

DOI:10.3876/j.issn.1000-1980.2026.01.020

基于轻量化 SuperPoint 网络的水下光学图像特征提取

刘 艳,朱昌盛,余 彬,霍冠英

(河海大学信息科学与工程学院)

摘要: 针对水下光学图像质量下降导致的图像配准、三维重建等水下视觉任务中特征提取鲁棒性差的问题,提出了一种轻量化 SuperPoint 网络,该网络针对水下光学图像普遍存在颜色失真、模糊等细节退化问题,利用注意力机制,构建频域-空间域动态注意力融合模块,融合频域与空间域的特征信息,提升网络在水下退化图像中的特征提取能力;构建残差特征增强深度可分离卷积模块,以降低模型复杂度并增强网络的特征提取能力。验证结果表明:该网络较 SuperPoint 网络参数量减少了 13.8%,计算量降低了 8.0%,帧率提升 31.7%,光照变化和视角变化下的重复率分别提高了 2.3%和 2.1%,在 SQUID 和 FLSea 数据集上的特征点检测与匹配性能评估中具有较好的特征提取鲁棒性。

关键词: 水下光学图像;SuperPoint 网络;轻量化网络;特征提取;特征融合

Underwater optical image feature extraction based on lightweight SuperPoint network

Liu Yan, Zhu Changsheng, Yu Bin, Huo Guanying

(College of Information Science and Engineering, Hohai University)

Abstract: In view of the poor robustness of feature extraction in underwater vision tasks such as image registration and 3D reconstruction caused by the decline in underwater optical image quality, an lightweight SuperPoint network was proposed. This network addressed the common challenges of detail degradation in underwater optical images, including color distortion and blurring. By leveraging an attention mechanism, it constructed a frequency-spatial dynamic attention fusion module that integrated feature information from both the frequency and spatial domains, thereby enhancing the network's capability for feature extraction in underwater degraded images. A residual feature enhancement depthwise separable convolutional module was constructed to reduce model complexity and enhance the feature extraction ability of the network. Verification results demonstrate that, compared with the SuperPoint network, the network proposed in this paper achieves a 13.8% reduction in the number of parameters, an 8.0% decrease in computational complexity, and a 31.7% improvement in frame rate. Meanwhile, its repeatability rates under illumination variation and viewpoint variation are increased by 2.3% and 2.1%, respectively. In addition, the network exhibits excellent robustness in feature extraction in the performance evaluation of feature point detection and matching on the SQUID and FLSea datasets.

Key words: underwater optical image; SuperPoint network; lightweight network; feature extraction; feature fusion

图像处理过程中,特征提取作为图像配准、运动跟踪、三维重建等视觉任务的基础,通常包含两个核心环节:特征点检测与特征描述符生成^[1]。特征点检测旨在定位图像中具有显著局部结构信息的关键点,这些点可作为后续匹配与分析的基础^[2];特征描述符旨在对关键点邻域进行独特数学表征,为特征点提供可区分、可匹配的描述符^[3]。水下环境会对图像成像质量造成显著影响:①光在水中的强烈衰减导致红光被严重吸收而蓝光吸收较少,使图像普遍呈现蓝绿色调;②水体散射效应降低了图像对比度;③水中的有机物质和悬浮颗粒等杂质进一步造成图像模糊^[4]。这些退化效应严重削弱了特征提取的鲁棒性,即不仅要求精确检测关键点,更需生成稳定且高区分度的描述符,因此如何提升模糊退化水下图像特征提取的鲁棒性,已成为亟待解决的关键问题。

为提高图像特征提取的鲁棒性,学者们提出了多种方法,主要分为传统图像特征提取方法和基于深度学习的图像特征提取方法。传统图像特征提取方法主要通过人工设计算法实现检测和描述图像中的关键局部特征^[5],如 Harris 等^[6]提出了一种特征点检测方法,通过计算图像局部窗口内的梯度信息检测角点,但其未

生成特征描述符;FAST(features from accelerated segment test)角点检测方法^[7]与 Harris 等^[6]的方法类似,只提取特征点而不生成描述符,通过检测每个像素周围的邻域区域是否满足一定的亮度变化判断特征点;Lowe^[8]提出的尺度不变特征变换(scale invariant feature transform,SIFT)解决了仅提取特征点的问题,检测出尺度不变的特征点后,还生成了描述符用于图像配准。加速稳健特征(speeded-up robust features,SURF)算法^[9]作为 SIFT 的改进,通过加速 Hessian 矩阵的计算,改善了 SIFT 计算耗时的问题。ORB^[10](oriented FAST and rotated BRIEF)利用 FAST 进行特征点检测,BRIEF^[11](binary robust independent elementary features)作为特征点的描述符,以提高特征提取效率。这些传统方法依赖于人工设计的规则,无法有效适应复杂和多变的环境,尤其在处理水下图像时,由于光照衰减、颜色失真、模糊等问题,表现常常受到较大限制^[12]。

随着深度学习技术的发展,卷积神经网络(CNN)在图像特征提取中得到了广泛应用,CNN 通过自动学习图像中的特征,能够在复杂和多变的环境中表现出更强的适应能力^[13]。近年来,研究者提出了多种基于 CNN 的特征提取网络,如 Dusmanu 等^[14]提出的 D2-Net 网络实现了特征点检测与描述符生成的端到端联合学习;Revaud 等^[15]通过改进 D2-Net 网络,引入鲁棒性与可重复性损失,通过自监督学习提升特征点检测的稳定性;De Tone 等^[16]提出 SuperPoint 网络,基于自监督训练框架,利用合成图像生成初始标签,以实现无监督的特征点检测与描述符提取,在光照变化、视角变化等复杂条件下表现出优异的稳定性和鲁棒性;Zhao 等^[17-18]提出 ALIKED 方法,引入了稀疏可变形描述符头(sparse deformable descriptor head,SDDH),能够自适应地捕捉关键点周围局部几何形变信息,从而提升匹配精度和跨模态场景的适应性;Yao 等^[19]提出的 EdgePoint2 网络则面向边缘计算场景,通过轻量化设计提高在嵌入式系统上的部署效率。上述方法虽然在不同场景下均取得了良好表现,但在水下图像处理方面仍面临诸多挑战:D2-Net 网络对纹理依赖性较强,在水下图像中容易遗漏低纹理区域的特征点;R2D2 网络的特征稠密性相对较低,匹配效果易受图像质量下降影响;而 SuperPoint 网络由于采用自监督训练机制与端到端联合架构,在面对光照变化与模糊干扰时表现出较强的鲁棒性,因而在水下图像处理任务中更具优势。Wu 等^[20]将自注意力机制引入 SuperPoint 网络,增强了其对水下模糊图像中退化特征的感知能力,进一步验证了 SuperPoint 网络在水下场景中的良好扩展性与应用潜力。

为提升 SuperPoint 网络水下光学图像特征提取的实时性和鲁棒性,本文基于 SuperPoint 网络提出一种轻量化 SuperPoint 网络(以下简称“本文网络”),其主要特点为:①针对水下图像通常存在颜色失真、模糊等细节退化问题,构建频域-空间域动态注意力融合模块(frequency-spatial dynamic attention fusion module,FSDAFM),利用注意力机制,融合频域与空间域的特征信息,提升网络对水下退化图像的特征提取能力;②借鉴深度可分离卷积^[21](depthwise separable convolution,DSC)思想,构建一种残差特征增强深度可分离卷积模块(residual feature enhancement depthwise separable convolution module,RFDS),在显著降低网络参数量、实现模型轻量化的同时,通过特征增强和残差学习机制提升网络的特征提取能力。

1 研究方法

1.1 网络总体结构

基于 SuperPoint 网络的轻量化特征提取网络,由改进共享编码层、特征点检测层、描述符解码层 3 部分构成,可同时检测及描述特征点,网络结构如图 1 所示,图中 H 、 W 分别为图像的高度、宽度。该网络使用自监督学习的方法进行训练,训练过程包括:①训练特征点检测器 MagicPoint 网络^[22],该网络由改进共享编码层和特征点检测层组成,不包含描述符解码层,训练数据集为 Synthetic Shapes 数据集。训练过程中,采用数据增强技术,对每幅图像进行随机缩放和旋转,加入随机高斯噪声、运动模糊、亮度变化等,提高网络对光照、视角变化的鲁棒性,使网络初步具有检测特征点的能力。②使用训练好的 MagicPoint 网络从真实场景图像数据集中检测每一幅图像的特征点,并将检测到的特征点视为真实标签。在此过程中,输入图像会进行 N 次单应性变换。每次单应性变换后,均会在变换后的图像上检测到特征点。利用逆单应性变换,将所有 N 次变换中检测到的特征点重新映射回原始图像,确保提取到足够数量的特征点。③在第②步获得的标注真实场景数据集上训练本文网络,通过训练输入图像及 MagicPoint 网络提供的真实标签,在特征点检测层输出特征点并在描述符解码层输出其对应的描述符。

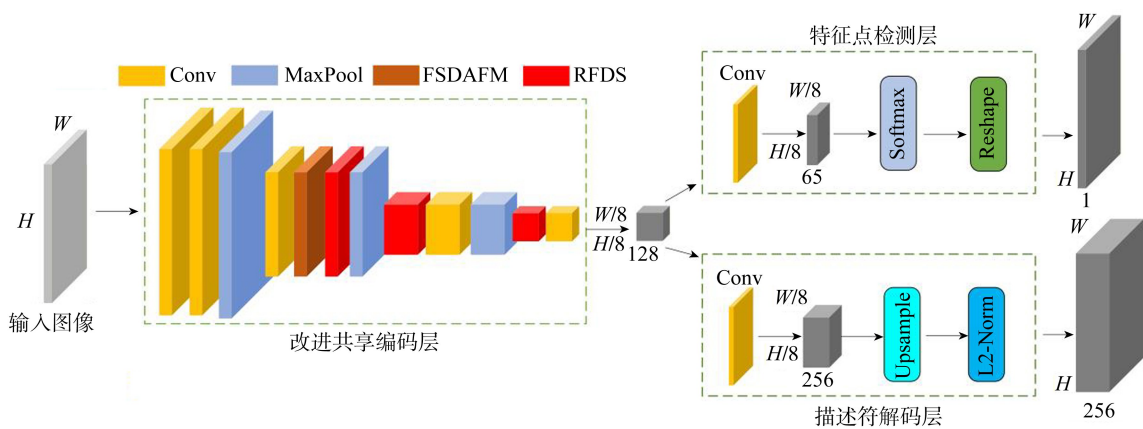


图 1 基于 SuperPoint 的轻量化特征提取网络框架

Fig. 1 Lightweight feature extraction network framework based on SuperPoint

1.2 基于 FSDAFM 和 RFDS 的改进共享编码层结构

1.2.1 FSDAFM

FSDAFM 的结构如图 2 所示,主要包含频域特征提取分支和空间域多尺度特征融合分支。①频域特征提取分支:设输入特征图 $X \in \mathbf{R}^{C \times H \times W}$,经傅里叶变换将特征图从空间域转换到频率域,得到实部图和虚部图,将二者拼接得到频域特征图,然后对频域特征图进行适当特征增强,并将增强后的频域特征图再次分为实部图和虚部图,重构为复数形式。通过傅里叶反变换将复数形式的频域图转换为特征图 $F_{\text{spatial}}^{\text{freq}}$,用于后续的特征融合。②空间域多尺度特征融合分支:设空间分支包含 3 个并行支路,每个支路考虑轻量化设计,采用深度可分离卷积体现轻量化设计。采用动态权重生成器计算每一个分支的权重,然后通过融合 3 个分支的特征,得到空间域多尺度融合输出特征图 $F_{\text{spatial}}^{\text{space}}$,拼接 $F_{\text{spatial}}^{\text{freq}}$ 和 $F_{\text{spatial}}^{\text{space}}$ 特征图并计算注意力图。

FSDAFM 针对水下图像常见的模糊、色彩失真等退化问题,基于水下退化在频域表现出的独特频率选择性进行设计。水下成像中,前向散射破坏低频成分导致全局模糊,而光吸收和后向散射则削弱高频细节信

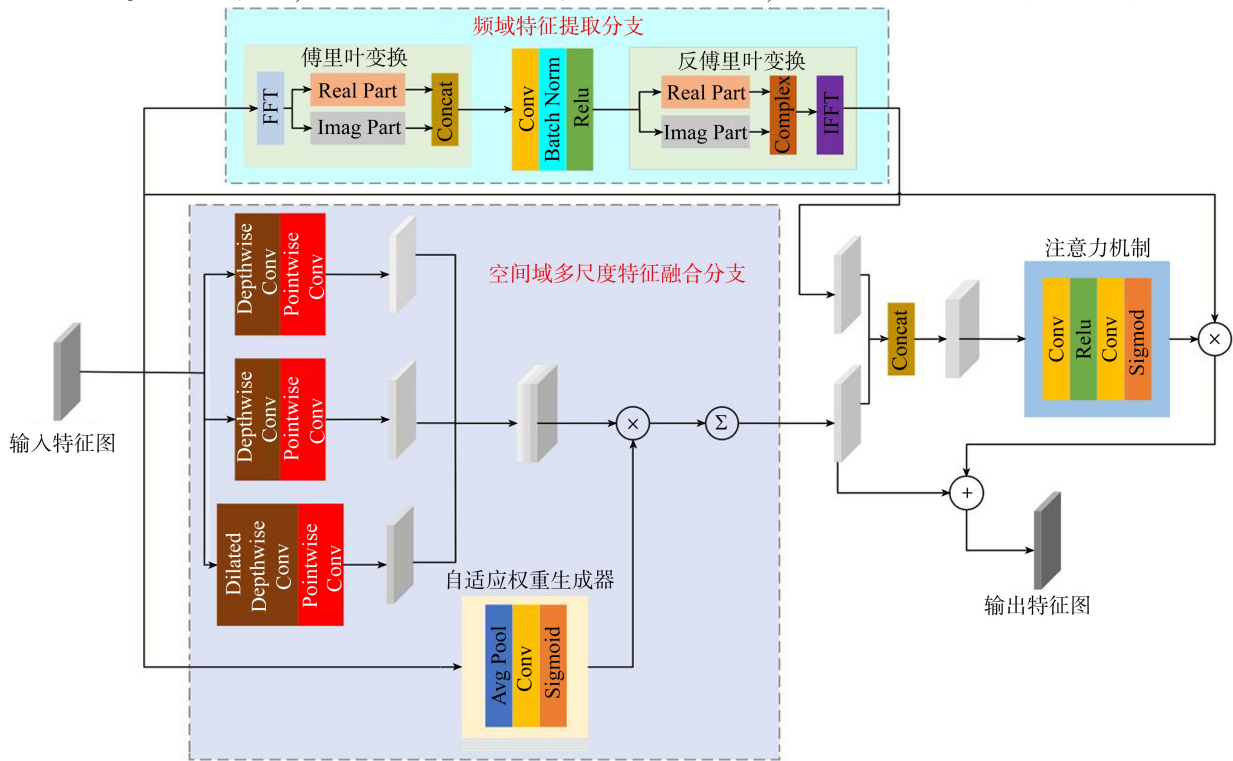


图 2 频域-空间域动态注意力融合模块

Fig. 2 FSDAFM

息。模块的频域分支通过傅里叶变换处理输入特征,将实部和虚部沿通道拼接,使 1×1 卷积层能够作为可学习频谱滤波器,自适应增强高频细节并抑制低频噪声。空间分支提取多尺度特征信息,通过注意力机制将频域优化后的特征与空间特征动态融合。这种物理机制引导的设计可提升模块对水下图像复杂退化的适应能力和特征提取鲁棒性。

1.2.2 RFDS

RFDS 的结构如图 3 所示,图中 C_{in} 、 C_{out} 分别为输入、输出通道数目。设输入特征图为 $X \in \mathbf{R}^{C_{in} \times H \times W}$,首先通过特征增强分支,生成每个通道的增强权重 w 。输入特征图 X 与 w ,逐通道乘积,得到特征增强后的输入特征图 $X_{enchanse}$ 。借鉴 DSC 思想^[21],对 $X_{enchanse}$ 进行深度卷积和点卷积,然后引入残差连接思想,设计通道适配函数 $\varphi(\cdot)$,当输入通道数与输出通道数不同,则输入特征图 X 通过适配层调整输入维度,否则残差 $X_{residual}$ 即为输入特征图 X 。将深度可分离卷积的输出与残差相加得到输出特征图。

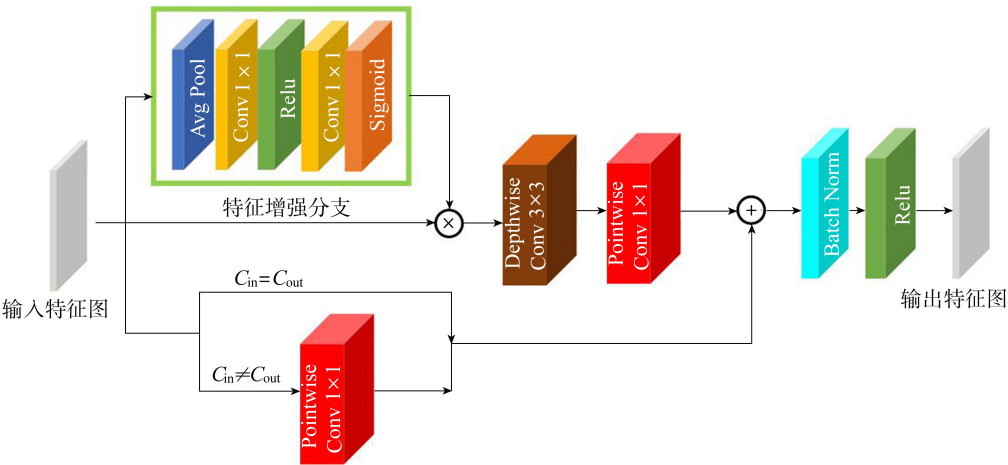


图 3 残差特征增强深度可分离卷积模块

Fig. 3 RFDS

RFDS 通过引入 DSC 降低了模型的参数量,从而实现了轻量化设计,确保了计算效率与模型性能的平衡。这一设计有效减少了计算资源的消耗,尤其在处理高维数据时,能够显著提升计算速度与存储效率。同时,RFDS 通过残差连接和特征增强分支,进一步增强了网络的特征提取能力。残差连接通过跨层信息的传递,缓解了梯度消失问题,保证信息的有效流动;而特征增强分支则通过轻量化的全局信息捕获机制,强化了网络的上下文感知能力,帮助网络更好地捕捉全局特征,提升模型在水下复杂任务中的表现。

1.2.3 共享编码层结构改进

原 SuperPoint 网络共享编码层采用类似 VGG^[23] 架构,包含 4 个处理阶段,每个阶段针对不同尺寸的特征图进行特征提取。尽管该架构设计简单直观,但其堆叠的普通卷积层导致网络深度较大、计算负担重、参数量大,难以满足水下检测任务中图像特征提取对实时性的需求。用 RFDS 替代原共享编码层中第 2~4 阶段的部分普通卷积操作,可降低网络参数量和计算量、提升推理速度,同时提升网络编码层的特征提取能力。考虑到水下退化图像在频域中呈现的独特特性,低频分量受散射噪声主导,而高频细节则因光吸收而减弱,在共享编码层第 2 阶段结构中引入了 FSDAFM。该模块通过其频域处理分支,有针对性地抑制低频噪声并增强高频结构信息;同时利用注意力机制将频域优化后的特征与空间域特征动态融合,从而提高网络在水下图像特征提取上的鲁棒性。改进共享编码层结构见表 1。

2 实例验证

实验硬件平台为 Inter Core i7-13700KF@3.4 GHz CPU 和 NVIDIA RTX 3080 GPU 设备,模型基于 PyTorch 架构实现,学习率为 0.001,图像分辨率设置为 240×320 像素,单应性变换次数 N 设置为 100。

2.1 数据集及数据增强

采用 Synthetic Shapes^[22]、RUOD^[24] 数据集训练网络检测特征点能力,并在 SQUID^[25]、FLSea^[26] 数据集进行试验。

表 1 改进共享编码层结构

Table 1 Improved shared encoding layer structure

处理阶段	编码层	输入尺寸	输出尺寸
第 1 阶段	Conv	$W \times H \times 1$	$W \times H \times 64$
	Conv	$W \times H \times 64$	$W \times H \times 64$
	MaxPool	$W \times H \times 64$	$W/2 \times H/2 \times 64$
第 2 阶段	Conv	$W/2 \times H/2 \times 64$	$W/2 \times H/2 \times 64$
	FSDAFM	$W/2 \times H/2 \times 64$	$W/2 \times H/2 \times 64$
	RFDS	$W/2 \times H/2 \times 64$	$W/2 \times H/2 \times 64$
	MaxPool	$W/2 \times H/2 \times 64$	$W/4 \times H/4 \times 64$
第 3 阶段	RFDS	$W/4 \times H/4 \times 64$	$W/4 \times H/4 \times 128$
	Conv	$W/4 \times H/4 \times 128$	$W/4 \times H/4 \times 128$
	MaxPool	$W/4 \times H/4 \times 128$	$W/8 \times H/8 \times 128$
第 4 阶段	RFDS	$W/8 \times H/8 \times 128$	$W/8 \times H/8 \times 128$
	Conv	$W/8 \times H/8 \times 128$	$W/8 \times H/8 \times 128$

Synthetic Shapes 数据集是一个用于特征点检测的人工合成功形状数据集。该数据集是在现有 Opencv 函数的基础上设计的,包含了多种简单的几何形状,如三角形、四边形、星形、直线、棋盘、三维立方体、椭圆等,其特点是图像简单,且能够提供特征点的真实标签,因此用于训练特征点检测器 MagicPoint 网络。

RUOD 数据集为水下真实场景数据集,涵盖一般水下场景,总共包含 14 000 张图像。本文将数据集划分为训练集和测试集,其中训练集包含 9 800 张图像,测试集包含 4 200 张图像。训练过程中,MagicPoint 特征点检测器从 RUOD 数据集中获取每张图像的真实标签,本文网络则利用这些水下图像及其对应标签进行训练。

SQUID 数据集包含来自 4 个不同地点的 57 对图像,分为 4 个子集,其中两个地点位于红海,代表热带水域,另外两个地点位于地中海,代表温带水域。在红海,Katzaa 是一个深度为 10~15 m 的珊瑚礁,包含 15 对图像,Satil 是一个深度为 20~30 m 的沉船,包含 8 对图像。在地中海,图像来自两个相隔 30 km 的岩石礁环境,其中 Nachsholim 的水深为 3~6 m,包含 13 对图像,Mikhmoret 的水深为 10~12 m,包含 21 对图像。但 SQUID 数据集在视角变化和光照变化方面存在不足,为此本文参考陆地场景特征提取性能测试基准中常用的 HPatches 数据集^[27],从视角变化和光照变化两个方向分别对 SQUID 数据集进行增强操作。通过调整图像亮度模拟光照变化,并利用仿射变换和尺度缩放模拟视角变化,从而完成对 SQUID 数据集的数据增强,增强后的 SQUID 数据集包含 57 个光照变化较大的场景和 57 个视角变化较大的场景。由于 HPatches 数据集仅包含陆地场景,因此在实验部分未采用该数据集进行模型性能评估。增强后的 SQUID 数据集的图像示例如图 4 所示。SQUID 数据集以及增强后的 SQUID 数据集用于评估算法的特征点检测能力及匹配性能。

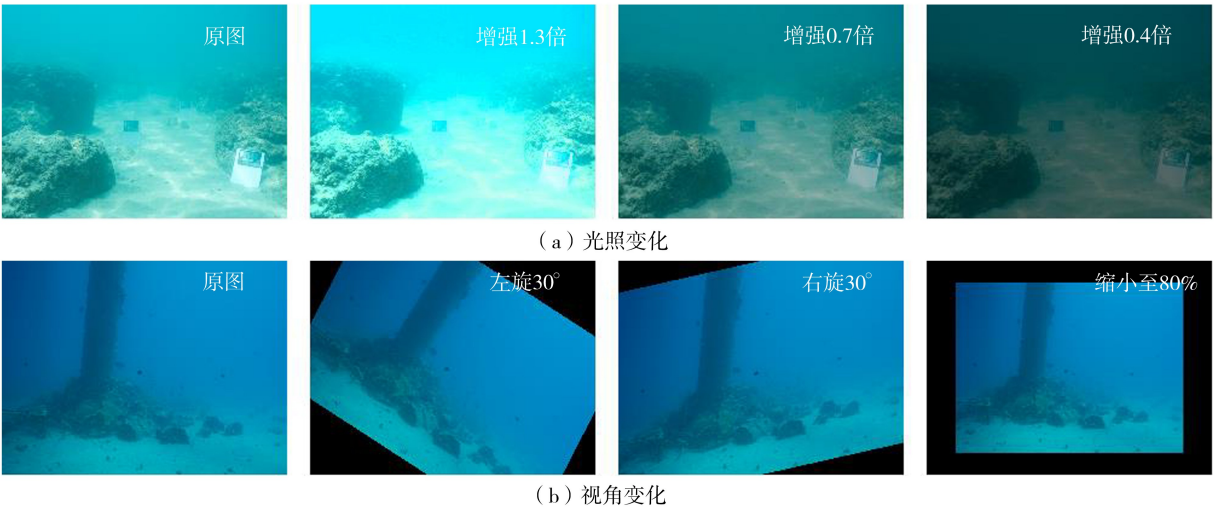


图 4 增强后的 SQUID 数据集的图像示例

Fig. 4 Image examples of enhanced SQUID dataset

FLSea 数据集也包含来自红海和地中海拍摄的水下双目图像,共包含 4 个子集。其中,Canyon1 和 Canyon2 分别为水下峡谷自然场景,包含 3 803 对和 2 362 对图像;Rock_garden1 和 Rock_garden2 为水下平坦沙区场景,分别包含 867 对和 305 对图像。同样用于评估算法的特征点检测能力及匹配性能。

2.2 评价指标

评价主要从算法的轻量化及实时性、算法特征点检测能力和匹配性能两方面开展。在轻量化和实时性方面,采用参数量、计算量和帧率作为评价指标。参数量表示模型的大小,参数量越小,表明模型越轻量化;计算量表示一次前向传播所需的浮点运算次数(单位为 10 亿次),计算量越低模型的计算复杂度越低,不需要高性能的硬件支持;帧率则用于衡量模型执行一次水下图像推理所需的时间,帧率越高模型的实时性越好。

在算法特征点检测能力和匹配性能方面,采用特征点重复率和描述符匹配准确率作为特征点检测和描述符匹配的性能指标。其中特征点重复率是衡量算法在不同视角、尺度和光照条件下对相同特征点检测一致性的指标。较高的特征点重复率表明算法在复杂的环境下仍能稳定地检测出相同的特征点,具有较强的鲁棒性。描述符匹配准确率则用于衡量特征点描述符匹配的精确度,该值越高,算法在匹配任务中能找到正确的匹配对的准确性越高。

2.3 结果分析

2.3.1 轻量化和实时性分析

采用参数量、计算量和帧率作为评价标准,将 SuperPoint 网络、SuperPoint + RFDS 网络、SuperPoint + FSDAFM 网络、本文网络进行比较,结果见表 2。将 RFDS 替换普通卷积后,SuperPoint 网络的参数量和计算复杂度显著下降,同时帧率大幅提升,表明 RFDS 的引入能够有效实现网络的轻量化,满足实时性需求。尽管 FSDAFM 会增加额外的模型参数和计算复杂度,但由于 RFDS 的优化作用,本文网络相较于 SuperPoint 网络在参数量上减少了 13.8%,计算量降低了 8.0%,帧率提升了 31.7%,仍能保持轻量化并提升实时性。

2.3.2 特征点检测的有效性分析

对不同方法采用特征点重复率进行验证,即当前图像上检测到的特征点经过光照变化或视角变化在下一张图像上仍能稳定被检测到的概率。计算重复率的方法:选择前 k 个特征点,使用单应性矩阵(光照变化场景下单应性矩阵为单位矩阵)将原图的关键点映射到变换后的图像上,检查对应点是否在给定距离阈值内匹配,再计算匹配点占总点数的比例。 k 为 300,距离阈值为 3 像素。针对水下光照变化、视角变化,在数据增强后的 SQUID 数据集上的对比试验结果见表 3。

由表 3 可知,在水下场景中,SuperPoint 网络相比于传统特征点检测器,在光照变化和视角变化方面表现出更好的鲁棒性。与深度学习方法对比,ALIKED 网络在视角变化方面的性能优于 SuperPoint 网络。传统特征点检测方法在水下环境中,由于图像质量下降,依赖的边缘和角点等特征变得模糊不清,从而使得特征点检测变得困难,导致特征点重复率较低。SuperPoint+FSDAFM 网络在特征点重复率上优于 SuperPoint 网络,尤其在视角变化的场景中重复率达到最高,表明 FSDAFM 能够有效应对水下复杂场景,提高模型在水下特征点检测任务中的鲁棒性。SuperPoint+RFDS 网络在参数量上有所减少,在光照、视角变化的水下场景中,其重复率均优于 SuperPoint 网络,表明 RFDS 的特征增强和残差连接机制确实能够提升特征提取能力。结合 RFDS 和 FSDAFM 的本文网络,在光照变化下达到了最高的重复率,虽然视角变化下略低于 SuperPoint+FSDAFM 网络,但在特征点重复率方面仍表现最佳。相较于 SuperPoint 网络,本文网络水下光照变化下的重复率提高了

表 2 各网络的消融实验结果

Table 2 Ablation experiment results for each network

网络	参数量/ 10^6 个	计算量/亿次	帧率
SuperPoint	1.304	66.2	192
SuperPoint+RFDS	1.088	55.8	291
SuperPoint+FSDAFM	1.339	71.3	245
本文网络	1.124	60.9	253

表 3 各方法的特征点重复率

Table 3 Repetition rate of feature point for each method

方法	重复率/%	
	光照变化	视角变化
SuperPoint ^[16]	84.2	72.9
D2-Net ^[14]	63.6	29.1
R2D2 ^[15]	66.9	61.6
ALIKED ^[17]	76.3	74.1
EdgePoint2 ^[19]	78.5	48.1
SIFT ^[8]	48.5	57.6
ORB ^[10]	56.6	68.0
Harris ^[6]	82.1	38.2
FAST ^[7]	50.2	62.3
SuperPoint+RFDS	85.2	74.0
SuperPoint+FSDAFM	84.7	75.2
本文网络	86.5	75.0

2.3%,视角变化下的重复率提高了 2.1%。

为了避免偶然性,不同阈值设定下,对比不同方法的特征点重复率,结果如图 5 所示。在小阈值条件下,ALIKED 在视角变化方面具有最高的特征点重复率,展现出较强的特征点检测能力。但在不同阈值设定下进行综合评估时,本文网络在水下光照和视角变化场景中的平均重复率表现最优,有效提升了水下图像特征点检测的整体性能。

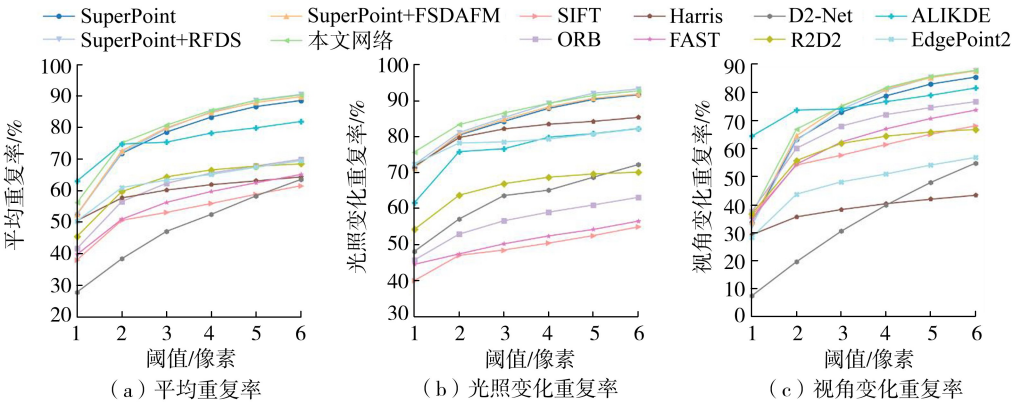


图 5 不同阈值下各方法的特征点重复率

Fig. 5 Feature point repetition rate for each method under different thresholds

2.3.3 特征点匹配结果及分析

分别对本文网络、SuperPoint 网络、D2-Net 网络、R2D2 网络、ALIKED 网络、EdgePoint2 网络以及传统的 ORB 方法、SIFT 方法在 FLSea 数据集和 SQUID 数据集上进行特征点匹配测试。

为客观评估算法生成的特征点描述符性能,采用描述符匹配准确率作为评价指标,该指标衡量匹配成功的特征点数与总特征点数的比率。为便于计算,在 FLSea 数据集和 SQUID 数据集上进行测试,采用暴力匹配算法进行描述符匹配,设置描述符最近邻匹配阈值为 0.8,最大检测特征点数目为 1 000,匹配结果见表 4 和表 5。基于学习生成的描述符能够在水下场景中有效捕捉到更具描述性和显著性的特征,而传统方法人工设计的描述符在水下图像中,由于常见的颜色失真、模糊等退化效应,其区分性和鲁棒性较差,导致匹配准确率较低。本文网络在 FLSea 数据集和 SQUID 数据集上的匹配准确率平均值分别达到 24.0%和 51.4%,虽然在部分子集上未达到最佳匹配率,但本文网络的平均匹配准确率仍高于 D2-Net 网络、R2D2 网络、ALIKED 网络以及 EdgePoint2 等。各方法的匹配结果如图 6 所示,在水下退化图像中,本文网络检测到的特征点数量多,且分布均匀,误匹配的情况少。对特征点匹配结果的量化和可视化分析表明,本文网络生成的特征点描述符具备更强的区分性、更高的鲁棒性以及更优的匹配性能。

表 4 FLSea 数据集中各方法的描述符匹配准确率对比 单位: %

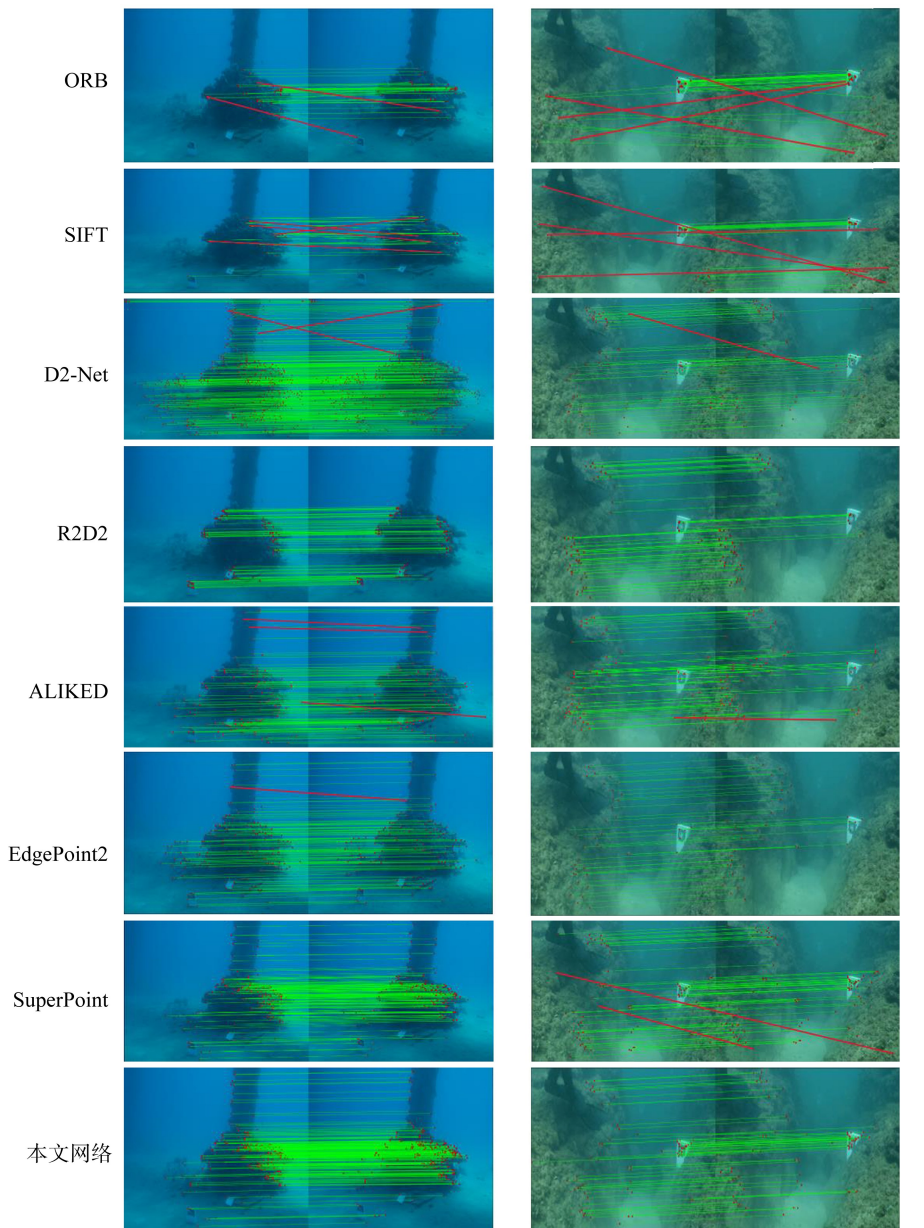
Table 4 Descriptor matching accuracy for each method in FLSea dataset unit: %

FLSea 数据集	SuperPoint ^[16]	SIFT ^[8]	ORB ^[10]	D2-Net ^[14]	R2D2 ^[15]	ALIKED ^[17]	EdgePoint2 ^[19]	本文网络
Canyon1	26.7	16.9	6.5	20.1	21.9	28.1	28.0	28.2
Canyon2	25.7	17.4	8.8	18.9	20.9	26.0	25.4	26.5
Rock_garden1	25.5	18.0	8.5	20.7	21.0	25.1	25.8	25.2
Rock_garden2	14.5	13.3	6.2	13.1	13.8	15.3	15.7	16.2

表 5 SQUID 数据集中各方法的描述符匹配准确率对比 单位: %

Table 5 Descriptor matching accuracy for each method in SQUID dataset unit: %

SQUID 数据集	SuperPoint ^[16]	SIFT ^[8]	ORB ^[10]	D2-Net ^[14]	R2D2 ^[15]	ALIKED ^[17]	EdgePoint2 ^[19]	本文网络
Katzaa	64.0	41.7	14.9	49.4	45.2	63.9	55.6	63.0
Michmoret	44.5	45.3	19.1	39.3	37.8	44.8	47.1	46.5
Nachsolim	35.2	20.4	7.5	25.6	20.6	37.5	36.1	35.5
Satil	57.5	37.8	11.8	41.0	45.6	57.1	60.1	60.7



(a) 水下偏蓝色调的图像对 (b) 受到水下杂质影响而变得模糊的图像对

注:算法检测出的特征点用红色标出,红色连线表示这些特征点存在误匹配。

图 6 特征匹配结果

Fig. 6 Feature matching results

3 结 语

本文提出了一种基于轻量化 SuperPoint 网络的水下光学图像特征提取方法。针对水下图像普遍存在的色彩失真、模糊等细节退化问题,利用注意力机制构建频域-空间域动态注意力模块(FSDAFM),通过融合频域与空间域特征信息,增强网络对水下退化图像的特征提取能力。为降低模型复杂度并提升实时性,设计了残差特征增强深度可分离卷积模块(RFDS)以取代普通卷积,在显著减少网络参数数量的同时实现模型轻量化;通过特征增强与残差学习机制,提升网络的特征提取能力。实验结果表明,本文网络的参数量相较于 SuperPoint 网络显著降低,成功实现了轻量化设计;在 SQUID 和 FLSea 数据集上的特征点检测与匹配评估中,展现了最佳的特征点检测重复率和更强的描述符区分性,对于水下光学图像有较好的特征提取鲁棒性。

参考文献:

[1] Zhao Xuanran, Wu Yan, Hu Xin, et al. A novel dual-branch global and local feature extraction network for SAR and optical image

- registration[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024, 17: 17637-17650.
- [2] 李嘉祥, 范影乐, 武薇. 基于生物视觉机制的图像特征点检测方法[J]. 计算机工程与应用, 2018, 54(7): 182-187. (Li Jiaxiang, Fan Yingle, Wu Wei. Image feature point detection method based on biological vision mechanism[J]. Computer Engineering and Applications, 2018, 54(7): 182-187. (in Chinese))
- [3] 徐晶, 包利东, 方明, 等. 基于 Gabor 表示特征描述符的红外可见光图像配准[J]. 激光与光电子学进展, 2024, 61(14): 1437007. (Xu Jing, Bao Lidong, Fang Ming, et al. Infrared visible image registration based on Gabor representation feature descriptor[J]. Laser & Optoelectronics Progress, 2024, 61(14): 1437007. (in Chinese))
- [4] Zhang Gangping, Li Chaofeng, Yan Jiajia, et al. ULD-CycleGAN: an underwater light field and depth map-optimized CycleGAN for underwater image enhancement[J]. IEEE Journal of Oceanic Engineering, 2024, 49(4): 1275-1288.
- [5] Quan Dou, Wang Zhe, Lyu Chonghua, et al. LM-Net: a lightweight matching network for remote sensing image matching and registration[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5229313.
- [6] Harris C G, Stephens M. A combined corner and edge detector[C]//Proceedings of the Alvey Vision Conference. Manchester: Alvey Vision Club, 1988: 1-6.
- [7] Rosten E, Drummond T. Machine learning for high-speed corner detection[C]//9th European Conference on Computer Vision. Graz, Austria: Springer, 2006: 430-443.
- [8] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [9] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF)[J]. Computer Vision and Image Understanding, 2008, 110(3): 346-359.
- [10] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF[C]//2011 International Conference on Computer Vision. Barcelona: IEEE, 2011: 2564-2571.
- [11] Calonder M, Lepetit V, Strecha C, et al. BRIEF: binary robust independent elementary features[C]//11th European Conference on Computer Vision. Heraklion, Crete, Greece: Springer, 2010: 778-792.
- [12] 张文静, 张振, 黄剑, 等. 基于图像语义分割的水位智能监测方法[J]. 河海大学学报(自然科学版), 2023, 51(5): 24-30. (Zhang Wenjing, Zhang Zhen, Huang Jian, et al. Intelligent water-level monitoring method based on image semantic segmentation[J]. Journal of Hohai University (Natural Sciences), 2023, 51(5): 24-30. (in Chinese))
- [13] 赵二峰, 李章寅, 袁冬阳. 基于双阶段注意力机制的大坝变形深度学习预测模型[J]. 河海大学学报(自然科学版), 2023, 51(6): 44-52. (Zhao Erfeng, Li Zhangyin, Yuan Dongyang. Deep learning model for deformation prediction of dam based on dual-stage attention mechanism[J]. Journal of Hohai University (Natural Sciences), 2023, 51(6): 44-52. (in Chinese))
- [14] Dusmanu M, Rocco I, Pajdla T, et al. D2-Net: a trainable CNN for joint description and detection of local features[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 8084-8093.
- [15] Revaud J, Weinzaepfel P, De Souza C, et al. R2d2: repeatable and reliable detector and descriptor[C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2019: 1113.
- [16] De Tone D, Malisiewicz T, Rabinovich A. SuperPoint: self-supervised interest point detection and description[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018: 224-236.
- [17] Zhao Xiaoming, Wu Xingming, Chen Weihai, et al. ALIKED: a lighter keypoint and descriptor extraction network via deformable transformation[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 5014016.
- [18] Zhao Xiaoming, Wu Xingming, Miao Jinyu, et al. ALIKE: accurate and lightweight keypoint detection and descriptor extraction[J]. IEEE Transactions on Multimedia, 2023, 25: 3101-3112.
- [19] Yao Haodi, He Fenghua, Hao Ning, et al. EdgePoint2: compact descriptors for superior efficiency and accuracy[EB/OL]. (2025-04-24). <https://arxiv.org/abs/2504.17280v1>.
- [20] Wu Di, Su Boxun, Hao Lichao, et al. A feature detection network based on self-attention mechanism for underwater image processing[J]. Ocean Engineering, 2024, 311: 118949.
- [21] Chollet F. Xception: deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 1800-1807.
- [22] Detone D, Malisiewicz T, Rabinovich A. Toward geometric deep SLAM[EB/OL]. (2017-07-24). <http://arxiv.org/abs/1707.07410>.
- [23] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[C]//3rd International Conference on Learning Representations. San Diego: 3rd International Conference on Learning Representations (ICLR), 2015.
- [24] Fu Chenping, Liu Risheng, Fan Xin, et al. Rethinking general underwater object detection: datasets, challenges, and solutions[J]. Neurocomputing, 2023, 517: 243-256.

[25] Berman D,Levy D,Avidan S,et al. Underwater single image color restoration using haze-lines and a new quantitative dataset[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2021,43(8):2822-2837.

[26] Randall Y. FLSea:underwater visual-inertial and stereo-vision forward-looking datasets[D]. Haifa:University of Haifa,2023.

[27] Balntas V,Lenc K,Vedaldi A,et al. HPatches:a benchmark and evaluation of handcrafted and learned local descriptors[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu:IEEE,2017:3852-3861.

(收稿日期:2025-06-18 编辑:刘晓艳)

(上接第118页)

[19] 李宏波,张冬月,葛学元. 基于 PSO-LSSVM 算法的隧道掘进机掘进参数预测方法[J]. 科学技术与工程,2023,23(14):6230-6237. (Li Hongbo,Zhang Dongyue,Ge Xueyuan. Tunnel boring machine tunneling parameters prediction method based on PSO-LSSVM algorithm[J]. Science Technology and Engineering,2023,23(14):6230-6237. (in Chinese))

[20] Zhang Liamo,Guo Jing,Fu Xianlei,et al. Digital twin enabled real-time advanced control of TBM operation using deep learning methods[J]. Automation in Construction,2024,158:105240.

[21] 满轲,曹子祥,刘晓丽,等. 地质突变条件下基于组合模型的围岩等级和 TBM 掘进参数预测[J]. 河海大学学报(自然科学版),2024,52(1):55-62. (Man Ke,Cao Zixiang,Liu Xiaoli,et al. Prediction of surrounding rock grades and TBM tunnelling parameters based on combined model under geological mutation conditions[J]. Journal of Hohai University (Natural Sciences),2024,52(1):55-62. (in Chinese))

[22] 徐卫亚,陈世壮,张贵科,等. 基于 PSO-LSSVM-BP 模型的高边坡力学参数反分析及稳定性评价[J]. 河海大学学报(自然科学版),2024,52(5):52-59. (Xu Weiya,Chen Shizhuang,Zhang Guike,et al. Inverse analysis of mechanical parameters of high slope based on PSO-LSSVM-BP model and stability evaluation[J]. Journal of Hohai University (Natural Sciences),2024,52(5):52-59. (in Chinese))

[23] 赵宇,郑东健. 基于 QR-KOA-ITransformer-BiLSTM 的混凝土坝变形区间预测模型[J/OL]. 河海大学学报(自然科学版),2025-02-24. <https://link.cnki.net/urlid/32.1117.tv.20250224.1500.004>. (Zhao Yu,Zheng Dongjian. A deformation interval prediction model for concrete dams based on QR-KOA-ITransformer-BiLSTM[J/OL]. Journal of Hohai University (Natural Sciences),2025-02-24. <https://link.cnki.net/urlid/32.1117.tv.20250224.1500.004>. (in Chinese))

[24] Liu Feitony,Ting Kaiming,Zhou Zhihua. Isolation forest[C]//Proceedings of the 2008 Eighth IEEE International Conference on Data Mining. Pisa,Italy:IEEE,2008:413-422. doi:10.1109/ICDM.2008.17.

(收稿日期:2025-07-02 编辑:熊水斌)

(上接第128页)

[28] 邹敏,尹平保,贺炜. 交通荷载下镍铁渣-黏土路基动力响应特性分析[J]. 公路与汽运,2024,40(6):71-75. (Zou Min,Yin Pingbao,He Wei. Analysis of characterization of dynamic response of nickel-iron slag-clay roadbed under traffic loading[J]. Highways & Automotive Applications,2024,40(6):71-75. (in Chinese))

[29] Gao Xuhe,Tian Weiping,Li Jiachun,et al. Analysis of deformation characteristics of filled subgrade treated by micro-pile[J]. Advances in Civil Engineering,2021,2021(1):5755796.

[30] 马晓凡,陈红,刘瑾,等. CT 扫描视域下黏土干湿循环劣化机理[J]. 河海大学学报(自然科学版),2023,51(5):111-118. (Ma Xiaofan,Chen Hong,Liu Jin,et al. Deterioration mechanism of clay in dry-wet cycle under CT scanning[J]. Journal of Hohai University (Natural Sciences),2023,51(5):111-118. (in Chinese))

[31] 黄玉纯. 降雨入渗对土工格室加筋路基动力响应的影响研究[J]. 公路交通科技,2023,40(11):61-71. (Huang Yuchun. Study on influence of rainfall infiltration on dynamic response of geocell-reinforced subgrade[J]. Journal of Highway and Transportation Research and Development,2023,40(11):61-71. (in Chinese))

[32] 余云燕,刘武通,付艳艳,等. 交通荷载作用红层填料道路沉降细观模拟[J]. 地震工程学报,2023,45(1):1-11. (Yu Yunyan,Liu Wutong,Fu Yanyan,et al. Meso-simulation of road settlement of red beds filling under traffic load[J]. China Earthquake Engineering Journal,2023,45(1):1-11. (in Chinese))

[33] 陈剑,苏跃宏. 交通荷载作用下公路路基动力特性的数值模拟研究[J]. 公路交通科技,2011,28(5):44-48. (Chen Jian,Su Yuehong. Numerical simulation of dynamic performance of highway subgrade under traffic loads[J]. Journal of Highway and Transportation Research and Development,2011,28(5):44-48. (in Chinese))

(收稿日期:2025-06-08 编辑:熊水斌)