



基于主干网络浅深层特征的无人机海上分割算法

沈昊, 葛泉波, 吴高峰

引用本文:

沈昊, 葛泉波, 吴高峰. 基于主干网络浅深层特征的无人机海上分割算法[J]. *智能系统学报*, 2025, 20(3): 605-620.
SHEN Hao, GE Quanbo, WU Gaofeng. Unmanned aerial vehicle-driven sea segmentation based on the shallow and deep features of the backbone[J]. *CAAI Transactions on Intelligent Systems*, 2025, 20(3): 605-620.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202405021>

您可能感兴趣的其他文章

动态云台摄像机无人机检测与跟踪算法

Drone detection and tracking in dynamic pan-tilt-zoom cameras

智能系统学报. 2021, 16(5): 858-869 <https://dx.doi.org/10.11992/tis.202103032>

基于反馈注意力机制和上下文融合的非模式实例分割

Feedback attention mechanism and context fusion based amodal instance segmentation

智能系统学报. 2021, 16(4): 801-810 <https://dx.doi.org/10.11992/tis.202007042>

基于生成对抗网络的机载遥感图像超分辨率重建

Super-resolution reconstruction of airborne remote sensing images based on the generative adversarial networks

智能系统学报. 2020, 15(1): 74-83 <https://dx.doi.org/10.11992/tis.202002002>

基于跳跃连接金字塔模型的小目标检测

Skip feature pyramid network with a global receptive field for small object detection

智能系统学报. 2019, 14(6): 1144-1151 <https://dx.doi.org/10.11992/tis.201905041>

高斯核函数卷积神经网络跟踪算法

Convolutional neural network tracking algorithm accelerated by Gaussian kernel function

智能系统学报. 2018, 13(3): 388-394 <https://dx.doi.org/10.11992/tis.201612040>

行人重识别研究综述

Survey on pedestrian re-identification research

智能系统学报. 2017, 12(6): 770-780 <https://dx.doi.org/10.11992/tis.201706084>

DOI: 10.11992/tis.202405021

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20250421.0849.002>

基于主干网络浅深层特征的无人机海上分割算法

沈昊¹, 葛泉波^{2,3,4}, 吴高峰²

(1. 南京信息工程大学计算机学院, 江苏南京 210044; 2. 南京信息工程大学自动化学院, 江苏南京 210044; 3. 大数据分析 with 智能系统江苏省高校重点实验室, 江苏南京 210044; 4. 大气环境与装备技术协同创新中心, 江苏南京 210044)

摘要: 为提高复杂海洋环境中无人机自主降落时分割目标的实时性和精确性, 研究主干网络和浅深层特征对分割算法性能的影响问题, 基于 DeepLabV3+ 框架建立一种基于主干网络浅深层特征的无人机海上分割 (shallow and deep features of backbone, SDFB) 算法。首先, 针对风浪扰动降低目标稳定性的问题, 优化 MobileNetV2 结构提出一种特征提取方法, 解决了算法无法处理短时间目标变化较大图像的问题; 然后, 针对深层特征输出通道数较多且存在不均匀分布大气湍流噪声的问题, 利用本地全局信息选择性地聚合特征, 提出一种特征筛选机制, 剔除冗余通道的同时解决了算法对环境噪声敏感度高问题; 其次, 针对光照不均降低目标边界清晰度问题, 从浅层空间维度和深层通道维度中提取轮廓信息建立一种并行轮廓学习机制, 解决了算法利用轮廓特征效率低的问题; 最后, 针对障碍物遮挡破坏目标特征完整性问题, 融合优化后的条带池化建立一种特征融合机制, 解决了算法无法联系离散分布特征问题。实验表明, SDFB 算法的实时性和精确性均高于其他算法, 能够更好地适应海上场景无人机分割目标需求。

关键词: 复杂海上场景; 语义分割; 无人机降落; 船舶目标; DeepLabV3+; 注意力机制; 深度学习; 卷积神经网络
中图分类号: TP751 **文献标志码:** A **文章编号:** 1673-4785(2025)03-0605-16

中文引用格式: 沈昊, 葛泉波, 吴高峰. 基于主干网络浅深层特征的无人机海上分割算法 [J]. 智能系统学报, 2025, 20(3): 605-620.

英文引用格式: SHEN Hao, GE Quanbo, WU Gaofeng. Unmanned aerial vehicle-driven sea segmentation based on the shallow and deep features of the backbone[J]. CAAI transactions on intelligent systems, 2025, 20(3): 605-620.

Unmanned aerial vehicle-driven sea segmentation based on the shallow and deep features of the backbone

SHEN Hao¹, GE Quanbo^{2,3,4}, WU Gaofeng²

(1. School of Computer, Nanjing University of Information Science and Technology, Nanjing 210044, China; 2. School of Automation, Nanjing University of Information Science and Technology, Nanjing 210044, China; 3. Jiangsu Provincial University Key Laboratory of Big Data Analysis and Intelligent Systems, Nanjing 210044, China; 4. Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology (CICAEET), Nanjing 210044, China)

Abstract: To improve the real-time and accurate segmentation of targets during the autonomous landing of UAVs in complex marine environments, studying the impact of backbone and shallow/deep features on the performance of algorithms is crucial. Based on the DeepLabV3+ framework, a Shallow and Deep Features of Backbone (SDFB) algorithm is established for maritime scene segmentation. First, to address the issue of reduced target stability caused by wind-wave disturbances, a feature extraction method is proposed by optimizing the MobileNetV2 structure, and this method resolves the issue of low processing speed of single frame images in the algorithm. Second, to address the issue of numerous deep feature output channels and the uneven distribution of atmospheric turbulence noise, a feature filtering mechanism is proposed by selectively aggregating features using local and global information, thereby eliminating redundant features while solving the high sensitivity issue of the algorithm to environmental noise. Third, to address the issue of uneven lighting reducing the clarity of target boundaries, a parallel contour learning mechanism is established by extracting contour information from shallow spatial dimensions and deep channel dimensions, thereby solving the low-efficiency issue regarding the utilization of contour features. Finally, to address the issue of background occlusion disrupting the integrity of target features, a multi-scale feature fusion mechanism is established through the fusion optimization of strip pooling, and this solves the connection issue of the algorithm to discrete distribution features. Finally, relevant experiments reveal that the LMSC algorithm exhibits higher real-time accuracy than other algorithms and can better adapt to the segmentation requirements of UAVs in maritime scenes.

Keywords: complex maritime scene; semantic segmentation; drone landing; ship target; DeepLabV3+; attention mechanism; deep learning; convolutional neural network

收稿日期: 2024-05-15. 网络出版日期: 2025-04-21.

基金项目: 国家自然科学基金项目 (62033010, 62303233); 江苏省高校青蓝工程项目 (R2023Q07).

通信作者: 葛泉波. E-mail: quanboge@163.com.

无人机技术在海上应用越发广泛^[1], 安全降落作为保证无人机海上作业^[2]完整性的必要条

件,得到人们的普遍关注。海上待降落无人机首先根据预先设定好的航线飞行到目标地点附近,随后通过搭载的相机识别出目标图像并利用算法将图像分割为不同的语义区域如海面、障碍物、平台等,以更精确地确定降落区域,最后基于语义分割^[3]结果,规划降落路径以规避障碍物,调整自身姿态和飞行速度降落到合适的位置。因此,海面目标场景分割是无人机降落的重要环节。

海上场景分割算法面临高精度、实时性的挑战。降落过程中需要算法精确且实时地识别出目标区域和障碍物以规划安全的降落路径^[4],无人机通过路径信息及时调整飞行状态保证降落的安全性,因此需要分割算法具有极高的精确性和实时性。然而,海上气候情况复杂多变,极端风浪气候易造成降落目标剧烈运动^[5],大气湍流噪声引起的涟漪效应也会导致图像中的空间扭曲模糊^[6],不同气候下光照对比度不同导致图像亮度动态范围变化^[7],这些海上气候因素导致算法精确性和实时性下降;此外,无人机搭载的相机及传感器感知能力和范围有限限制了自身的计算能力^[8],目标区域被飞鸟等障碍物遮挡破坏了目标特征的完整性^[9],这些无人机、降落目标因素导致算法精确性和实时性下降。

国内外针对无人机降落需求的图像分割^[10]仍处于发展阶段。分割算法是一种将图像中的像素划分成具有相似特征对象以实现对图像语义理解的方法。Chen 等^[11]基于位置、颜色和纹理的像素特征使用密集随机场和掩膜孔修复伪标签的缺陷区域,补充图像的粒度和可靠度,然而受海上风浪气候干扰,目标船舶进入摇摆运动状态易偏离期望区域^[12],该方法虽然提高了信息提取的精度,但是处理单帧图像特征速度较慢,无法应对分割目标短时间内变化较大的情况;康玥等^[13]对局部几何特征进行分组提取,并通过动态筛选和修剪冗余特征通道减少无用特征,然而受海洋大气湍流噪声影响,包含目标特征的图像退化严重^[14],该方法虽然在一定程度上消除了冗余特征的影响,但是无法筛除因大气湍流噪声干扰而模糊的特征;王潇棠等^[15]采用索伯算子计算各像素点不同方向的边缘梯度值,并结合双三次插值算法分割图像边缘细节,然而海上不同气候下光照对比度不同,目标图像明暗区别较大^[16],该方法基于环境不变假设,虽然对明亮图像边界分割效果较好,但是分割明暗变化较大图像时准确率不足;Han 等^[17]通过一种多尺度融合模块提取全局上下文信息进行对象搜索,并学习多级相似

性图之间的相互依赖关系,然而海上目标易被其他障碍物遮挡导致目标部分特征消失^[18-19],该方法虽然实现了多层相似图之间的交互融合,但是无法联系因障碍物遮挡而离散化的目标特征。

针对上述问题,本文基于 DeepLabV3+ 框架^[20]建立一种高精度高实时的基于主干网络浅深层特征的无人机海上分割算法 SDFB(shallow and deep features of backbone)。海上场景风浪扰动降低了目标稳定性,短时间内目标图像变化大,需要算法能够在极短的时间内提取初始特征从而提高算法的实时性,因此本文建立一种基于 MobileNetV2 的初始特征提取方法,对主干网络进行结构性优化,尽可能保留特征提取能力的同时降低运行参数,传统算法采用堆叠混合卷积的方式提取特征信息^[21],虽然不同的卷积核可以在不同尺度的感受野中提取更加丰富的特征,但是堆叠的方式计算复杂度较高,相比较本文方法处理单帧图像速度更快,能够实时输出变化目标的结果,算法抗风浪扰动能力得到增强。海上大气湍流噪声引起的光学畸变和涟漪效应降低了图像清晰度,大量干扰噪声不均匀分布在通道特征中,需要算法能够从众多通道特征中筛选出高质量特征,因此本文提出一种基于本地全局信息的特征筛选机制,构建通道权重学习模块分别从本地和全局的角度学习通道权重,传统方法采用对比边缘特征点数目与设定阈值筛选候选特征^[22],虽然计算方式比较简单,但是阈值的设定受场景影响较大,且这种方法极易受噪声干扰,相比较本文方法能够多角度学习通道重要性,筛选出有效特征,减少冗余通道的同时算法抗噪声干扰能力更强。海上场景受光照明暗以及光照角度的变化影响,目标阴影以及边缘与海平面颜色相近,需要算法能够精确学习目标轮廓,因此本文建立一种基于浅层空间和深层通道信息的轮廓学习机制,从浅层空间和深层通道两个维度中提取轮廓信息以精细化目标轮廓,传统方法利用超像素的种子点连接形成分割边界^[23],虽然具有较强的区域一致性,有助于减少过度分割现象,但是结果过度依赖初始图像超像素质量,对复杂场景适应性差,相比较本文方法能够充分利用浅深层特征空间和通道两个不同维度的轮廓信息,算法的光照适应性更强,能够更好地分割明暗变化较大的目标边界。海上障碍物遮挡破坏了目标特征的完整性,需要算法能够聚合离散分布的特征信息,因此本文建立一种融合优化条带池化的特征融合机制,优化传统条带池化结构,为离散分布的特征建立长期

依赖关系以聚合不同特征信息,传统方法采用微缩金字塔平衡特征金字塔融合特征^[24],虽然能够融合不同尺度特征,但是处理遮挡目标存在特征冲突问题,相比较本文方法能够联合分散的目标特征,抗障碍物遮挡能力更强。实验结果表明,该算法兼顾了图像分割实时性和精确性,更加适合无人机降落过程需要面对的复杂海上环境。

1 问题分析

1.1 问题描述

分割算法先用编码器将输入图像逐步压缩为特征图从而提取出图像特征,随后用解码器从特征图中恢复像素级别的语义信息从而输出与原始图像尺寸相同的分割结果^[25]。传统针对海上场景分割算法流程如图 1 所示。

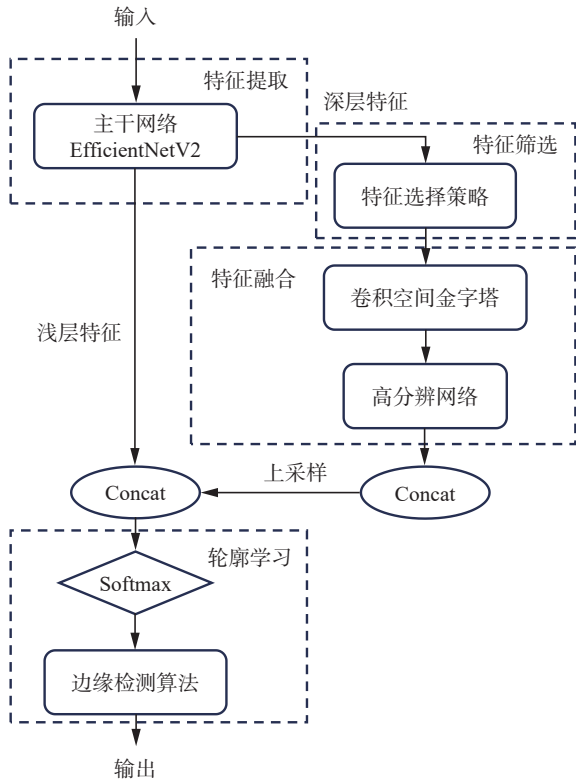


图 1 传统方案流程

Fig. 1 Traditional plan flowchart

特征提取模块采用轻量级 EfficientNetV2 提取初始语义信息,其渐进式学习方法可对正则化因子进行自适应调节^[26],缓解了极度正则化造成的模型欠拟合和过拟合的问题。

特征筛选模块引入稀疏组套索方法对通道特征进行选择,并用多层特征加权融合策略结合低层结构特征和高层语义特征^[27],提高了特征跟踪的精度,缓解了通道冗余和特征不相关性问题。

特征融合模块利用注意力机制增强了关键目标特征信息的权重,并将卷积空间金字塔的不同

尺度通道和高分辨网络相融合^[28],起到了保持图像高分辨率特征信息的效果。

轮廓学习模块将处理后的深层和浅层特征相加融合,经由 Softmax 函数激活后输出后与 Canny 算子相结合,建立一种边缘检测算法提取水线轮廓^[29],起到了分割水体和船体区域的效果。

1.2 研究动机

现有方案虽然在一定程度上缓解了复杂环境对分割算法实时性和精确性的影响,但在解决海上场景特有的风浪扰动、噪声干扰、光照不均匀以及障碍物遮挡问题时,仍存在以下不足。

主干网络庞大的结构降低了算法运行速度,导致分割算法无法实时获取变化中的目标位置。海水受到风力作用而产生的波动形成了海上场景特有的风浪^[30],风浪扰动导致降落目标剧烈运动,目标稳定性下降,由此导致短时间内图像变化较大。因此,需要算法能够在极短的时间内提取单帧图像目标特征。针对特征提取效率问题,图 1 的主干网络虽然可以缓解模型过拟合和欠拟合的问题,但是因为模型采用过多参数关注信息处理结构的有效性,导致算法处理单帧图像特征时间仍然过长,无法满足分割算法高实时性需求。

深层特征输出通道数较多且包含不均匀分布的大气湍流噪声,降低了特征利用率。近海面风与海面摩擦产生风速的垂直梯度^[31]引发大气湍流,而由此引起的光学畸变和涟漪效应导致目标图像空间扭曲和失真,这些模糊噪声不均匀分布在深层特征中,并且主干网络输出的通道数较多存在许多冗余通道信息。因此,需要算法能够剔除冗余通道并从中筛选出高质量特征。针对特征筛选问题,图 1 的通道选择策略虽然一定程度上缓解了不相关通道的影响,但是因为模型选择方法计算复杂度高,难以处理高维特征图像,且该模型对环境噪声敏感度高,难以筛除大气湍流噪声,无法满足分割算法高精度要求。

边缘检测算法不能充分利用浅层空间和深层通道信息,降低了算法分割模糊边界的能力。海上场景不同气候下光照明暗变化较大,且由于大气层密度不均匀阳光发生折射导致阳光透过大气空气不均匀^[32],图像较暗部分目标边界与海面颜色相近难以分割。因此,需要算法加强模糊边界处理能力。针对轮廓学习问题,图 1 的边缘检测算法虽然能够分割明显的水线轮廓,但是先输出结果再与边缘检测算法相结合的方式加大了算法运行负担,且忽视了浅层空间和深层通道维度的轮廓信息,特征信息利用率低,难以正确提取完

整目标区域轮廓,无法满足分割算法高精度要求。

拼接融合方法依赖特征之间的共变性,导致算法无法处理离散特征信息。海上场景存在飞鸟等障碍物遮挡目标区域^[33]造成目标特征完整性缺失问题,因此,需要算法能够为离散分布的特征建立长期依赖关系。针对离散特征融合问题,图 1 的高分辨网络虽然保持了图像高分辨率特征效果,但是因为模型依赖特征之间的相关性和共变性,对于离散特征存在信息丢失问题,难以关注距离较远的关联特征,大量离散有效特征被忽略,无法满足分割算法高精度要求。

改进后的算法有更强的抗风浪扰动能力,在

大气湍流噪声干扰方面有更好的特征突出能力且能更好地剔除冗余通道,在光照不匀方面有更好的轮廓表征能力,在障碍物遮挡方面有更好的离散特征处理能力,能够有效应对复杂海上场景对分割算法造成的挑战。对无人机自主降落海上目标的实现具有重大意义。

1.3 解决方案

为保障分割算法的鲁棒性,消除海上场景风浪扰动、大气湍流噪声干扰、光照不匀以及障碍物遮挡对分割结果带来的影响,本文建立基于主干网络浅深层特征的无人机海上分割算法。主要包括 4 部分内容,算法整体结构如图 2 所示。

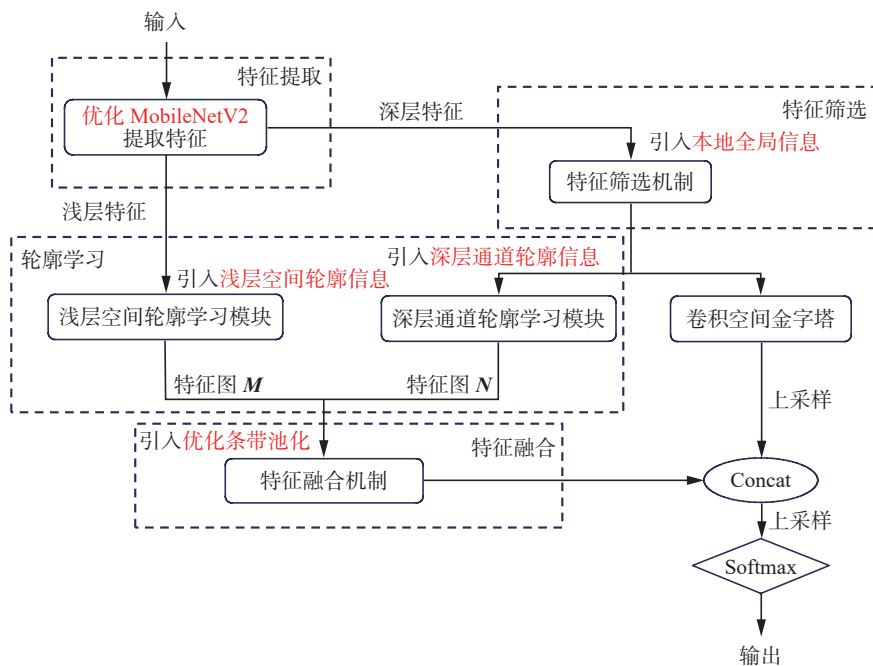


图 2 研究思路框架

Fig. 2 Research framework

1) 针对海上目标受风浪扰动稳定性降低问题,研究高实时提取特征方法,提出一种基于优化 MobileNetV2 初始特征提取。将裁剪掉最后一层,并优化其余各层参数以及结构的 MobileNetV2 模型替换传统的 Xception 模型,保留一定特征提取能力的同时极大降低了运算参数,使得算法能够在较短的时间内快速且精确地分割目标区域。

2) 针对主干网络输出通道数较多且包含不均匀分布大气湍流噪声问题,研究高精度筛选特征方法,提出一种基于本地全局信息的特征筛选机制。建立轻量化通道权重学习模块,分别从本地和全局角度学习通道重要性,剔除冗余通道,抑制大气湍流噪声,突出有效特征,从而筛选出高质量特征,提高了算法处理特征的利用率和速度。

3) 针对海上光照不匀导致图像较暗部分目标边界与海面颜色相近问题,研究高精度分割模糊边界方法,提出一种基于浅层空间和深层通道的并行轮廓学习模块。在不增加过多参数的同时,从空间和通道两个不同的维度中提取目标轮廓信息,并将此模块与金字塔模块并行放置构成并行特征处理结构,从梯度信息中学习轮廓,以提取精细物体边界,从而解决边界模糊问题。

4) 针对海上目标受障碍物遮挡特征完整性被破坏问题,研究高精度联系离散特征方法,提出一种融合优化条带池化的特征融合机制。可反向学习参数的优化条带池化拥有更强的特征表现能力,融合该池化结构的特征融合机制,能够为不同图像中离散分布的特征之间建立长期依赖关系,从而解决障碍物遮挡问题。

2 基于优化 MobileNetV2 的特征提取

传统 DeepLabV3+模型采用 Xception 作为主干网络, 然而 Xception 的结构巨大, 引入大量参数对每次计算效果进行详尽对比, 主要目的为审视结构的有效性^[34], 力求更全面地评估计算结

果。因此, 这种模型无法处理短时间变化较大的目标, 难以满足无人机降落任务所需的实时性。Xception 结构如表 1 所示。为解决该问题, 本节构建了基于优化 MobileNetV2 的初始特征提取方法, 尽可能保留初始特征提取能力的同时, 减少训练参数, 以此满足无人机分割算法所需的实时性。

表 1 原始主干网络 Xception 结构
Table 1 Original backbone network Xception structure

层号	输入	操作	重复次数	卷积步长	输出
1	512×512×3	二维卷积	1	2	256×256×32
2	256×256×32	二维卷积	1	1	256×256×64
3	256×256×64	深度可分离卷积	2	2	128×128×128
4	128×128×128	深度可分离卷积	2	2	64×64×256
5	64×64×256	深度可分离卷积	2	2	32×32×728
6	32×32×728	深度可分离卷积	24	1	32×32×728
7	32×32×728	二维卷积	2	1	16×16×1024
8	16×16×1024	二维卷积	2	1	16×16×2048

MobileNet 系列网络主要采用深度可分离卷积^[35], MobileNetV2^[36] 在此基础之上引进反向残差块模块^[37]。该模块由升维丰富特征信息的 1×1Conv、跨像素点提取特征的 3×3Conv、降维减少运算参数的 1×1Conv 3 部分构成, 通过堆叠该模块与残差结构进行混合使用。这种方法会极大程度减少运算参数, 模型的主要目的变为压缩参数和提速。优化后的 MobileNetV2 模型舍去最后

一层只采用 4 次下采样, 以减少参数量达到提高运行速度的目的; 同时修改第 2、4、7、14 块的卷积步长, 图像尺寸以相同尺度缩小且通道以相同倍数增加, 以避免损失过多特征影响模型预测的精确性。将优化后的主干网络应用在无人机海上场景分割算法中, 尽可能保留图像特征信息的同时, 降低了 10 倍左右的系统开销。优化后的 MobileNetV2 结构如表 2 所示。

表 2 优化后的 MobileNetV2 结构
Table 2 Optimized MobileNetV2 structure

层号	输入	操作	重复次数	卷积步长	T	输出
1	512×512×3	二维卷积	1	2	否	256×256×3
2	256×256×3	反向残差块	1	1	是	256×256×6
3	256×256×6	反向残差块	2	2	是	128×128×10
4	128×128×10	反向残差块	4	2	是	64×64×20
5	64×64×20	反向残差块	4	2	是	32×32×40
6	32×32×40	反向残差块	3	1	是	32×32×80
7	32×32×80	二维卷积	3	1	是	16×16×160
8	16×16×160	二维卷积	1	1	是	16×16×320

注: T表示模块是否进行 1×1 卷积。

综上所述, 可以得出结论:

1) 离散化的 Xception 结构参数数量过大, 运算过程繁琐, 提取图像特征的效率颇低, 而优化后的 MobileNetV2 参数数量较少, 结构紧凑, 运算过程集中, 能够有效提高模型收敛速率;

2) Xception 需要进行 36 层计算, 最大输出通道数为 2048, 而 MobileNetV2 仅需要进行 18 层计

算, 最大输出通道数也仅为 320, 极大地缩减了计算量, 提高了模型分割速度。

因此, 本节提出的优化 MobileNetV2 作为主干网络能够尽可能保证初始特征提取效率的同时, 极大减少模型运算参数, 提高了模型运算速度, 使得算法能够分割短时间变化较大的目标, 满足无人机海上场景降落所需的高实时性要求。

3 基于本地全局信息的特征筛选机制

随着主干网络对图像操作次数的增加,感受野逐渐增加且感受野之间的重叠区域也在不断增加,此时像素点代表的信息是一个区域的信息,虽然语义信息更加丰富,但是受海上大气湍流噪声干扰,模糊特征也增加了。此外,经过优化的主干网络深层特征输出通道数虽然大大减少,但仍存在较多冗余通道。传统 DeepLabV3+模型忽略了众多通道之间的联系以及深层特征中的噪声特征,未加处理地直接传入后续模块。一方面,模糊特征的干扰降低了算法分割海上目标的精确性;另一方面,冗余的通道信息造成了资源的严重浪费,降低了算法分割目标的实时性。因此,为了解决上述问题,提高算法处理深层特征的精确性与实时性,本节提出了基于本地全局信息的特征筛选机制。通过 3×3 感受野获取目标及其周围各一个邻居的特征信息作为目标的本地上下文特征以利用更多的细粒度特征,保证本地信息能够捕获更多细节,从而从本地角度帮助机制抵抗大气湍流噪声的干扰;通过全局平均池化生成对应于每个通道的汇聚特征值以得到图像的全局表示,保留图像的整体特征和空间信息,建立特征之间的远距离依赖,协助更好地描述本地语义信息,从而从全局角度帮助机制抵抗大气湍流噪声的干扰;通过权重学习模型学习通道重要性,从而剔除冗余通道的影响。特征筛选能够强化必要特征信息,抑制非必要特征信息,从而筛选出高质量特征。

3.1 通道权重学习模型

通道权重学习模型利用点卷积和参数化修正线性单元 (parametric rectified linear unit, PReLU) 相结合的方式学习并保存通道权重系数:

$$W(M) = PWC_2(\gamma(LN(PWC_1(M)))) \quad (1)$$

式中: M 为输入图像; PWC_1 为 1×1 点卷积, 作用为对输入特征进行降维操作, 放缩至输入尺寸的 $1/r$ 倍, r 为设定参数; LN 为 Layer Normalization 层, 作用为防止网络过拟合; γ 为 PReLU 激活函数; PWC_2 为 1×1 点卷积, 作用为对特征进行升维操作, 恢复至输入时尺寸; W 为模型输出结果。

文献 [38] 采用点卷积、ReLU 和批量标准化 (batch normalization, BN) 的组合进行像素权重学习。而本文利用点卷积首先对每个空间位置的逐点通道进行交互, 聚合通道上下文信息, 使用较小的参数量学习保存通道权重; 利用 PReLU 在正域部分取原值, 负域部分取原值乘以可学习参数

的方法代替文献 [38] 的 ReLU, 可以更加突出通道信息之间的差异, 便于模块学习; 利用训练时不需要使用 mini-batch 统计信息的 LN 代替文献 [38] 的 BN, 使得通道权重学习模块对于本文使用的小批量样本仍保持较好的归一化效果。改进后的通道权重学习模块结构如图 3 所示。

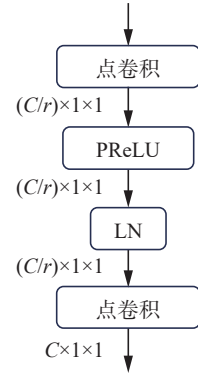


图 3 通道权重学习模块
Fig. 3 Channel weight learning module

3.2 特征筛选机制

特征筛选机制利用本地和全局信息训练通道权重, 以寻求更强大的特征表现力, 筛选出高质量通道特征, 提高图像分割的精准性。

特征筛选机制的具体流程如图 4 所示。先采用全局池化的方法从全局信息中训练通道权重:

$$Glo(M) = W(GP(M)) \quad (2)$$

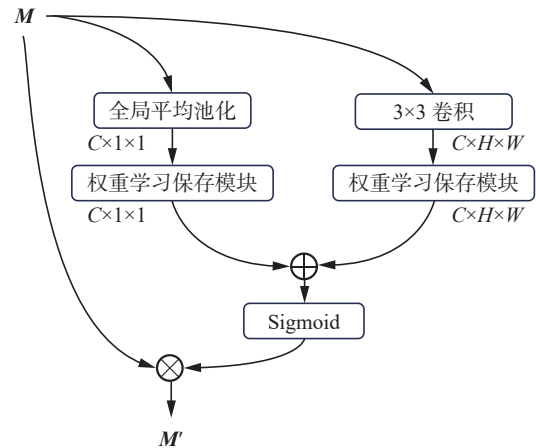


图 4 特征筛选机制
Fig. 4 Feature screening mechanism

然后从本地信息关注图像细节训练通道权重:

$$Loc(M) = W((Conv(M))) \quad (3)$$

全局特征信息和局部特征信息相加后经由 Sigmoid 函数激活, 并与原始输入图像进行逐元素点乘, 得到最终通道加权后的结果:

$$M' = M \otimes \sigma(Glo(M) \oplus Loc(M)) \quad (4)$$

式中: M 为输入图像; GP 为全局平均池化; W 为通道权重学习模块; $Conv$ 为 3×3 卷积, 将输入图

像每个像素作为目标, 对于周围缺失邻居的目标采取补 0 法将其缺失的特征信息补全; Glo 为全局信息; Loc 为本地信息; \oplus 为逐元素广播相加; σ 为 Sigmoid 激活函数; \otimes 为逐元素点乘。

相比较原始算法未加处理地直接使用所有深层特征信息, 本文所构建的基于本地全局信息的特征筛选机制, 从局部和全局多角度学习通道特征的重要性, 强有力地关注了其中有效特征, 抵抗了大气湍流噪声的干扰且剔除了冗余通道的影响, 提高了分割算法的精确性与实时性。点卷积和 PReLU 激活函数的使用, 使得整个特征筛选机制轻量化, 满足分割算法所需的实时性。

4 并行轮廓学习模块

分割海上场景目标时, 受光照不匀的影响, 图像较暗部分目标边界与海面颜色相近, 易产生边界模糊问题。因此, 算法的轮廓学习能力需要得到提高, 兼顾分割精确性和鲁棒性, 以此保证无人机安全降落于目标区域。然而, 传统 DeepLabV3+ 模型仅采用空洞卷积空间金字塔处理深层语义特征, 忽略了包含轮廓信息浅层空间和深层通道, 特征信息捕捉不充分, 无法准确分割模糊边界。为此本文利用文献 [39] 对位置以及通道特征信息的处理建立一种基于浅层空间和深层通道的轻量化并行轮廓学习模块。主干网络浅层输出感受野较小, 而且此时特征图每个像素点对应的感受野重叠区域也较小, 故浅层特征富含较高特征分辨率的空间信息, 因此提取空间轮廓信息可以增加模块从浅层空间维度分割模糊边界的精确性; 主干网络深层输出感受野较大且感受野重叠区域也较大, 故深层特征富含较低特征分辨率的语义信息, 而某些深层通道映射的高级特征可以视为轮廓类特定的响应, 且经过上一节特征筛选机制筛选后已剔除大气湍流噪声和冗余通道的影响, 因此提取通道轮廓信息可以增强模块从深层通道维度表征图像细节和边缘的能力。并行轮廓学习模块在尽可能减少参数的同时, 提取浅层空间和深层通道的轮廓信息, 关注重要信息。

将轻量级轮廓学习模块与空洞卷积空间金字塔并行放置构成并行特征处理结构, 使得模型保留空洞卷积空间金字塔多个比例捕获图像特征能力的同时, 选择性地聚合每个通道的特征以及再分配卷积信道之间的资源, 极大地增加了模型轮廓特征表达能力, 提高了模型处理边界模糊图像的能力。并行特征处理结构如图 5 所示。

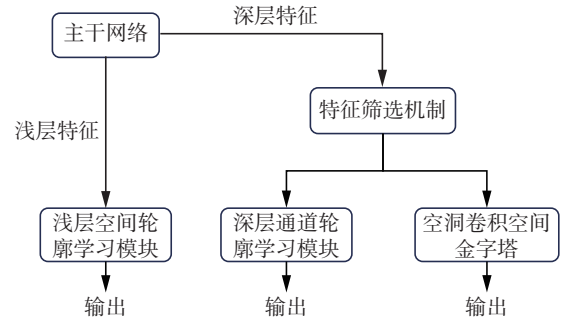


图 5 并行特征处理结构

Fig. 5 Parallel feature processing architecture

4.1 并行特征处理结构

主干网络处理输出的浅层特征传入空间轮廓学习模块, 输出的深层特征经由特征筛选机制通道加权后, 分别传入通道轮廓学习模块和空洞卷积空间金字塔模块, 3 个特征处理模块并行放置组成并行特征处理结构, 并行结构如图 5 所示。

浅层空间轮廓学习模块计算公式为

$$b_{nm} = \frac{\exp(s_n \cdot Q_m)}{\sum_{m=1}^N \exp(s_n \cdot Q_m)} \quad (5)$$

$$p_n = \alpha \sum_{m=1}^N (b_{nm} \cdot C_m) + M_s \quad (6)$$

式中: M_s 表征浅层主干输入空间轮廓学习模块的单帧特征图; s_n 为浅层特征图 M_s 通过 1×1 卷积降维之后, 传入多尺寸池化金字塔后经过第 1 层池化后得到的特征图; Q_m 为特征图 M_s 经过第 1 层池化后与后 3 层堆叠并全连接后得到的特征图; b_{nm} 表征特征图中第 m 个中心和第 n 个像素之间的关系; C_m 为前述所有特征经由全连接后得到的特征图; α 为设置变量值; p_n 为浅层空间轮廓学习模块最终输出结果。

浅层空间轮廓学习模块通过关联每个像素值与周围像素值之间的关系, 构建出每个像素点与部分聚合中心之间的联系, 梯度中学习轮廓, 从浅层空间角度补充轮廓空间细节信息。并且 1×1 卷积和多尺寸池化金字塔的使用, 有效减少了参数量, 提高了模型的运算速度。

深层通道轮廓学习模块计算公式为

$$X_{nm} = \frac{\exp(M_d \cdot E_m)}{\sum_{m=1}^c \exp(M_d \cdot E_m)} \quad (7)$$

$$F_n = \beta \sum_{m=1}^c (X_{nm} \cdot E_m) + M_d \quad (8)$$

式中: M_d 表征深层主干输入通道轮廓学习模块的单帧特征图; E_m 由深层特征图 M_d 经过 1×1 卷积降维后得到, E_m 的每个通道都被视为一个聚集中

心; \mathbf{X}_{nm} 表征特征图中第 m 个中心和第 n 个通道之间的关系; β 为设置变量值; \mathbf{F}_n 为深层通道轮廓学习模块最终输出结果。

深层通道轮廓学习模块中每个卷积核对应的通道之间存在或大或小的依赖, 利用原始特征图中任意两个通道映射之间的依赖关系, 学习加强有效依赖, 从深层通道角度补充轮廓信息。

综上所述, 本节提出的结构通过并行放置浅层空间轮廓学习模块、深层通道轮廓学习模块和空间金字塔模块, 保留以多通道的方式将经过学习的不同层级的特征图与解码器特征融合能力的同时, 有效地从浅层空间和深层通道信息中学习轮廓, 适应海上场景目标区域边界模糊的状况, 极小的运算量也能保证模型分割目标区域的实时性和精确性, 有效保障了无人机的安全降落。

5 融合优化条带的特征融合机制

海上目标极易被飞鸟等障碍物遮挡出现大量关键特征信息被离散化的情况, 导致算法错误地将目标分割成多个区域或将障碍物和目标当成整体。第 4 节的并行轮廓学习模块会分别输出两种不同的特征图, 对于这两种不同尺寸的特征图, 采用加法、逐点相乘以及串联等传统融合方法, 会忽略上下文之间的语义关联, 其简单的结构很难捕获关键信息, 更无法统筹距离较远的特征信息, 导致分割算法的精确性无法得到保障。为了解决上述问题, 本节优化传统条带池化^[40]方法, 并融合该池化方法建立一种特征融合机制, 为离散分布的特征信息之间建立长期依赖关系。

5.1 优化条带池化结构

池化一般作用于卷积之后, 对于输入进来的特征图像进行最大池化或者平均池化等操作, 降维压缩图像的同时保留主要特征, 减少参数和计算量。池化本质就是采样, 选择有效特征, 过滤无用信息, 以便加快运算速度。而全局池化就是池化滑窗的大小和整幅输入特征图的大小相同, 输入 $W \times H \times C$ 的特征图, 得到 $1 \times 1 \times C$ 的输出, 也分为全局最大池化和全局平均池化等。全局池化已在多种任务尤其是逐像素预测任务中, 被证明捕获远程上下文语义信息效果显著。然而方形的内核形状利用上下文语义信息的能力是有限的, 其局限性使之难以直接用于特征学习的网络构建块。由于海上场景船只长程带状外形的特殊性以及背景干扰, 如果只是简单的放大方形内核, 会不可避免地受到来自不相关区域的噪声影响。

传统条带池化分别使用两种不同形状的池化

结构输出并融合^[40], 本文在传统结构上进行改进。传统条带池化水平和垂直输出结果融合后仅采用一次 1×1 卷积和 Sigmoid 激活函数的组合, 虽减小了一部分运算参数量, 但丢失了许多重要特征信息, 且 Sigmoid 激活函数计算复杂, 输出值在 $0 \sim 1$, 容易产生梯度消失以及梯度爆炸的同时, 不利于辨别特征有效性。本文对融合扩张后的特征图先进行点卷积降维, 减少运算量, 降低过拟合的风险, 然后经由 PReLU 函数进行非线性激活, 突出了有效特征信息, 减弱了无关特征信息的影响, 最后通过点卷积恢复特征图原始尺寸, 扩展特征图的深度。整个过程的参数被保存且可被反向训练, 使得该结构可以学习到更复杂的特征表示, 提高该结构的表达能力。

假设给定输入目标二维张量 $\mathbf{M} \in \mathbf{R}^{H \times W}$, 经由水平条带池化输出 $\mathbf{q}^x \in \mathbf{R}^H$ 为

$$\mathbf{q}_i^x = \frac{1}{W} \sum_{0 \leq j < W} \mathbf{M}_{i,j} \quad (9)$$

经由垂直条带池化输出 $\mathbf{q}^y \in \mathbf{R}^H$ 为

$$\mathbf{q}_j^y = \frac{1}{H} \sum_{0 \leq i < H} \mathbf{M}_{i,j} \quad (10)$$

将扩张后的特征图叠加在一起:

$$\mathbf{q}_{n,i,j} = \mathbf{M}_{n,i}^x + \mathbf{M}_{n,j}^y \quad (11)$$

$$\mathbf{Q} \in \mathbf{R}^{C \times H \times W} \quad (12)$$

经由 1×1 点卷积降维, 放缩至输入尺寸的 $1/r$ 倍后, 利用 PReLU 函数进行激活, 再由 1×1 点卷积升维恢复至原始尺寸。激活后与原始输入图像逐层点积进行融合后输出结果:

$$\mathbf{T} = \mathbf{M} \otimes (\text{PWC}_2(\gamma(\text{PWC}_1(\mathbf{q})))) \quad (13)$$

式中: \mathbf{M} 为原始输入图像; W 为设置水平池化长度; \mathbf{q}^x 为单个水平条带池化后的输出; H 为设置垂直条带池化高度; \mathbf{q}^y 为单个垂直条带池化后的输出; \mathbf{Q} 为扩张后叠加的特征图; PWC_1 为 1×1 点卷积降维; γ 为 PReLU 激活函数; PWC_2 为 1×1 点卷积升维; \otimes 为矩阵逐元素乘法; \mathbf{T} 为最终输出。

优化条带池化的建立, 其长而窄的窗口能够有效联系背景干扰下离散的远程细节特征信息, 改进后的融合扩张特征图的方法, 能够为离散分布的特征之间建立长期依赖关系, 相比较直接考虑整个特征图的方形内核, 该方法避免了连接大多数相距很远且无效的特征, 最大程度的缓解不相关区域的噪声扰动, 有效提高了分割算法的精确性; 该方法也因此避免了大量无效的计算, 使得整个过程更加的轻量化, 有效满足了无人机降落任务所需的精确性和实时性。

5.2 特征融合机制

文献 [38] 仅采用点卷积和 ReLU 的组合处理本地特征, 无法融合因障碍物遮挡而特征离散的不同尺寸特征图, 因此本文提出融合优化条带池化的方式处理本地特征, 建立一种特征融合机制。

设浅层空间轮廓学习模块输出的特征图为 M , 深层通道轮廓学习模块输出的特征图为 N , 首先将两种特征图相加融合:

$$V = M \oplus N \tag{14}$$

然后对融合特征图取全局池化并经由权重学习模块关注特征图全局信息:

$$Glo(V) = W(GP(V)) \tag{15}$$

再对融合特征图做优化条带池化操作, 然后经由权重学习模块关注特征图局部信息:

$$Loc(V) = W(T(V)) \tag{16}$$

全局信息与局部信息相加融合, 并由 Sigmoid 函数进行激活后分别与原始特征图加权融合:

$$z = Loc(V) + Glo(V) \tag{17}$$

$$Z = (1 - \sigma(z)) \otimes M + \sigma(z) \otimes N \tag{18}$$

式中: GP 为全局平均池化; W 为通道权重学习模块; Glo 为全局信息输出; T 为优化条带池化; Loc 为局部信息输出; \oplus 为不同特征图 M 、 N 进行逐元素广播相加; σ 为 sigmoid 激活函数; Z 为特征融合机制最终输出结果。由于 $\sigma(z)$ 输出真实值位于 0 到 1 之间, 为了特征融合机制能够更加平滑地训练选择相应的权重, 采用 $1 - \sigma(z)$ 的方法对输入特征图做加权平均。整个多尺度特征融合机制的实现过程如图 6 所示。

综上所述, 传统 DeepLabV3+ 算法采用例如加法、逐点相乘以及串联等作为融合方法, 忽略了上下文语义特征联系, 损失了大量的有效特征。

海上场景极易出现障碍物遮挡目标问题, 传统融合方法因无法处理离散化的特征易将目标分割成多个区域或将障碍物和目标当成整体, 降低了分割算法的精确性。融合优化条带池化的特征融合机制, 既关注全局远程特征, 也关注局部细节特征, 特殊长呈带状的优化条带池化在减少运算量的同时, 能够联系不同特征图中离散分布的特征, 可反向训练参数的结构使得其在复杂特征中仍有良好的表现力, 提高了算法的实时性与精确性。

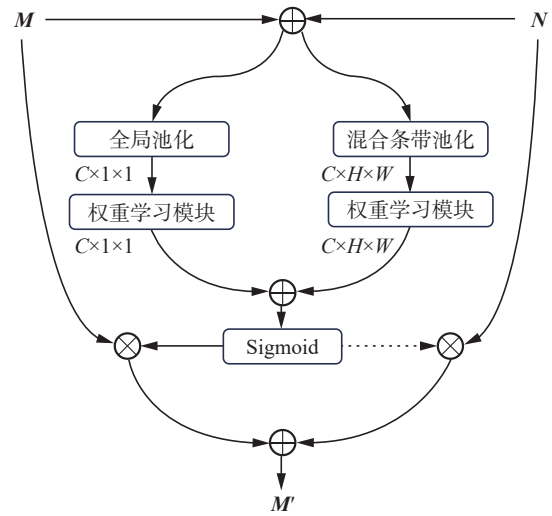


图 6 多尺度特征融合机制
Fig. 6 Multi-scale feature fusion

6 简要分析

本文针对原始 DeepLabV3+ 算法分割海上目标的不足, 为解决海上场景风浪扰动、噪声干扰、光照不匀、背景遮挡的问题, 提出一种基于主干网络浅深层特征的无人机海上分割算法, 算法主要包括 4 个新工作, 整体算法流程如图 7 所示。

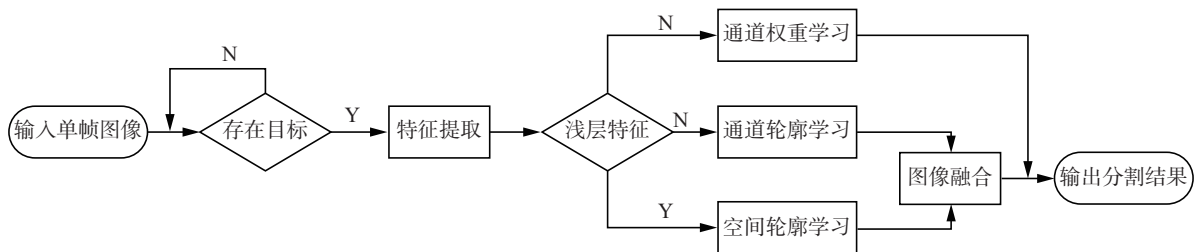


图 7 整体算法流程
Fig. 7 Overall algorithm flowchart

6.1 算法小结

本文所提出算法首先利用 MobileNetV2 结构的轻量化, 对其整体结构进行优化以减少运算时间。现有主干网络 Xception 主要目的在于验证结构的合理性, 需要大量的计算资源进行训练和推

理。但是本文优化后的主干网络结构, 在保留一定特征处理能力的同时, 极大减少了运算量。

然后基于本地全局信息建立一种特征筛选机制。该方法是对文献 [38] 中的特征加权机制进行改进, 通过点卷积与 PReLU 激活函数以及 LN 的

组合使用, 构建出通道权重学习模块, 用于学习各个通道的重要性权重, 然后分别从全局信息 $Glo(\mathbf{M})$ 以及局部信息 $Loc(\mathbf{M})$ 中训练通道权重, 以此来调整通道之间的关系, 从而在众多特征中筛选出高质量特征, 解决了大气湍流噪声不均匀堆积和通道冗余影响算法精确性与实时性的问题。

其次基于浅层空间和深层通道维度建立一种轮廓学习模块。该模块通过改进文献 [39] 中的位置和通道注意模块, 从浅层特征的空间信息和深层特征的通道信息中学习目标轮廓, 并将其与空间金字塔并行放置, 以此加强原始算法的边界处理能力, 保留模型多个比例捕获图像特征的能力的同时, 有效增强了算法的模糊边界处理能力。

最后融合优化条带池化建立一种特征融合机制。该机制是对文献 [40] 的传统条带池化进行改进, 利用点卷积和 PReLU 组合后可反向训练参数的结构, 增强了条带池化突出有效特征信息的能力。多尺度特征融合机制对不同尺度的图像分别从全局信息和本地信息角度突出有效特征信息。本地信息利用优化条带池化为离散分布的特征之间建立长期依赖关系, 整体机制能够有效联系不同特征图之间的上下文语义联系。

6.2 创新性分析

第 1 个工作是提出了一种基于优化 MobileNetV2 的初始特征提取方法。该方法是对文献 [37] 中反向残差块的重复次数和卷积步长进行修改, 减少原始模型层数的同时修改各层输入输出特征尺寸, 将优化后的 MobileNetV2 模型替换原始主干网络 Xception 进行初始特征提取, 在尽可能保留图像特征信息的同时, 降低了 10 倍左右的开销, 有效提高了分割算法的实时性。

第 2 个工作利用文中式 (1) 的通道权重学习模块训练学习各个通道的重要性权重。现有的通道加权方法 SE 全模块采用压缩激励机制, 其激励阶段通过全连接层来学习权重, 全连接层的参数量取决于输入特征图的通道数量, 而深层特征输入通道数为 320, 若使用 SE 模块会引入大量的参数, 因此文中式 (1) 仅使用 2 个点卷积和 1 个 PReLU 激活函数以及 1 个 LN 进行组合, 仅用较少的参数即可学习每个通道的权重, 极大减少了运算量。现有的通道加权方法仅从全局角度考虑权重, 忽视了局部信息对通道权重的影响, 因此本文利用文献文中式 (2) ~ (3) 分别从全局信息和本地信息角度训练通道权重, 文中式 (4) 将学习到的权重值与原始输入图像进行逐元素点乘, 得到最终通道加权结果, 高质量特征得到突出, 以

此达到特征筛选的效果。

第 3 个工作是利用文献 [39] 对位置和通道特征信息的处理, 通过式 (5) ~ (6) 和 (7) ~ (8) 分别从浅层空间信息和深层通道信息中提取轮廓特征提高算法的模糊边界处理能力。现有轮廓学习模块 BASNet 仅从局部范围内处理边界信息, 忽略了整体场景的全局上下文信息以及通道信息, 但是本文所使用的式 (5) ~ (8), 不仅在浅层空间维度关注了局部与全局上下文中的轮廓信息, 也保留了深层通道维度中的轮廓信息。

第 4 个工作是利用式 (9) ~ (13) 优化传统条带池化, 并与式 (14) ~ (18) 相结合改进传统图像融合方案。现有跳跃结构融合方法中, 低层与高层的特征传递不及时, 影响特征融合的效果, 因此利用公式 (9) ~ (12) 从水平与垂直方向池化特征图并叠加组合输出新的特征图, 然后由式 (13) 为离散分布的远程细节特征信息建立连接, 经由此操作的式 (16) 能够获得更加全面的局部特征信息, 结合式 (15) 的全局信息, 由式 (17) ~ (18) 输出最终特征图融合结果。

7 仿真实验

7.1 算法平台

本次实验的硬件平台采用 CPU 为 12th Gen Intel Core i7, 内置 GPU 为 NVIDIA GeForce RTX 4070 Ti, 软件平台选择以 PyCharm2022、Python3.6 在 PyTorch 框架上进行训练。

7.2 数据集构造

本文主要以各类典型的目标为对象, 其中包括军舰、海船、货船、轮船等可降落无人机目标。由于本文设定的海面场景的特殊性, 不同气候条件下的目标情况直接影响算法的训练结果, 因此, 收集图像时充分考虑了目标的光照、位置、距离、角度以及气候等不同因素。降落目标数据集如图 8 所示。

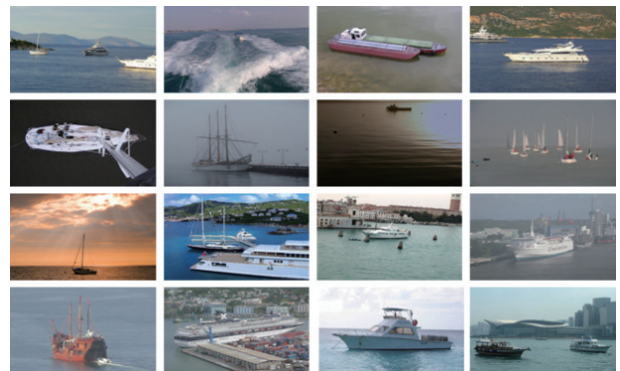


图 8 降落目标数据集
Fig. 8 Drop the target dataset

主要的数据来源包括 4 部分: 南信大自主无人系统研究团队现有设备采集的多场景水上实船图像、由英国牛津大学计算机视觉小组创建并在 PASCAL VOC 挑战赛中使用的通用标准数据集 Vocdevkit 中含有船舶的图像^[41]、由微软团队提供的可以用于图像分割的通用标准数据集 COCO 中含有船舶的图像^[42] 以及网络上公开船舶图像, 通过 labelme 进行标定。数据集图像总数为 8280 张, 其中 6624 张作为训练集, 1656 张作为测试集。

7.3 实验设计

实验设计分为以下几部分:

1) 在本节提出的实验环境下训练模型, 观察其在训练数据集以及测试训练集上的表现;

2) 构建两个消融模块“全局+全局”和“局部+局部”, 验证通道权重学习模块采用“本地+全局”策略聚合上下文的有效性。

3) 对比本文的通道权重学习模块与 SE 通道加权机制的性能, 验证特征筛选机制消除冗余通道和大气湍流噪声的高效性。

4) 可视化分割结果, 对比改进前后算法与前沿算法分割无人机降落过程获取的远近及光暗区别目标图像的效果, 验证轮廓学习模块分割因光照不匀影响而模糊的边界的有效性。

5) 在相同的实验环境, 将 SDFB 整体算法与算法改进前后以及近年来提出的前沿语义分割模型进行性能对比, 验证 SDFB 算法分割无人机降落过程机载相机获取目标图像的实时性和精确性。

7.4 性能指标

实验采用平均交并比 F_{MIoU} (mean intersection over union, MIoU)、类别平均像素准确率 F_{MPA} (mean pixel accuracy, MPA)、像素准确率 F_{PA} (pixel accuracy, PA)、时延 L (latency) 作为性能指标。

1) 平均交并比 (MIoU)

单类的交并比就是指该类的真实标签和预测结果的交和并的比值, 平均交并比 F_{MIoU} (MIoU) 就是指数据集中每一个单类交并比的平均。其值越大, 说明分割精度越高。 F_{MIoU} 的计算公式为

$$F_{MIoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{S_{TP}}{S_{FN} + S_{FP} + S_{TP}} \quad (19)$$

式中: S_{TP} 是真正例, S_{FP} 是假假例, S_{FN} 是假反例。用 P_{ij} 表示将 i 类别预测为 j 类别, 替换变量后的公等价于

$$F_{MIoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ij}}{\sum_{i=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}} \quad (20)$$

2) 类别平均像素准确率 (MPA)

MPA 首先对图像每个像素点类别进行准确率计算, 然后再对所有像素点类别准确率求均值。MPA 的数值越大, 表示模型对图像的分割能力越好。 F_{MPA} 的计算公式为

$$F_{MPA} = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij}} \quad (21)$$

3) 像素准确率 (PA)

PA 是指像素预测正确数在总像素数中占有比例。 F_{PA} 的计算公式为

$$F_{PA} = \frac{\sum_{i=0}^k P_{ii}}{\sum_{i=0}^k \sum_{j=0}^k P_{ij}} \quad (22)$$

4) 时延 (L)

在设定的实验软硬件配置下, 模型对于传入图像进行分割所需平均时长。

7.5 模型训练

SDFB 在本文数据集上进行训练, 其在训练数据集以及测试数据集上的损失函数如图 9 所示。横轴为算法的运行次数, 纵轴为训练损失值。

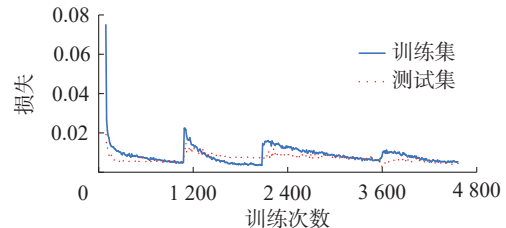


图 9 损失函数变化情况
Fig. 9 Changes in loss function

算法使用原始 Vocdevkit 数据集预训练的 MobileNetV2 模型进行初始化权重。初期设置了相对较大学习率即 0.01 实现模型训练初期损失函数快速下降, 其间出现损失函数梯度消失、爆炸和震荡等问题, 经过多次测试, 最后选取较小的学习率即 0.007 进行训练。前 1000 轮训练冻结了主干网络, 仅训练后续结构参数, 第 800 轮损失逐渐平稳收敛; 在第 1000 轮解冻主干网络, 训练整体模型所有参数, 出现第一次损失激增, 仍采用学习率为 0.007, 第 3000 轮损失逐渐平稳收敛; 在第 3500 轮缩小学习率到 0.00007 左右, 第 4000 轮整体损失基本收敛。

7.6 本地全局策略聚合上下文有效性研究

为了验证“本地+全局”方式聚合上下文信息的有效性, 本节构建“全局+全局”和“本地+本地”两个消融模块, 如图 10 所示。

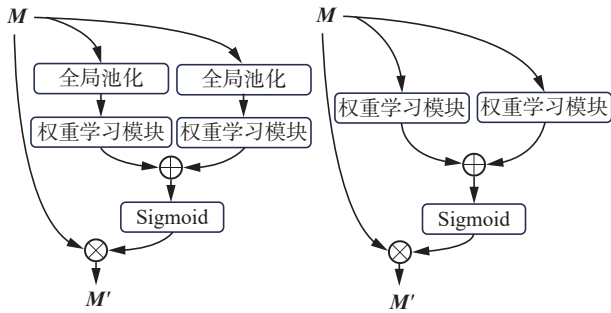


图 10 消融模块
Fig. 10 Ablation module

设定在同样实验环境下训练相同次数，采用相同的学习率，均冻结主干网络，唯一区别是上下文聚合方式，分别在 Vocdevkit 和 COCO 数据集上进行实验，如图 11 所示，横轴为训练

迭代次数，图 11(a)、(b) 纵轴为 MPA 评价指标值，图 11(c)、(d) 纵轴为 MIoU 评价指标值。根据图 11 消融模块分别在 Vocdevkit 和 COCO 数据集上的表现可以发现，在两种不同的数据集上，“本地+全局”的 MIoU 和 MPA 性能参数在整个过程中均优于另外两种结构，验证了“本地+全局”结构的有效性以及稳定性。并且由图 11(a) 可以发现，“本地+全局”策略的 MPA 指标在训练 2000 次基本可以达到“本地+本地”策略需要训练 3000 次才能达到的效果，而且远超“全局+全局”策略训练 3000 次的效果，更加验证了“本地+全局”策略的有效性。因此，本文“本地+全局”策略更加适用于聚合上下文特征信息。

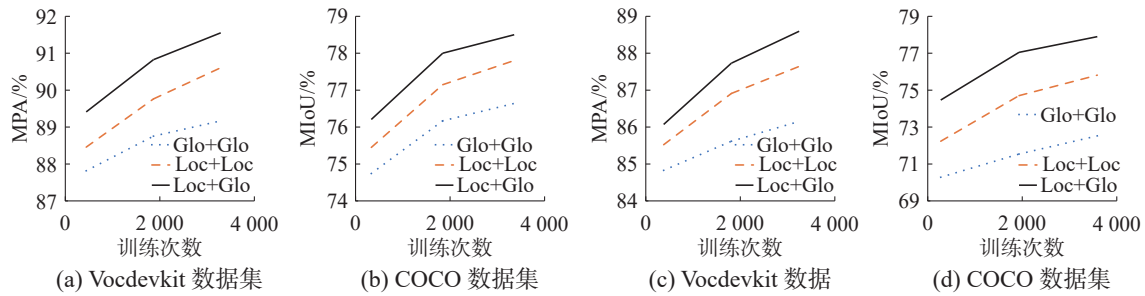


图 11 消融模块在不同数据集表现

Fig. 11 Performance of different datasets of the ablation module

7.7 本地全局特征筛选机制有效性研究

为了验证基于本地全局信息的特征筛选机制消除冗余通道和大气湍流噪声的有效性，与现有 SE 通道加权模块对比性能指标与参数大小。对比实验均采用本文构造的无人机降落海上场景相机获取含有噪声干扰的船舶数据集。

由表 3、4 可知，同样的实验环境和本文创建的数据集下，本文所建立的基于本地全局的特征筛选机制在 MIoU 性能指标上高出 SE 模块 2.1%，在 MPA 性能指标上高出 SE 模块 2.3%，但所占取内存仅为 SE 模块的 56%。本节实验结果表明，基于本地全局信息的特征筛选机制能够有效筛选出无人机降落海上场景机载相机获取目标图像中的高质量特征，并且在一定程度上节约了资源开销。

表 3 性能指标对比
Table 3 Comparison of performance indicators %

模型	MIoU	MPA
SE	76.93	88.59
特征筛选	78.54	90.62

表 4 参数大小对比

Table 4 Comparison of parameter sizes MB

模型	储存空间
SE	579
特征筛选	325

7.8 可视化分割验证轮廓学习模块

在相同的实验环境下，消融轮廓学习模块，从可视化效果来验证轮廓学习模块的有效性。图像分别为自主无人系统研究团队实船以及百度公开船舶图像，最终分割效果如图 12 所示。图 12 第 1 列为输入原始图像，其中第 1 个和第 5 个为无人机海上降落过程中获取图像目标较近的场景，与第 3 个无人机海上降落过程获取图像目标较远场景形成对比，第 2 个为无人机获取图像目标较明亮场景，与第 4 个无人机获取图像目标较昏暗场景形成对比；第 2 列为原始 DeepLabV3+算法分割结果图；第 3 列为添加了浅层空间轮廓学习模块的 DeepLabV3+算法分割结果图；第 4 列为本文添加了浅层空间和深层通道轮廓学习模块的 SDFB 算法；第 5 列为 HRNet 算法分割结果

图; 第 6 列为 U²Net 算法分割结果图; 第 7 列为 SETR 算法分割结果图; 第 8 列为 SegFormer 算法

分割结果图; 第 9 列为 Swin-UNet 算法分割结果图。



图 12 分割结果可视化对比

Fig. 12 Visual comparison of segmentation results

对比分割结果可发现, 仅添加浅层空间轮廓学习模块对模糊边界处理能力就有一定程度的提升, 同时添加浅层空间和深层通道轮廓学习模块的 SDFB 分割质量更是比原始 DeepLabV3+有明显的提高, 且 SDFB 的海上模糊边界分割效果优于其他算法, 抗海上光照不均的能力更强。

7.9 本文算法与前沿算法的定量对比

将 SDFB 与改进前模型, 以及近 5 年前沿算法 HRNet、U²Net、SETR、SegFormer、Swin-Net 等在目标区域分割的精准性和实时性方面进行定量对比。在本文所构建海上无人机降落目标数据集上, 不同算法得到的 MIoU、MPA、PA 以及时延结果如表 5 所示, 各模型储存空间如表 6 所示。

表 5 模型性能对比

Table 5 Model performance comparison

模型	MIoU/ %	MPA/ %	PA/ %	L/ms
Xception+ASPP	80.25	92.73	95.59	400
MobileNetV2+ASPP	76.78	87.06	92.79	210
SDFB	81.73	93.96	96.68	230
HRNet	78.20	88.28	90.38	393
U ² Net	73.52	83.94	86.69	250
SETR	77.12	85.30	87.71	518
SegFormer	80.30	90.72	93.27	1 548
Swin-UNet	77.25	86.59	90.18	5 634

在对比实验中, 输入图片经裁剪后大小统一为 512×512, SDFB、Xception+ASPP 以及 MobileNet+ASPP 的 MIoU 分别排在第 1、3、7 名。采用 MobileNetV2 作为主干网络比采用 Xception 作为主干网络处理速度提升了 47.5%, 但处理精度会下降很多, SDFB 处理速度虽然相较于主干网络

为 MobileNetV2 下降 8.7%, 但相比较于主干网络为 Xception 提升 42.5%, 且整体性能有所提升, 平均交并比提高 1.8%, 类别平均像素准确率提高 1.3%, 相比较其他前沿分割算法, SDFB 算法性能也是最优的。SDFB 算法优化了主干网络参数, 在不增加过多模型参数的同时, 保证了无人机分割海上目标的实时性需求, 并对其输出的浅层特征和深层特征进行充分利用, 提高了模型分割的精确性, 总体看来, 改进后的算法更加满足无人机海上降落需求。

表 6 模型储存空间对比

Table 6 Comparison of operating parameters MB

模型	储存空间
Xception+ASPP	4389
MobileNetV2+ASPP	1250
SDFB	1661

8 结束语

为提高无人机分割算法在复杂海洋环境降落过程中的实时性和精确性, 本文基于 DeepLabV3+ 算法建立一种更适应海面场景的基于主干网络浅深层特征的分割算法 SDFB。首先针对风浪扰动问题提出基于优化 MobileNetV2 的特征提取方法, 其次针对噪声干扰以及通道冗余问题提出一种基于本地全局信息的特征筛选机制, 然后针对光照不均问题提出一种基于浅层空间和深层通道的轮廓学习模块, 最后针对障碍物遮挡问题提出一种融合优化条带池化的特征融合机制。SDFB 算法处理单张图像仅需 230 ms, 相比较原始模型提高了 42.5%, SDFB 算法分割目标的平均交并比精度高达 81.73%, 相比较原始模型提高了 1.8%。

SDFB 算法充分利用了主干网络及其输出的浅层特征和深层特征,兼容了运行实时性与分割精度,可以满足无人机降落海上场景目标分割的所需的精确性与实时性要求。由于无人机相机对海洋环境敏感度高,因此接下来计划将本文思想应用于更加极端的海洋环境中,进一步研究海洋气候对目标稳定性以及图像空间清晰度的影响,解决极端气候下目标特征的模糊和失真等问题。

参考文献:

- [1] LIU Yalin, DAI Hongning, WANG Q, et al. Unmanned aerial vehicle for Internet of everything: opportunities and challenges[J]. *Computer communications*, 2020, 155: 66–83.
- [2] RAOULT V, COLEFAX A P, ALLAN B M, et al. Operational protocols for the use of drones in marine animal research[J]. *Drones*, 2020, 4(4): 64.
- [3] 饶颖露, 邢金昊, 张恒, 等. 基于视觉的无人机板载自主实时精确着陆系统[J]. *计算机工程*, 2021, 47(10): 290–297.
RAO Yinglu, XING Jinhao, ZHANG Heng, et al. Vision-based autonomous real-time precise landing system for UAV-borne processors[J]. *Computer engineering*, 2021, 47(10): 290–297.
- [4] 原辉, 裴楚, 王帅, 等. 基于地标的轻量化无人机精准降落算法[J]. *计算机工程与应用*, 2022, 58(17): 306–313.
YUAN Hui, PEI Chu, WANG Shuai, et al. Landmark-based lightweight UAV accurate landing algorithm[J]. *Computer engineering and applications*, 2022, 58(17): 306–313.
- [5] 金建海, 周则兴, 张波, 等. 无人艇航行仿真关键技术研究[J]. *系统仿真学报*, 2021, 33(12): 2846–2853.
JIN Jianhai, ZHOU Zexing, ZHANG Bo, et al. Research on USV navigation simulation key technologies[J]. *Journal of system simulation*, 2021, 33(12): 2846–2853.
- [6] MAO Zhiyuan, CHIMITT N, CHAN S H. Accelerating atmospheric turbulence simulation via learned phase-to-space transform[C]//2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 14739–14748.
- [7] KUBRYAKOV A A, KUDRYAVTSEV V N, STANICHNY S V. Application of Landsat imagery for the investigation of wave breaking[J]. *Remote sensing of environment*, 2021, 253: 112144.
- [8] GHAZALI M H M, RAHIMAN W. An investigation of the reliability of different types of sensors in the real-time vibration-based anomaly inspection in drone[J]. *Sensors*, 2022, 22(16): 6015.
- [9] 龙翔, 陈华杰, 吴浩宇, 等. 基于特征增强的海面强干扰目标检测[J]. *激光与光电子学进展*, 2025, 62(2): 1–20.
Long Xiang, Chen Huajie, Wu Haoyu, et al. Strong Interference Target Detection on the Sea Surface Based on Feature Augmentation[J]. *Laser & optoelectronics progress*, 2025, 62(2): 1–20.
- [10] 李伟, 杨敏. 基于轻量级改进网络的无人机航拍图像语义分割研究[J]. *微电子学与计算机*, 2024, 41(7): 29–36.
LI Wei, YANG Min. Research on semantic segmentation of aerial images based on improved lightweight networks[J]. *Microelectronics & computer*, 2024, 41(7): 29–36.
- [11] CHEN Guangchen, SHI Benjie, ZHANG Yinhuai, et al. CGSNet: Cross-consistency guiding semi-supervised semantic segmentation network for remote sensing of plateau lake[J]. *Journal of network and computer applications*, 2024, 230: 103974.
- [12] 董微微. 考虑横向风浪干扰的舰船航向稳定性自动控制算法[J]. *舰船科学技术*, 2024, 46(9): 143–146.
DONG Weiwei. Automatic control algorithm for ship heading stability considering lateral wind and wave interference[J]. *Ship science and technology*, 2024, 46(9): 143–146.
- [13] 康玥, 杨军. 可学习动态分组卷积神经网络的大规模点云分割[J]. *计算机工程与应用*, 2024, 60(10): 217–226.
KANG Yue, YANG Jun. Large-scale point cloud segmentation by learnable dynamic grouping convolutional neural network[J]. *Computer engineering and applications*, 2024, 60(10): 217–226.
- [14] 苗壮, 崔浩然, 张启阳, 等. 基于对比学习的大气湍流退化图像复原方法[J]. *计算机科学*, 2025, 52(5): 171–178.
MIAO Zhuang, CUI Haoran, ZHANG Qiyang, et al. Restoration of atmospheric turbulence-degraded images based on contrastive learning[J]. *Computer science*, 2025, 52(5): 171–178.
- [15] 王潇棠, 闫河, 刘建骐, 等. 一种边缘梯度插值的双分支 deepLabv3+语义分割模型[J]. *智能系统学报*, 2023, 18(3): 604–612.
WANG Xiaotang, YAN He, LIU Jianqi, et al. A new deepLabv3+ semantic segmentation model of edge gradient interpolation with double branch structure[J]. *CAAI transactions on intelligent systems*, 2023, 18(3): 604–612.
- [16] 华怡坦, 黄影平, 过文昊. 基于 CNN 和 Transformer 点

- 云图像融合的道路检测[J]. 应用科学学报, 2024, 42(4): 695–708.
- HUA Yitan, HUANG Yingping, GUO Wenhao. Fusion of point-cloud and image for road segmentation using CNN and Transformer[J]. Journal of applied sciences, 2024, 42(4): 695–708.
- [17] HAN Shuo, GAO Lisha, WU Yue, et al. SiamSMN: Siamese cross-modality fusion network for object tracking[J]. *Information*, 2024, 15(7): 418.
- [18] CHEN Chen, LI Ying, WANG Tianqi. Real-time tracking and berthing aid system with occlusion handling based on LiDAR[J]. *Ocean engineering*, 2023, 288: 115929.
- [19] ŽUST L, KRISTAN M. Learning with weak annotations for robust maritime obstacle detection[J]. *Sensors*, 2022, 22(23): 9139.
- [20] CHEN L C, ZHU Yukun, PAPANDEOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]// Computer Vision–ECV 2018. Cham: Springer International Publishing, 2018: 833–851.
- [21] JÖRGES C, BERKENBRINK C, GOTTSCHALK H, et al. Spatial ocean wave height prediction with CNN mixed-data deep neural networks using random field simulated bathymetry[J]. *Ocean engineering*, 2023, 271: 113699.
- [22] 张远, 赵建国. 基于边缘特征点筛选的被遮挡盲道检测算法[J]. 科学技术与工程, 2021, 21(34): 14657–14664.
- ZHANG Yuan, ZHAO Jianguo. Algorithm for occluded blind track detection based on edge feature points screening[J]. Science technology and engineering, 2021, 21(34): 14657–14664.
- [23] SHI Qingwu, YIN Shoulin, WANG Kun, et al. Multichannel convolutional neural network-based fuzzy active contour model for medical image segmentation[J]. *Evolving systems*, 2022, 13(4): 535–549.
- [24] 邵凯, 王明政, 王光宇. 基于 Transformer 的多尺度遥感语义分割网络[J]. 智能系统学报, 2024, 19(4): 920–929.
- SHAO Kai, WANG Mingzheng, WANG Guangyu. Transformer-based multiscale remote sensing semantic segmentation network[J]. CAAI transactions on intelligent systems, 2024, 19(4): 920–929.
- [25] 王卓, 瞿绍军. 基于注意力机制与多尺度池化的实时语义分割网络[J]. 计算机工程, 2023, 49(10): 222–229, 238.
- WANG Zhuo, QU Shaojun. Real-time semantic segmentation network based on attention mechanism and multi-scale pooling[J]. Computer engineering, 2023, 49(10): 222–229, 238.
- [26] 赵为平, 陈雨, 项松, 等. 基于改进的 DeepLabv3+图像语义分割算法研究[J]. 系统仿真学报, 2023, 35(11): 2333–2344.
- ZHAO Weiping, CHEN Yu, XIANG Song, et al. Image semantic segmentation algorithm based on improved DeepLabv3+[J]. Journal of system simulation, 2023, 35(11): 2333–2344.
- [27] 刘佳榆, 余华, 徐舒. 基于通道特征选择的加权响应目标跟踪[J]. 计算机仿真, 2022, 39(12): 253–258.
- LIU Jiayu, YU Hua, XU Shu. Weighted response object tracking based on channel feature selection[J]. *Computer simulation*, 2022, 39(12): 253–258.
- [28] 方豪, 伍鹏, 谢凯, 等. 基于注意力机制的道路环境语义分割算法[J]. 计算机仿真, 2023, 40(3): 122–128.
- FANG Hao, WU Peng, XIE Kai, et al. Semantic segmentation algorithm of road environment based on attention mechanism[J]. Computer simulation, 2023, 40(3): 122–128.
- [29] 吕悦, 张永忠. 基于图像语义分割的船舶水线检测[J]. 计算机仿真, 2022, 39(5): 216–219, 233.
- LYU Yue, ZHANG Yongzhong. Ship waterline detection based on image semantic segmentation[J]. Computer simulation, 2022, 39(5): 216–219, 233.
- [30] 周阳, 叶钦, 杨斌, 等. 海湾台风浪特性及其对船舶的影响[J]. 中国造船, 2020, 61(3): 216–227.
- ZHOU Yang, YE Qin, YANG Bin, et al. Wave characteristics in the bay during typhoons and its effect on mooring ships[J]. *Shipbuilding of China*, 2020, 61(3): 216–227.
- [31] 柯熙政, 李鑫. 激光束光斑漂移反演大气湍流各向异性实验研究[J]. 光电工程, 2024, 51(4): 42–53.
- KE Xizheng, LI Xin. Experimental study on inversion of atmospheric turbulence anisotropy by laser beam spot wander[J]. Opto-electronic engineering, 2024, 51(4): 42–53.
- [32] STEPHENS G L, HAKUBA M Z, KATO S, et al. The changing nature of Earth's reflected sunlight[J]. *Proceedings of the Royal Society A: mathematical, physical and engineering sciences*, 2022, 478(2263): 20220053.
- [33] TERŠEK M, ŽUST L, KRISTAN M. eWaSR—an embedded-compute-ready maritime obstacle detection network[J]. *Sensors (basel)*, 2023, 23(12): 5386.
- [34] LO W W, YANG Xu, WANG Yapeng. An xception con-

- volutional neural network for malware classification with transfer learning[C]//2019 10th IFIP International Conference on New Technologies, Mobility and Security. CANARY ISLANDS: IEEE, 2019: 1–5.
- [35] HOWARD A G, ZHU Menglong, CHEN Bo, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2024-05-15]. <https://arxiv.org/abs/1704.04861v1>.
- [36] SANDLER M, HOWARD A, ZHU Menglong, et al. Inverted residuals and linear bottlenecks: Mobile networks for classification, detection and segmentation[EB/OL]// (2018-04-02)[2024-05-05]. <https://arxiv.org/abs/1801.04381>.
- [37] SANDLER M, HOWARD A, ZHU Menglong, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4510–4520.
- [38] DAI Yimian, GIESEKE F, OEHMCKE S, et al. Attentional feature fusion[C]//2021 IEEE Winter Conference on Applications of Computer Vision. Waikoloa: IEEE, 2021: 3560–3569.
- [39] FU Jun, LIU Jing, JIANG Jie, et al. Scene segmentation with dual relation-aware attention network[J]. *IEEE transactions on neural networks and learning systems*, 2021, 32(6): 2547–2560.
- [40] HOU Qibin, ZHANG Li, CHENG Mingming, et al. Strip pooling: rethinking spatial pooling for scene parsing[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 4003–4012.
- [41] EVERINGHAM M, ESLAMI S M A, VAN GOOL L, et al. The pascal visual object classes challenge: a retrospective[J]. *International journal of computer vision*, 2015, 111: 98–136.
- [42] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft coco: Common objects in context[C]//Computer Vision–ECCV 2014: 13th European Conference. Zurich: Springer International Publishing, 2014: 740–755.

作者简介:



沈昊, 硕士研究生, 主要研究方向为计算机视觉。E-mail: hshen_nuist@163.com。



葛泉波, 教授, 中国自动化学会青年工作委员会副主任委员、中国人工智能学会自主无人系统专业委员会副主任委员、中国自动化学会人工智能与机器人教育专业委员会副秘书长、中国自动化学会混合智能专业委员会副秘书长等。江苏省“青蓝工程”中青年学术带头人、浙江省万人计划“青年拔尖”入选者、中国自动化学会第四届青年科学家奖、浙江省杰出青年科学基金获得者等。主要研究方向为状态估计与信息融合、自主智能无人系统、飞行器测试数据分析和电力 IOT 技术。主持国家级科研项目 12 项, 省部级科研项目 7 项, 其余类别科研项目 10 项。发表学术论文 92 篇, 制定专利/标准 76 项。E-mail: quanboge@163.com。



吴高峰, 讲师, 主要研究方向为航空火力控制、无人机任务规划与控制、机器学习及其在无人系统中应用。E-mail: wgf@nuist.edu.cn。