0.887	32.6
MLP ₄	$n-256-256-256-2$	0.949	0.944	0.898	35.6
MLP ₅	$n-32-2$	0.972	0.950	0.900	20.4

结构,其 AP 值开始下降。同样,多层感知器 MLP_4 的 AP 值也较 MLP_2 有所下降。这是由于网络层数加大,会造成一定的过拟合现象。 MLP_5 以更轻量的网络结构($n=32-2$)实现与 MLP_2 相差无几的 AP 值,但平均时间更少。为了减少模型参数,文中采用更简单的 MLP_5 的网络结构作为船舶检测的船舶分类器。图 6 为不同网络结构的 MLP 在船舶检测中的 Precision-Recall 曲线图,描述了不同 MLP 结构在检测过程中对精确率和召回率的影响。

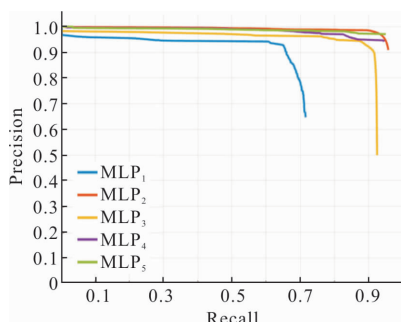


图 6 不同网络结构的多层感知器在船舶检测中的 Precision-Recall 曲线图。不同颜色的曲线代表不同的网络结构,网络结构细节对应表 2

Fig.6 Tradeoff between precision and recall in different network structures of multi-layer perceptron in ship detection. The curves of different colors denote different network structures, and the network structure details correspond to Table 2

3.3 不同特征实验对比

不同特征的优势各不相同,此节采用 MLP_5 的网络结构对各特征进行实验试验。HOG 特征^[7]分块统计目标的梯度量,从而能描述目标的局部轮廓和边缘。LBP 特征计算领域中灰度的对比性,描述目标的局部纹理,具有一定的光照鲁棒性。如表 3 所示,文中算法(HOG-LBP+MLP)组合 HOG 特征与 LBP 特征,结合 MLP 分类器实现了最高的 AP(0.900),而只用 HOG 特征或者 LBP 特征都使 AP 明显降低。这表明,综合使用 HOG-LBP 特征,比单纯用 HOG 特征或 LBP 特征,更能有效表达目标信息。虽然,与单一特征相比,提取 HOG-LBP 特征相对费时。但是,三种特征的提取过程都表现出较好的实时性,对算法效率整体影响不大。因此,将 AP 作为文中船舶检测算法性能评估的首要因素。图 7 为不同特征联合多层感知器在船舶检测中的 Precision-Recall 曲线图,

描述了在统一 MLP 结构下,采用不同特征对船舶检测的精确率和召回率的影响。

表 3 不同特征结合 MLP_5 的网络结构的船舶检测性能对比

Tab.3 Comparison of ship detection performance with different features combined with MLP_5 network structure

Methods	Precision	Recall	AP	Time/ms
Proposed (HOG-LBP+MLP)	0.972	0.950	0.900	20.4
HOG+MLP	0.919	0.855	0.798	15.4
LBP+MLP	0.870	0.803	0.757	13.2

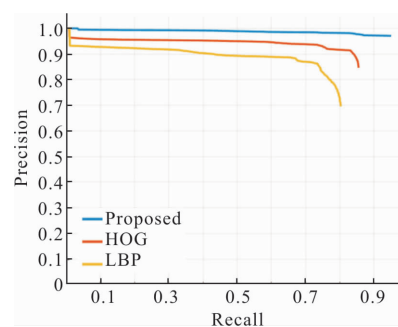


图 7 不同特征联合多层感知器在船舶检测中的 Precision-Recall 曲线图。不同颜色的曲线代表不同的特征

Fig.7 Tradeoff between precision and recall in different features combined with multilayer perceptron in ship detection. Curves of different colors denote different features

3.4 不同检测算法实验对比

将文中算法与几种经典检测算法进行对比,如表 4 所示。四种方法都能满足实时性的要求,因此主要将 AP 作为评价指标。文中算法获得最高的 AP (0.900),远高于其他三种算法。图 8 为各检测算法在多种场景下(目标紧邻与局部、视角变化、海浪和尺度变化)的检测结果。

表 4 不同检测算法的船舶检测性能对比

Tab.4 Comparison of ship detection performance with different detection algorithms

Methods	Precision	Recall	AP	Time/ms
Proposed (HOG-LBP+MLP)	0.972	0.950	0.900	20.4
HOG-LBP+SVM	0.765	0.846	0.697	19.5
HOG+SVM	0.689	0.894	0.592	14.6
GMM	0.189	0.617	0.267	11.7

GMM(Gaussian mixture model)能对静态背景下的运动目标实现有效的检测。一旦背景变化,该方法假阳性会急速上升如图 8(e5)所示。即便目标处于静止状态,也常常出现漏检,如图 8(e2)和图 8(e4)所示。此外,该算法检测的目标矩形框并不能很好地包围实际目标,如图 8(e1)和图 8(e3)所示。综合来看,该算法 AP 值较低,只有 0.267,性能较差。

HOG+SVM 算法结合 HOG 特征与 SVM 分类器进行目标检测,AP 值为 0.592,比 GMM 效果稍好。但由于特征单一,容易受到目标紧邻、海浪影响,出

现误检,如图 8(d1)和图 8(d5)所示。此外,该算法对目标的描述不够全面,无法很好地产生完整的包围框,如图 8(d2)和图 8(d3)所示。

HOG-LBP+SVM 算法综合 HOG 与 LBP 特征,并使用 SVM 分类器进行目标检测。如图 8(c5)所示,该算法对背景抑制效果比 HOG+SVM 算法要好,但对船舶包围框的生成效果依然不理想。该算法 AP 值为 0.697,高于 HOG+SVM 的算法,由此看出 HOG-LBP 能提高船舶检测性能。该算法 AP 值低于文中算法,主要差别在于分类器的结构设计。综上,

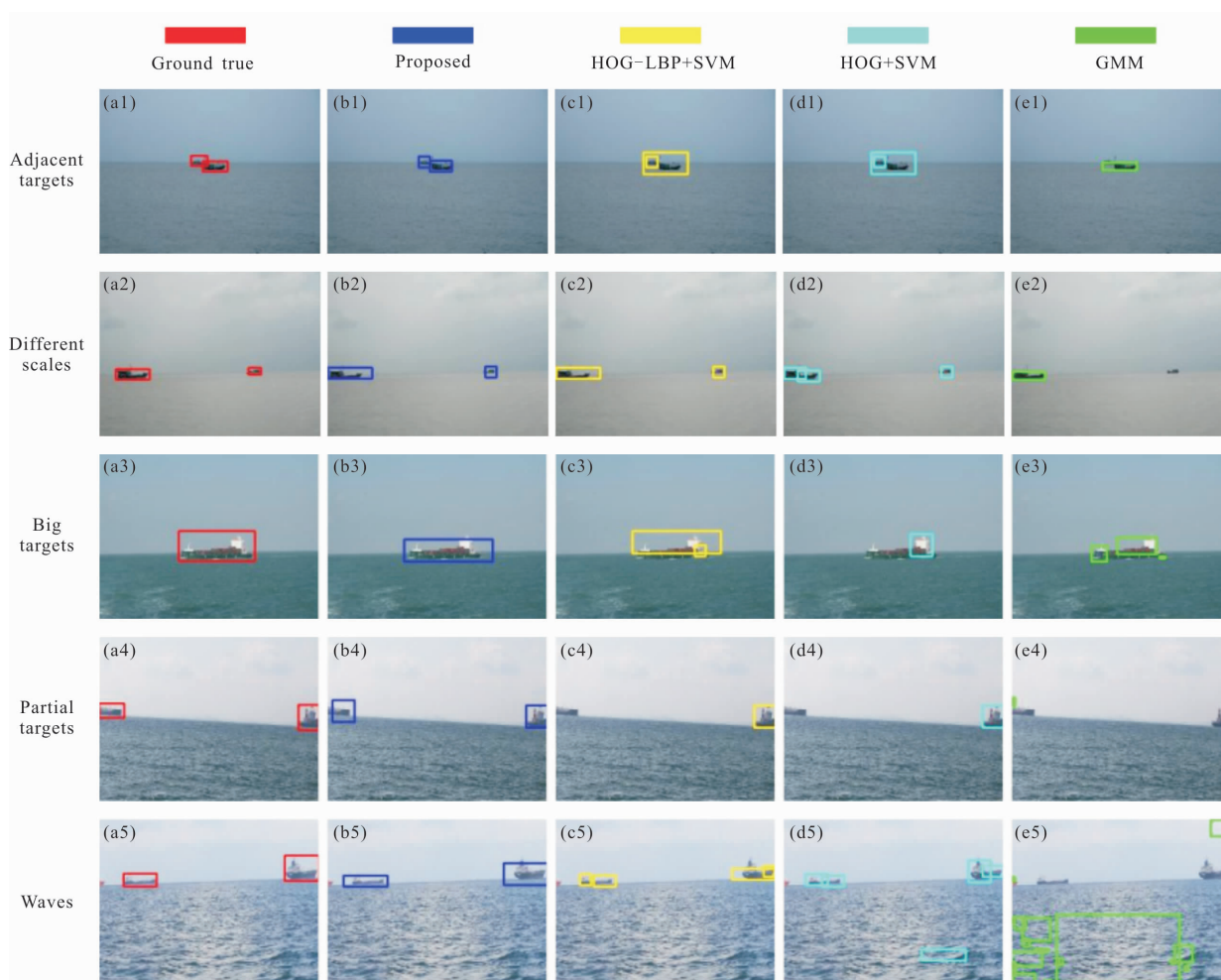


图 8 不同检测算法结果。不同颜色代表不同检测算法的结果,其中红色代表船舶标注信息,蓝色为文中算法结果

Fig.8 Results of different detection algorithms. Different colors denote the results of different detection algorithms, where red boxes indicate the ground true and blue boxes indicate the results of proposed algorithm

对于船舶目标检测,MLP 综合性能优于 SVM。

文中算法 AP 值最高,计算开销虽相对较大,但仍可以满足实时性的需求。从图 8(b5)可以看出,文

中算法抑制背景效果较好,也能产生包围效果更好的目标框。

从以上实验分析可得,文中算法在不同的场景

中(目标紧邻与局部、视角变化、海浪和尺度变化)比其他检测算法具有更好的效果,图 9 为不同算法在船舶检测中的 Precision-Recall 曲线图,显示了文中算法在综合多种场景中比其他算法具备更高的精确率和召回率。

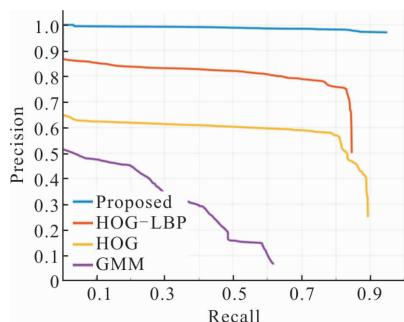


图 9 不同检测算法在船舶检测中的 Precision-Recall 曲线图
(不同颜色的曲线代表不同的检测算法)

Fig.9 Tradeoff between different detection algorithms in ship detection (Curves of different colors denote different detection algorithms)

4 结论

文中提出基于梯度纹理直方图与多层感知器的船舶快速检测方法,为了加快检测的速度,训练一个基于二值梯度特征的船舶候选框模型。实验从特征和分类器的角度出发,针对单一特征的不足和线性分类器区分能力的不足,融合梯度与纹理特征,形成能描述目标整体轮廓到局部结构的特征空间,结合多层感知器实现更高的船舶检测性能,并从实验中进行验证。

实验结果显示,文中算法比其他算法(HOG-LBP+SVM、HOG+SVM 和 GMM)具有更高的精确率和召回率,从而实现平均精确率达 90.0%。实时性虽然不及其他算法,但仍能满足实时性的要求,平均执行时间为 20.4 ms/frame。虽然文中算法能比其他算法实现更好的精确率,但是,由于数据集的局限,文中算法的通用性有待进一步验证。下阶段,仍需要扩大船舶数据集,研究具备更好泛化性能的船舶检测算法。

参考文献:

[1] Ding Peng, Zhang Ye, Jia Ping, et al. Ship detection on sea surface based on multi-feature and multi-scale visual

attention [J]. *Optical and Precision Engineering*, 2017, 45 (1): 167-172. (in Chinese)

丁鹏,张叶,贾平,等.基于多尺度多特征视觉显著性的海面舰船检测[J].*光学精密工程*,2017,45(1):167-172.

[2] Wang Huili, Zhu Ming, Lin Chunbo, et al. Ship detection of complex sea background in optical remote sensing images[J]. *Optical and Precision Engineering*, 2018, 26(3): 723-732. (in Chinese)

王慧利,朱明,蔺春波,等.光学遥感图像中复杂海背景下的舰船检测[J].*光学精密工程*,2018,26(3):723-732.

[3] Zhang Zhongyu, Jiao Shuhong. Infrared ship target detection method based on multiple feature fusion [J]. *Infrared and Laser Engineering*, 2015, 44(S): 29-34. (in Chinese)

张仲瑜,焦淑红.多特征融合的红外舰船目标检测方法[J].*红外与激光工程*,2015,45(S):29-34.

[4] Prasad D K, Prasath C K, Rajan D, et al. Object detection in a maritime environment: performance evaluation of background subtraction methods [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 20(5): 1787-1802.

[5] Zhang Yang, Li Qingzhong, Zang Fengni. Ship detection for visual maritime surveillance from non-stationary platforms [J]. *Ocean Engineering*, 2017, 141: 53-63.

[6] Li J, Tian Y, Huang T. Visual saliency with statistical priors [J]. *International Journal of Computer Vision*, 2014, 107 (3): 239-253.

[7] Wang X, Han T X, Yan S. An HOG-LBP human detector with partial occlusion handling [C]//2009 IEEE 12th International Conference on Computer Vision. IEEE, 2010: 32-39.

[8] Cheng Quan, Fan Yu, Liu Yuchun, et al. Moving target detection method based on block projection matching [J]. *Infrared and Laser Engineering*, 2018, 47 (10): 1026004. (in Chinese)

程全,樊宇,刘玉春,等.分块投影匹配的运动目标检测方法[J].*红外与激光工程*,2018,47(10):1026004.

[9] Lu Fuxing, Chen Xin, Chen Guilin, et al. Dim and small target detection based on background adaptive multi-feature fusion [J]. *Infrared and Laser Engineering*, 2019, 48 (3): 0326002. (in Chinese)

陆福星,陈忻,陈桂林,等.背景自适应的多特征融合的弱小目标检测[J].*红外与激光工程*,2019,48(3):0326002.

[10] Wu Tianshu, Zhang Zhijia, Liu Yunpeng, et al. A lightweight small object detection algorithm based on improved SSD[J]. *Infrared and Laser Engineering*, 2018, 47

- (7): 0703005. (in Chinese)
- 吴天舒, 张志佳, 刘云鹏, 等. 基于改进 SSD 的轻量化小目标检测算法[J]. 红外与激光工程, 2018, 47(7): 0703005.
- [11] Guliyev N J, Ismailov V E. On the approximation by single hidden layer feedforward neural networks with fixed weights [J]. *Neural Networks*, 2017, 98: 296–304.
- [12] Hosang J, Benenson R, Dollár P, et al. What makes for effective detection proposals? [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 38 (4): 814–830.
- [13] Cheng M M, Zhang Z, Lin W Y, et al. BING: Binarized normed gradients for objectness estimation at 300 fps [C]// *Computer Vision & Pattern Recognition*, 2014.
- [14] Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent [J]. *Journal of Statistical Software*, 2010, 33(1): 1.
- [15] Mdakane L, Bergh F V D. Extended local binary pattern features for improving settlement type classification of QuickBird images [J]. *Comptes Rendus Des Séances De La Société De Biologie Et De Ses Filiales*, 2012, 148(21–22): 1851–1852.
- [16] Vogl T P, Mangis J K, Rigler A K, et al. Accelerating the convergence of the back-propagation method [J]. *Biological Cybernetics*, 1988, 59(4–5): 257–263.
- [17] Uijlings J R R, van de Sande K E A. Selective search for object recognition [J]. *International Journal of Computer Vision*, 2013, 104(2): 154–171.
- [18] Alexe B, Deselaers T, Ferrari V. Measuring the objectness of image windows [J]. *IEEE Transactions on Software Engineering*, 2012, 34(11): 2189–2202.