# 基于混合尺度感知与边缘特征交互的伪装目标检测

蒲东兵[1],张金峰[1],张美花[1],王一可[2],孙英娟[3]

(1.东北师范大学 信息科学与技术学院,长春 130117;2.沈阳城市学院 智能与工程学院,沈阳 110112;3.长春师范大学 计算机科学与技术学院,长春 130032)

**摘 要**:针对已有伪装目标检测(COD)方法存在的边缘特征信息匮乏、多层次特征融合不充分等问题,提出了一种新颖的基于混合尺度感知与边缘特征交互的伪装目标检测网络(MSPEFI-Net)。MSPEFI-Net由混合尺度感知模块(MSP)、边缘特征交互模块(EFI)和全局特征聚合模块(GFA)组成。其中,MSP使用不同大小的卷积核提取不同尺度的特征,并进行跨尺度特征融合实现混合尺度感知,使模型能够有效扩大感受野,充分挖掘目标局部细节特征,以确保特征信息的完整性;EFI可以将高层语义信息与部分低层细节信息进行交互集成,以挖掘和补充边缘语义信息;GFA通过对上下文局部信息与边缘信息的跨层级特征聚合,以便有效地集成伪装目标及其边缘的特征信息,生成完整的预测图。实验结果表明,MSPEFI-Net在公开数据集CAMO、CHAMELEON、COD10K和NC4K上的$S_\alpha$、$F_\beta^w$、$F_\beta$、$E_\varphi$、MAE等指标整体优于其他12种对比模型,具有更好的检测性能,消融实验也验证了各模块对模型性能提升的有效性。

**关键词**:伪装目标检测;混合尺度感知;边缘特征交互;特征增强

**中图分类号**:TP391.4 **文献标志码**:A **文章编号**:1671-5497(2026)07-1996-10

**DOI**:10.13229/j.cnki.jdxbgxb.20241319

# Camouflaged object detection based on mixed scale perception and edge feature interaction

PU Dong-bing[1],ZHANG Jin-feng[1],ZHANG Mei-hua[1],WANG Yi-ke[2],SUN Ying-juan[3]

(1. *College of Information Science and Technology*, *Northeast Normal University*, *Changchun* 130117, *China*; 2. *School of Intelligence and Engineering*, *Shenyang City University*, *Shenyang* 110112, *China*; 3. *College of Computer Science and Technology*, *Changchun Normal University*, *Changchun* 130032, *China*)

**Abstract**: In view of the problems of lack of edge feature information and insufficient multi-level feature fusion in the existing amouflage object detection (COD) methods, a novel camouflaged object detection network MSPEFI-Net based on mixed-scale perception and edge feature interaction was proposed.

**收稿日期**:2024-12-11.

**基金项目**:国家自然科学基金项目(62277009);吉林省重点科技研发项目(20220201157GX);吉林省重点科技研发项目(20220201140GX);吉林省教育厅科学技术研究项目(JJKH20241014KJ).

**作者简介**:蒲东兵(1970-),男,副教授,博士.研究方向:计算机视觉,模式识别.E-mail:pudb@nenu.edu.cn

**通信作者**:孙英娟(1972-),女,教授,博士.研究方向:数据挖掘,人工智能.E-mail:happysunyj@foxmail.com

MSPEFI-Net consists of a mixed-scale perception module (MSP), an edge feature interaction module (EFI) and a global feature aggregation module (GFA). MSP extracts different scale features with different convolutional kernels, and fuses cross-scale features to achieve mixed-scale perception, so that MSPEFI-Net can effectively expand the receptive field and fully mine the local detail features of the object to ensure the integrity of feature information. EFI can interact and integrate high-level semantic information with some low-level detailed information to mine and supplement edge semantic information. GFA aggregates cross-level features from context local information and edge information to effectively integrate camouflage object features and its edge features, and generates a full prediction plot. Experimental results prove that the $S_\alpha$、$F_\beta^w$、$F_\beta$、$E_\varphi$ and MAE of MSPEFI-Net on the public datasets CAMO, CHAMELEON, COD10K and NC4K are superior to 12 comparison methods and show better detection performance of MSPEFI-Net. Ablation experiments also verify the effectiveness of each module to improve the performance of MSPEFI-Net.

**Key words**: camouflage object detection; mixed-scale perception; edge feature interaction; feature enhancements

## 0　引　言

自然界的一些动物利用自身颜色、纹理及覆盖物融入周围环境中,对捕食者准确辨识这种伪装目标提出了很多挑战。类似地,为了识别完美融入周围环境的生物,快速发现伪装的军事目标或有病灶的生理组织等[1],伪装目标检测(Camouflaged object detection, COD)技术受到越来越多研究者的重视,已经成为目标检测的热门研究方向之一。作为一种有效的预处理技术,COD在军事、医药、工业与娱乐等领域将得到广泛应用[1-4]。

伪装目标检测方法大致包括传统的COD和基于深度学习的COD[5-8]。传统COD主要基于颜色、纹理、光照等直接视觉特征检测伪装目标。而基于深度学习的COD在自动化特征学习、泛化能力和灵活性方面具有明显优势,能够更准确、高效地检测伪装目标。研究者们提出了很多基于深度学习的伪装目标检测方法,并且取得了显著效果。如Fan等[1]受动物的捕食过程启发,提出了搜索识别网络(Search identification network, SINet)模型,该模型主要由搜索模块(Search module, SM)和识别模块(Identification module, IM)构成,整个网络采取将搜索注意力和部分解码器组件(Partial decoder component, PDC)密集连接的方式对粗糙区域进行细化,但是这种密集连接方式的特征融合会引入大量特征,从而可能会造成特征冗余。通过对特征融合方式进行改进,Sun等[2]设计了基于上下文感知的跨层融合网络C$^2$FNet模型,该模型在粗略预测和细化两个阶段都使用了特征融合策略,利用注意引导的跨层次融合模块整合多尺度的特征,同时通过双分支全局上下文模块丰富的全局上下文信息,实现多尺度目标检测性能的提升。对于边缘信息,许多研究证明了边缘检测在一定程度上可以促进伪装目标检测性能的进步。Zhai等[3]提出了一种新颖的互图学习MGL网络模型,该模型将常规互图学习推广到图域,通过图形递归推理捕获高阶信息,利用区域诱导图推理和边缘压缩图推理共同挖掘有价值的补充信息,以提高COD真实边缘的可见性。Sun等[9]提出了BGNet模型,该模型探索了有价值且额外的对象相关的边缘语义,以指导COD的表示学习,这促使模型生成突出对象结构的特征,从而实现伪装目标检测的精确边界定位。

由于COD任务始终面临目标边缘不连续、轮廓模糊、边缘检测方法粗糙且不完整等问题,导致已有COD模型难以完整、有效地捕捉目标结构和细节特征信息。本文提出一种新的基于混合尺度感知与边缘特征交互的伪装目标检测网络模型(Mixed scale perception and edge feature interaction network for camouflage object detection, MSPEFI-Net),该模型主要由混合尺度感知模块(Mixed scale perception, MSP)、边缘特征交互模块(Edge feature interaction, EFI)以及全局特征聚合模块(Global feature aggregation, GFA)组成。其中,MSP用于感知局部混合尺度特征信息,挖掘并保留完整的细节特征信息。EFI通过对前景

信息与边缘信息相互补充的方式提取目标边缘信息,以集成不同层次的边缘语义特征,缓解边缘信息失衡问题。GFA通过注意力机制与跨层次连接方式,聚合和细化跨级别上下文信息,并将增强后的特征与边缘感知特征进行全局特征聚合。本文在4个基准数据集上进行了大量实验,定性、定量地分析和验证了MSPEFI-Net的先进性和有效性。

## 1 MSPEFI-Net模型

### 1.1 模型框架

如图1所示,MSPEFI-Net由MSP模块、EFI模块和GFA模块3部分组成。其中,MSP用于充分挖掘目标局部细节特征,以确保特征信息的完整性。EFI用于将高层语义信息与部分低层细节信息进行交互集成,以挖掘和补充边缘语义信息。GFA用于对上下文局部信息与边缘信息的跨层级特征聚合。对于输入图像$I\in R^{W\times H\times 3}$(其中$H$和$W$分别为图像高度和宽度),通过Res2Net提取多级特征$f_i;i\in\{1,2,3,4,5\}$。由于低级特征对模型最终性能贡献较小,且耗费计算资源,MSPEFI-Net舍弃了$f_1$和$f_2$以提高计算效率,$f_3$、$f_4$、$f_5$被馈送到3个MSP模块以扩大感受野并融合多尺度特征,获得增强后的特征$F_i;i\in\{3,4,5\}$。为降低模型计算量,增强特征的通道数被统一压缩为64。$f_5$被馈送到一个空间金字塔池化模块(Atrous spatial pyramid pooling,ASPP)和卷积模块C以保证信息的完整性。EFI模块用于挖掘目标边缘特征$f_e;e\in\{2,3,4,5\}$,包括局部边缘特征信息和全局位置语义信息。利用相邻的高层语义信息和低层细节线索指导当前层提取全局和局部信息,增强边界表示能力。将交互后的边缘信息与混合尺度高层语义信息馈送到GFA中,以自顶向下的方式逐步聚合跨级别的特征与边缘信息,生成最终预测图。

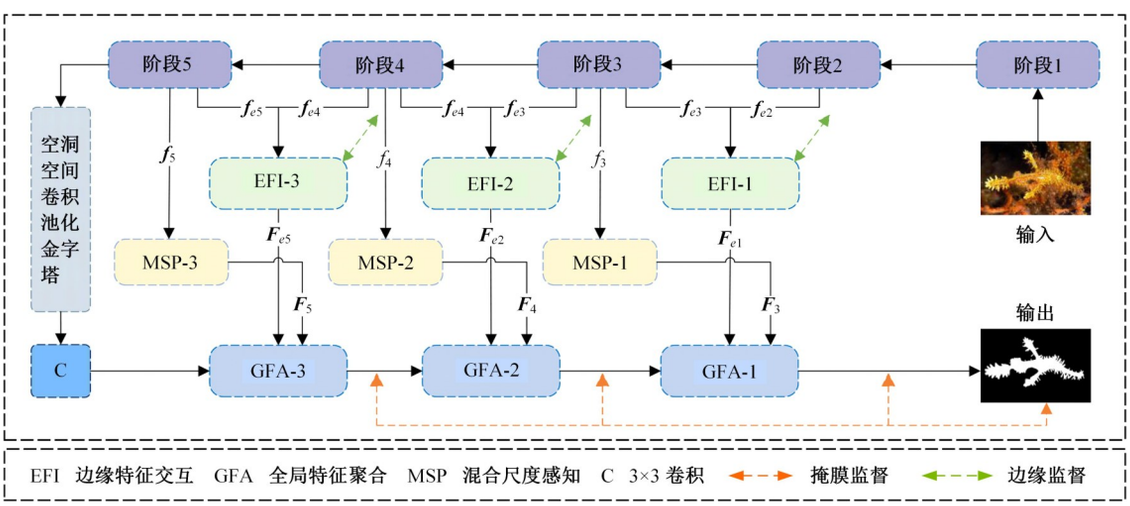



**图1 MSPEFI-Net框架**

**Fig. 1 MSPEFI-Net framework**

### 1.2 混合尺度感知模块

为增强单层特征并保持高质量的细节信息,本文在MSPEFI-Net中设计了MSP模块,使模型在感知混合尺度上下文信息的同时,保持更丰富的细节信息,如图2所示(彩图参见电子版,后文同)。

MSP通过4个分支并行处理输入特征$f_i$,用不同大小的卷积核扩大感受野以提高表示能力。每个分支包含一个1×1卷积层,用于压缩通道以提高计算效率,压缩后的特征表示为$f_p$;$p\in\{2,3,4\}$,$p$为分支序号。每个分支处理一个$f_p$,将$f_{p2}$和$f_{p3}$通过3×3卷积层以扩大感受野,提取有效的特征信息,同时对$f_{p4}$使用5×5卷积层进行特征提取,通过不同大小卷积核提取混合尺度特征,确保特征信息的完整性和有效性,如式(1)~(3)所示:

$$f_{r1}=\mathrm{Conv}_{k=1}(f_i) \tag{1}$$

$$f_{r2}=f_{r3}=\mathrm{Conv}_{k=3}(\mathrm{Conv}_{k=1}(f_i)) \tag{2}$$

$$f_{r4}=\mathrm{Conv}_{k=5}(\mathrm{Conv}_{k=1}(f_i)) \tag{3}$$

将提取信息的相邻分支进行连接,以融合不同分支的特征信息,保存更多有效特征。融合后

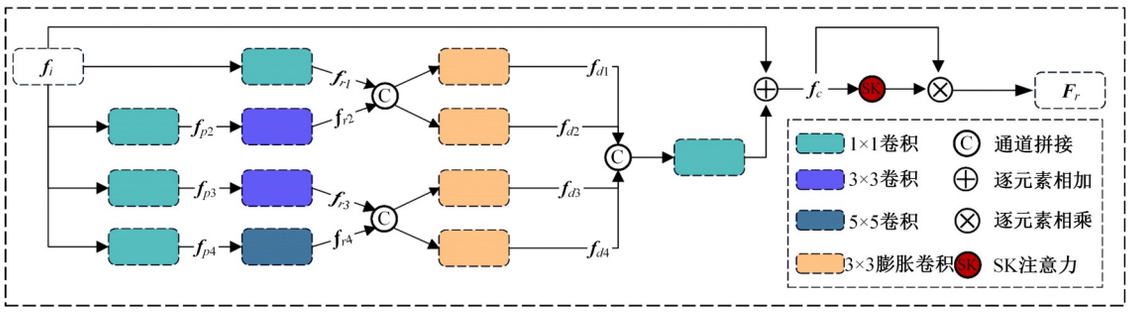



**图 2　MSP 模块**

**Fig. 2　MSP module**

的特征被馈送到不同的分支,$f_{r1}$、$f_{r2}$、$f_{r3}$、$f_{r4}$为特征馈送 $\mathrm{Conv}_{k=3,d=1}$、$\mathrm{Conv}_{k=3,d=3}$、$\mathrm{Conv}_{k=3,d=3}$、$\mathrm{Conv}_{k=3,d=5}$卷积分支,通过不同的卷积核提取特征,并融合分支之间的特征,再利用不同膨胀率卷积运算,模块获得的图像特征不仅充分保留完整的空间分辨率,同时也包含更多的上下文信息。该过程如下所示:

$$f_{d1}=\mathrm{Conv}_{k=3,d=1}\left(\mathrm{Cat}\left(f_{r1},f_{r2}\right)\right) \tag{4}$$

$$f_{d2}=\mathrm{Conv}_{k=3,d=3}\left(\mathrm{Cat}\left(f_{r1},f_{r2}\right)\right) \tag{5}$$

$$f_{d3}=\mathrm{Conv}_{k=3,d=3}\left(\mathrm{Cat}\left(f_{r3},f_{r4}\right)\right) \tag{6}$$

$$f_{d4}=\mathrm{Conv}_{k=3,d=5}\left(\mathrm{Cat}\left(f_{r3},f_{r4}\right)\right) \tag{7}$$

式中:$\mathrm{Conv}_{k,d}$为不同卷积核和膨胀率的卷积分支;$f_{di}, i\in\{1,2,3,4\}$为卷积后的特征;Cat为连接操作。

将不同分支提取的特征信息与原始特征信息进行连接,通过 1×1 卷积层与原始特征信息进行连接操作。原始特征信息具有极高的完整性,为了弥补卷积过程中的特征信息丢失,本文基于残差连接思想,引入动态选择注意力,使神经元可以自适应地调整其感受野大小,能够捕获不同尺度的目标,增强特征信息表达能力。首先,通过分离操作将输入特征通过不同大小的卷积核处理,然后将所有分支的特征图进行逐元素相加,接着对求和后的结果进行全局平均池化,通过两个全连接层(通常先降维后升维),处理池化后的特征并通过 Softmax 进行归一化,生成每个尺度的权重。最后,根据融合生成的尺度权重,动态选择并融合不同尺度的特征。通过选择性地应用不同尺度的卷积核,从而在不同层级上捕捉多尺度特征。该过程如下所示:

$$f_c=\mathrm{Conv}_{k=1}\left(\mathrm{Cat}\left(f_{d1},f_{d2},f_{d3},f_{d4}\right)\right)\oplus f_{i=5} \tag{8}$$

$$F_r=f_c\otimes\sigma\left(\mathrm{Conv}_{k=1}\left(f_c\right)\right) \tag{9}$$

式中:$F_r$为最终输出特征;$f_c$为连接操作后的特征;$\otimes$为逐元素相乘;$\oplus$为逐元素相加;$\sigma$为 Sigmoid 激活函数。

### 1.3　边缘特征交互模块

完整的边缘信息对于目标定位和检测十分重要,本文设计的 EFI 模块旨在突出伪装目标区域的边缘特征、明确目标边缘、聚合多层次特征、促进边界和区域特征的相互细化,提高边缘特征的表达能力,其结构如图 3 所示。EFI 对边缘特征 $f_e$; $e\in\{2,3,4,5\}$使用两个 1×1 卷积层压缩通道,通过级联操作将本层边缘特征与相邻层上采样后的边缘特征融合,融合后的特征被馈送到 3×3 卷积层,经归一化操作与 ReLU 激活函数获得相邻层融合后的边缘特征表示,融合后的特征被输入到 1×1 卷积层和 Sigmoid 函数得到最终边缘特征 $F_e$,如下所示。

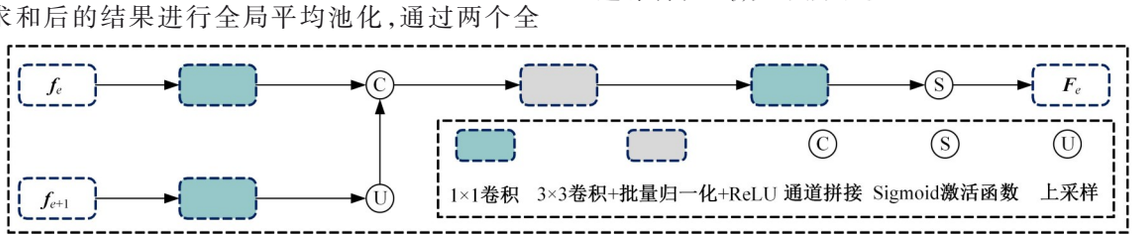



**图 3　EFI 模块**

**Fig. 3　EFI module**

$$F_e=\sigma\left(\mathrm{Conv}_{k=1}\left(\mathrm{ConvBR}\left(\mathrm{Cat}\left(\mathrm{UP}\left(\mathrm{Conv}_{k=1}\left(f_{e+1}\right),\left(\mathrm{Conv}_{k=1}\left(f_e\right)\right)\right)\right)\right)\right)\right) \tag{10}$$

式中:$Conv_k$为不同大小卷积核的卷积层;$f_e$为当前层输入的边缘特征;$f_{e+1}$为相邻层输入的边缘特征;UP为上采样;ConvBR为Conv+BN+ReLU。

### 1.4 全局特征聚合模块

在COD任务中,高级特征通常包含丰富的语义信息,而低级特征包含更详细的局部线索。为了同时利用语义信息和细节信息,集成多层次、多尺度与边缘等特征信息预测目标,本文设计了GFA模块,如图4所示。

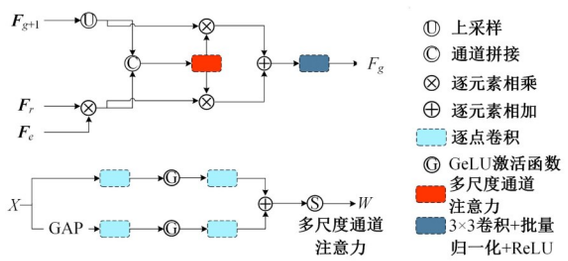



**图4 GFA模块**

**Fig. 4 GFA module**

GFA用于聚合多层次特征$F_g$、多尺度特征$F_r$以及边缘特征$F_e$,挖掘层次间的相关性和上下文语义信息。对于第$g\left(g\in\{3,4\}\right)$级的GFA模块,将经过当前级相邻$F_e$与$F_r$进行逐元素相乘,然后与上采样后的$F_{g+1}$进行连接得到$F_c$。当$g\in\{5\}$时,输入则是经过空洞空间金字塔池化(Atrous spatial pyramid pooling,ASPP)获得的包含丰富语义信息的特征图,如下所示:

$$F_c=\mathrm{Cat}\left(\mathrm{UP}\left(F_{g+1}\right),\left(F_r\otimes F_e\right)\right) \tag{11}$$

式中:$\otimes$为逐元素相乘;$F_{g+1}$为上一级GFA输出特征;$F_c$为连接操作后的特征。

将$F_c$输入多尺度通道注意机制(Multi-scale channel attention,MSCA)[10],聚合跨层级上下文特征信息,将其与上一级特征进行逐元素相乘操作,充分补充不同层次的特征,得到综合特征表示。然后,与融合边缘特征信息的当前级特征进行逐元素相加后,送入到3×3卷积层,进行批量归一化和RELU激活函数处理,最后得到邻近层融合特征$F_g$,如下所示:

$$F_g=\mathrm{ConvBR}\Big(m\left(F_c\right)\otimes F_{g+1}\oplus m\left(F_r\otimes F_e\right)\Big)\otimes m\left(F_c\right) \tag{12}$$

式中:$F_{g+1}$为相邻层模块输出特征;$m$表示MSCA模块处理。

### 1.5 损失函数

MSPEFI-Net由混合训练损失函数进行训练,分别是加权二分类交叉熵损失($L_{\mathrm{BCE}}^{w}$)、不确定性感知损失($L_{\mathrm{UAL}}$)以及Dice损失函数($L_{\mathrm{Dice}}$)。$L_{\mathrm{BCE}}^{w}$损失比标准二分类交叉熵损失更加关注较难的像素,而不是赋予每个像素相同的权重。由于伪装目标通常受背景干扰,模型在预测时会产生严重的模糊性和不确定性,本文引入不确定性感知损失($L_{\mathrm{UAL}}$)作$L_{\mathrm{BCE}}^{w}$的辅助,以提高模型对隐藏目标的预测能力。$L_{\mathrm{UAL}}$在0和1附近具有平缓的梯度,同时采用余弦策略对平衡系数$\lambda$进行动态调整。对于边缘监督,采用Dice损失($L_{\mathrm{Dice}}$),以缓解图像边缘之间的失衡问题。总体训练损失$L_{\mathrm{total}}$计算公式为:

$$L_{\mathrm{total}}=\sum_{i=2}^{5}\left(L_{\mathrm{BCE}}^{w}\left(\boldsymbol{P}_i,\boldsymbol{G}_o\right)+\lambda L_{\mathrm{UAL}}\left(\boldsymbol{P}_i,\boldsymbol{G}_o\right)\right)+\beta L_{\mathrm{Dice}}\left(\boldsymbol{P}_e,\boldsymbol{G}_e\right) \tag{13}$$

式中:$\lambda$为平衡系数,$\beta$为权衡参数,实验中设置$\lambda=0.5,\beta=3$;$\boldsymbol{P}_i$为预测图;$\boldsymbol{G}_o$为真实标签;$\boldsymbol{P}_e$为边缘预测图;$\boldsymbol{G}_e$为边缘真实标签。

## 2 实验验证和结果分析

### 2.1 实验环境及配置

MSPEFI-Net基于Linux、cuda11.1及PyTorch 1.8.0实现,使用具有单个NVIDIA GTX 4090 GPU的机器进行训练。Res2Net特征提取器的参数使用ImageNet上预训练的权重进行初始化。输入图像统一处理成384×384的分辨率,通过随机翻转和旋转策略增加训练数据以避免过度拟合,批量大小为14,采用Adam优化器,学习率从$5\times10^{-5}$开始,整个训练时间约为5 h,总共训练40轮次。

### 2.2 数据集与评价指标

本文基于4个常用的COD数据集(CAMO、CHAMELEON、COD10K和NC4K)评估MSPEFI-Net的性能。CAMO数据集共包括1 250张伪装图像,其中1 000张伪装图像用于训练,250张图像用于测试;CHAMELEON数据集包括76张用于测试的伪装图像;COD10K数据集共包括5 066张伪装图像,其中3 040张图像用于训练,2 026张图像用于测试;NC4K数据集包括4 121张图像,将全部用于测试。评价指标分别为$S$-measure($S_\alpha$)、加权$F$-measure($F_\beta^w$)、平均绝对误差(MAE)、$E$-

measure($E_\varphi$)和$F$-measure($F_\beta$)。此外，使用$F$-measure阈值曲线、精确率-召回率(PR)曲线与$E$-measure阈值曲线来更全面地评估模型。

### 2.3　MSPEFI-Net性能分析

为了证明MSPEFI-Net的有效性，本文将其与12种最先进的伪装目标检测模型进行比较，包括SINet[1]、MGL-R[3]、$C^2$FNet[2]、SINetv2[11]、BGNet[9]、ZoomNet[4]、DGNet[6]、FEDER[5]、FDNet[7]、FSPNet[8]、FINet[16]、UEDGNet[15]等网络模型。为了确保比较的公平性，上述方法的所有预测结果都是使用作者提供的原预测图像或者使用作者提供的源代码重新训练所得的预测图像进行评估。

MSPEFI-Net在4个基准数据集上与12种最新方法之间的性能对比结果如表1所示，其中，加粗数值表示该项指标的最优值，下划线数值表示该项指标的次优值，后文同。表1结果表明，MSPEFI-Net在4个数据集上的大多数评估指标都明显优于其他先进模型。例如，与最优的竞争方法UEDGNet相比，在CAMO数据集MSPEFI-Net将$F_\beta^w$、$F_\beta$分别提高了1.1%、1.5%，并将MAE降低了6.1%；在COD10K数据集MSPEFI-Net将$S_\alpha$、$F_\beta$分别提高了1.2%、3.5%；在其他数据集上与目前存在的COD检测模型相比，MSPEFI-Net的$S_\alpha$、$F_\beta^w$、$F_\beta$、$E_\varphi$以及MAE5个评价指标均有不错的检测结果，证明了MSPEFI-Net模型相比于先进方法的有效性和先进性。

**表1 基于CAMO、COD10K、CHAMELEON、NC4K数据集的性能对比**

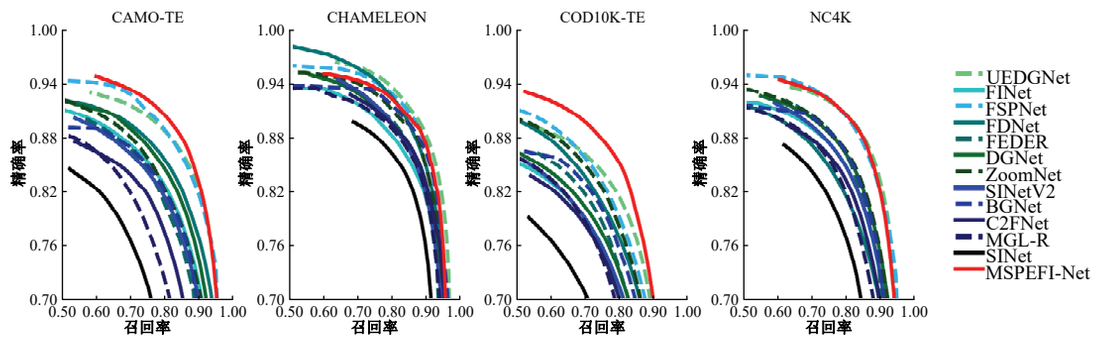
**Table 1　Performance comparison on CAMO, COD10K, CHAMELEON, and NC4K datasets**

| 方法 | CAMO | | | | | CHAMELEON | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $S_\alpha\uparrow$ | $F_\beta^w\uparrow$ | MAE↓ | $F_\beta$ | $E_\varphi\uparrow$ | $S_\alpha\uparrow$ | $F_\beta^w\uparrow$ | MAE↓ | $F_\beta$ | $E_\varphi\uparrow$ |
| SINet[1] | 0.745 | 0.644 | 0.092 | 0.702 | 0.829 | 0.872 | 0.806 | 0.034 | 0.827 | 0.946 |
| MGL-R[3] | 0.775 | 0.673 | 0.088 | 0.726 | 0.842 | 0.893 | 0.812 | 0.031 | 0.833 | 0.941 |
| $C^2$FNet[2] | 0.796 | 0.719 | 0.080 | 0.762 | 0.864 | 0.888 | 0.828 | 0.032 | 0.844 | 0.946 |
| SINetv2[11] | 0.820 | 0.743 | 0.070 | 0.782 | 0.895 | 0.888 | 0.816 | 0.030 | 0.835 | 0.961 |
| BGNet[9] | 0.812 | 0.749 | 0.073 | 0.789 | 0.882 | 0.901 | 0.850 | 0.027 | 0.860 | 0.954 |
| ZoomNet[4] | 0.820 | 0.752 | 0.066 | 0.794 | 0.892 | 0.902 | 0.845 | 0.023 | 0.864 | 0.958 |
| DGNet[6] | 0.839 | 0.769 | 0.057 | 0.806 | 0.915 | 0.890 | 0.816 | 0.029 | 0.834 | 0.956 |
| FEDER[5] | 0.802 | 0.738 | 0.071 | 0.781 | 0.873 | 0.887 | 0.834 | 0.030 | 0.851 | 0.954 |
| FDNet[7] | 0.841 | 0.775 | 0.063 | 0.807 | 0.908 | 0.897 | 0.835 | 0.027 | 0.854 | 0.964 |
| FSPNet[8] | 0.856 | 0.799 | 0.050 | 0.830 | 0.928 | 0.908 | 0.851 | 0.023 | 0.867 | 0.965 |
| FINet[16] | 0.828 | 0.752 | 0.065 | 0.791 | 0.900 | 0.883 | 0.808 | 0.031 | 0.825 | 0.957 |
| UEDGNet[15] | **0.867** | 0.816 | 0.049 | 0.841 | **0.931** | 0.909 | **0.862** | 0.023 | **0.874** | **0.969** |
| MSPEFI-Net | **0.867** | **0.825** | **0.046** | **0.854** | 0.929 | **0.910** | **0.862** | **0.020** | **0.874** | 0.964 |

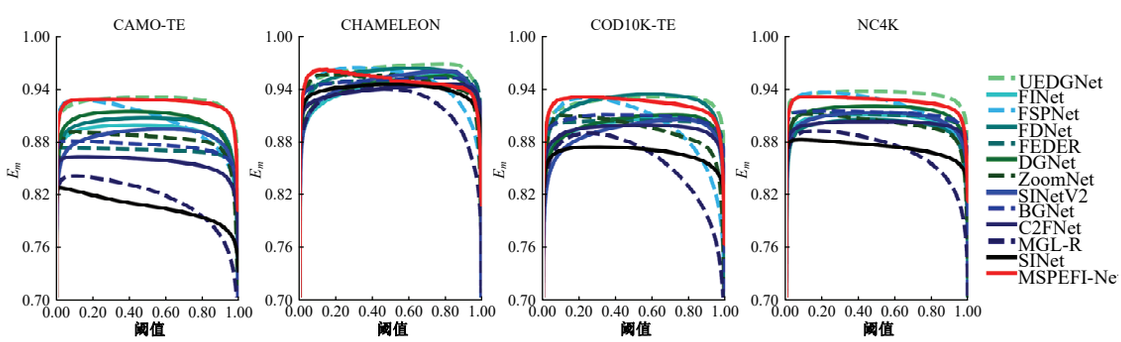
| 方法 | COD10K | | | | | NC4K | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $S_\alpha\uparrow$ | $F_\beta^w\uparrow$ | MAE↓ | $F_\beta$ | $E_\varphi\uparrow$ | $S_\alpha\uparrow$ | $F_\beta^w\uparrow$ | MAE↓ | $F_\beta$ | $E_\varphi\uparrow$ |
| SINet[1] | 0.776 | 0.631 | 0.043 | 0.679 | 0.874 | 0.808 | 0.723 | 0.058 | 0.769 | 0.883 |
| MGL-R[3] | 0.814 | 0.666 | 0.035 | 0.710 | 0.890 | 0.833 | 0.739 | 0.053 | 0.782 | 0.893 |
| $C^2$FNet[2] | 0.813 | 0.686 | 0.036 | 0.723 | 0.900 | 0.838 | 0.762 | 0.049 | 0.795 | 0.904 |
| SINetv2[11] | 0.815 | 0.680 | 0.037 | 0.718 | 0.906 | 0.847 | 0.770 | 0.048 | 0.805 | 0.914 |
| BGNet[9] | 0.831 | 0.722 | 0.033 | 0.753 | 0.911 | 0.851 | 0.788 | 0.044 | 0.820 | 0.916 |
| ZoomNet[4] | 0.838 | 0.729 | 0.029 | 0.766 | 0.911 | 0.853 | 0.784 | 0.043 | 0.818 | 0.912 |
| DGNet[6] | 0.822 | 0.693 | 0.033 | 0.728 | 0.911 | 0.857 | 0.784 | 0.042 | 0.814 | 0.922 |
| FEDER[5] | 0.822 | 0.716 | 0.032 | 0.751 | 0.905 | 0.847 | 0.789 | 0.044 | 0.824 | 0.915 |
| FDNet[7] | 0.840 | 0.729 | 0.030 | 0.757 | 0.935 | 0.834 | 0.750 | 0.052 | 0.784 | 0.905 |
| FSPNet[8] | 0.851 | 0.735 | 0.026 | 0.769 | 0.930 | 0.879 | 0.816 | 0.035 | 0.843 | 0.937 |
| FINet[16] | 0.817 | 0.686 | 0.034 | 0.723 | 0.908 | 0.847 | 0.771 | 0.047 | 0.805 | 0.911 |
| UEDGNet[15] | 0.856 | 0.760 | 0.026 | 0.785 | 0.932 | **0.882** | 0.828 | 0.035 | 0.852 | **0.938** |
| MSPEFI-Net | **0.867** | **0.784** | **0.022** | **0.813** | **0.935** | 0.879 | **0.830** | **0.034** | **0.857** | 0.936 |

注：评价指标包括$S$-measure、$E$-measure、加权$F$-measure、$F$-measure和平均绝对误差，对于一个确定的样本，$S_\alpha$、$F_\beta^w$、$E_\varphi$越大，MAE越小则检测效果越好；“↑”和“↓”分别表示越大越好和越小越好。

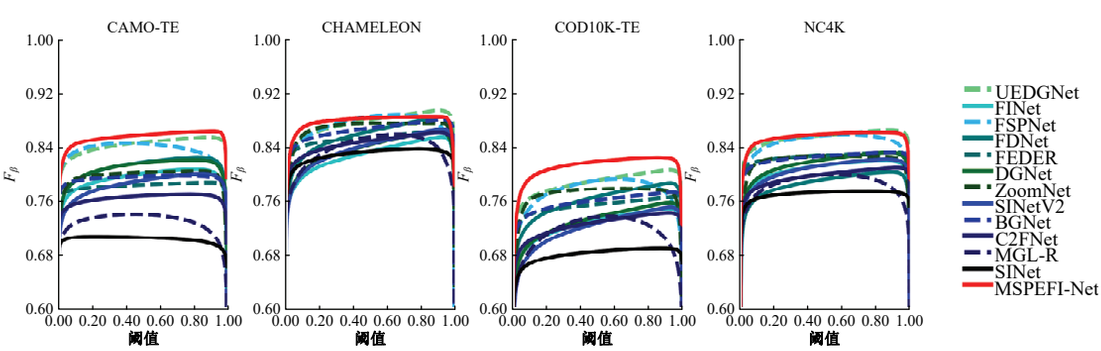
MSPEFI-Net和其他方法在4个数据集上的$PR$、$F_\beta$和$E_m$曲线对比结果如图5所示。由图5(a)可见,精准率数值与召回率数值越大,表示模型效果越好;由图5(b)可见,在相同阈值下,$E_m$值越大,表示模型效果越好;由图5(c)可见,在相同阈值下,$F_\beta$值越大,表示模型效果越好。图5表明,MSPEFI-Net在CAMO、COD10K、NC4K上均优于其他的竞争方法,在CHAMELEON数据集上次优,这是由于该数据集较小,验证效果具有特殊性。




(a)$FR$曲线对比




(b)$E_m$曲线对比




(c)$F_\beta$曲线对比

**图5 MSPEFI-Net与最先进方法在4个数据集上的$PR$、$F_\beta$、$E_m$曲线对比**

**Fig. 5 $PR$、$F_\beta$、$E_m$ comparison curves of MSPEFI-Net against state-of-the-art methods on four datasets**

## 2.4 可视化对比分析

为了直观分析MSPEFI-Net的性能,本文对MSPEFI-Net与其他几种最先进的算法进行了定性比较,本文可视化了MSPEFI-Net和其他先进的方法生成的一些具有挑战性的场景和结果,如图6所示。其中,图6(a)为数据集图片、图6(b)为真实标签、图6(c)为MSPEFI-Net检测效果、图6(d)~(k)分别为不同方法的检测效果。图6中圆圈所示区域表明,MSPEFI-Net在处理边缘细节方面比其他先进的方法具有更好的效果,如FEDER第4幅图左下角检测有粘连,第8幅图右上角有细节缺失,而MSPEFI-Net在这些细节处理方面具有更好的性能。相比之下,MSPEFI-Net对比其他方法在边界高度模糊、背景高度相似的情况下依然可以准确地预测不同尺度的对象,特别是对于具有复杂边缘图像的检测,预测结果具有更清晰

的边界和区域形状,展现出了优越的性能。

图 7 为本文模型 MSPEFI-Net 与 BGNet、MGL在边界相关边缘提取方面的视觉比较。由图7可以看出,虽然BGNet是一种针对边缘线索的模型,但在面对复杂、细长图像时,检测效果不佳;MGL是一种基于复杂图模型的辅助边缘检测网

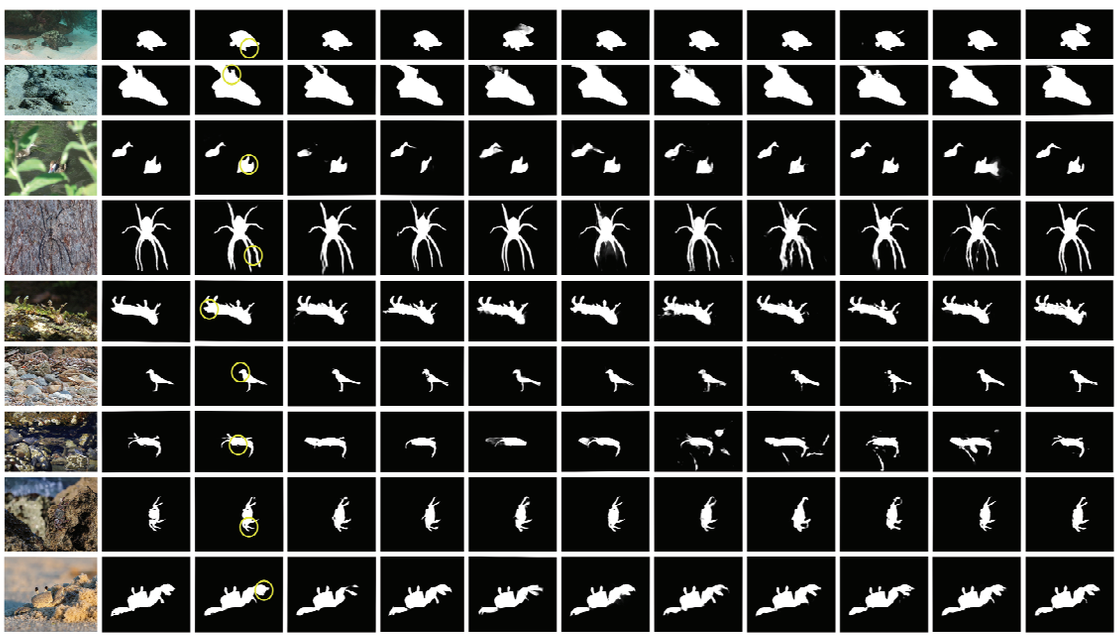



**图 6　MSPEFI-Net与其他先进方法可视化结果对比**

**Fig. 6　Visual comparisons between MSPEFI-Net and State-of-the-Art**

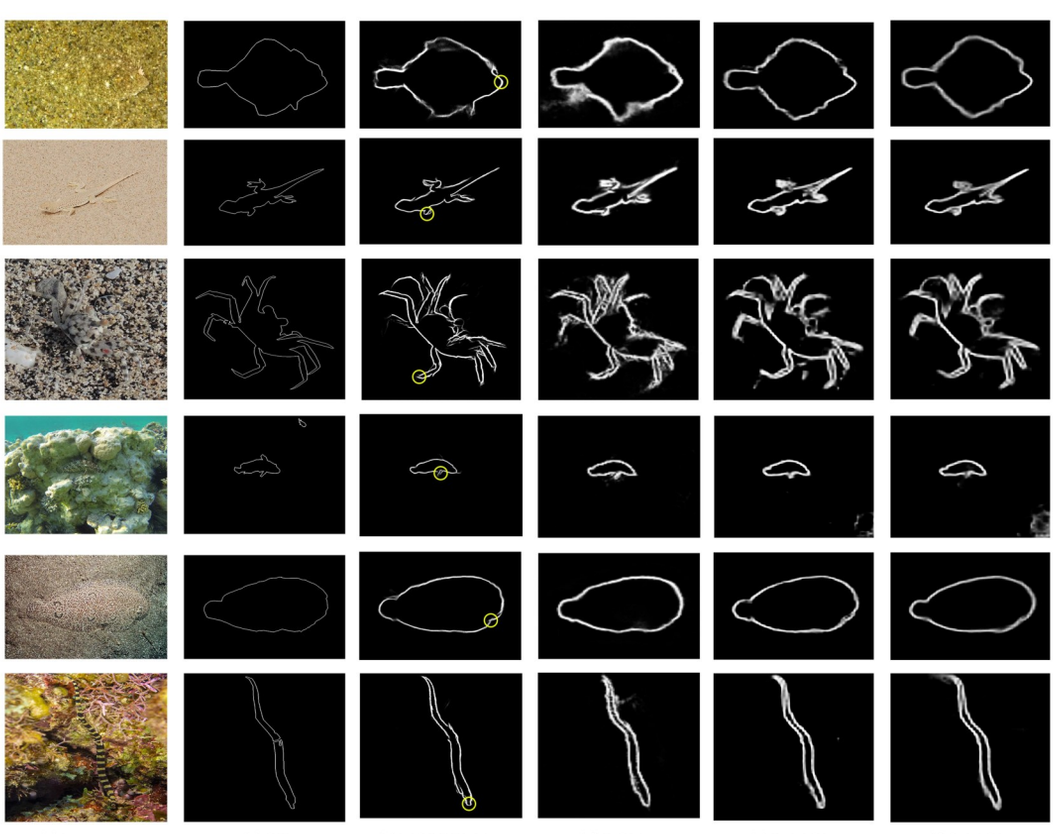



**图 7　MSPEFI-Net与其他方法感知边缘提取的可视化结果**

**Fig. 7　Visualization results of perceptual edge extraction by MSPEFI-Net versus other methods**

络,但它仍然丢失了很多结构细节,导致预测中边界定位不佳。实验证明,MSPEFI-Net在物体相关边缘信息挖掘和隐藏物体预测方面性能更佳。

### 2.5 消融实验分析

为了验证MSPEFI-Net各个模块的有效性,本文在CAMO、COD10K、NC4K上进行了消融实验,结果如表2和表3所示,表3中,模型a为基线模型、模型b为Basic+MSP、模型c为Basic+MSP+EFI、模型d为Basic+MSP+EFI+GFA(本文方法)。MSPEFI-Net采用的基线模型a仅由编码器-解码器结构以及ASPP模块及1×1卷积组成,其中编码器采用Res2Net,解码器采用自顶向下的方法逐层集成特征。如表2和表3所示,在基线模型a上添加MSP,在CAMO、COD10K、NC4K数据集上的$S_\alpha$分别提升了0.2%、0.9%、0.7%,说明MSP有效感知了混合尺度信息,并进行局部特征信息增强;通过在变体模型b上添加EFI模块,模型b比模型c在CAMO数据集上的$F_\beta^w$、$E_\varphi$分别提升了1.1%、1.2%,在COD10K数据集上的$F_\beta^w$提升了0.5%;在NC4K数据集上各项指标也有明显提升,说明EFI模块可以有效整合边缘语义信息,提高模型检测能力。最后,本文通过在变体模型c的上添加GFA模块进行自上而下的全局特征融合,用高层特征引导低层特征以帮助过滤低层特征中干扰信息,在CAMO、COD10K、NC4K数据集上各项指标也有明显提升,验证了GFA在提高模型性能上发挥了重要作用。

**表2 基于CAMO、COD10K的消融实验结果**

**Table 2 Ablation study results on CAMO and COD10K**

| 模型 | 方法 | | | | CAMO | | | | | COD10K | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Basic | MSP | EFI | GFA | $S_\alpha\uparrow$ | $F_\beta^w\uparrow$ | MAE↓ | $F_\beta$ | $E_\varphi\uparrow$ | $S_\alpha\uparrow$ | $F_\beta^w\uparrow$ | MAE↓ | $F_\beta$ | $E_\varphi\uparrow$ |
| a | √ | | | | 0.855 | 0.813 | 0.049 | 0.847 | 0.917 | 0.858 | 0.774 | 0.024 | 0.809 | 0.933 |
| b | √ | √ | | | 0.857 | 0.813 | 0.051 | 0.846 | 0.916 | 0.866 | 0.780 | 0.023 | 0.810 | 0.932 |
| c | √ | √ | √ | | 0.865 | 0.822 | 0.049 | 0.852 | 0.927 | 0.866 | 0.784 | 0.023 | 0.814 | 0.937 |
| d | √ | √ | √ | √ | 0.867 | 0.825 | 0.046 | 0.854 | 0.929 | 0.867 | 0.784 | 0.022 | 0.813 | 0.935 |

**表3 基于NC4K的消融实验结果**

**Table 3 Ablation study results on NC4K**

| 模型 | 方法 | | | | NC4K | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Basic | MSP | EFI | GFA | $S_\alpha\uparrow$ | $F_\beta^w\uparrow$ | MAE↓ | $F_\beta$ | $E_\varphi\uparrow$ |
| a | √ | | | | 0.869 | 0.820 | 0.037 | 0.853 | 0.931 |
| b | √ | √ | | | 0.875 | 0.824 | 0.035 | 0.853 | 0.932 |
| c | √ | √ | √ | | 0.878 | 0.827 | 0.035 | 0.855 | 0.934 |
| d | √ | √ | √ | √ | 0.879 | 0.830 | 0.034 | 0.857 | 0.936 |

为了验证目标边缘信息的有效性,本文比较表2中的模型b和模型c,可以看出,EFI可以进一步提高模型的检测性能,特别是在CAMO数据集上,$S_\alpha$、$F_\beta^w$、$E_\varphi$分别提升了0.9%、1.1%、1.2%。此外,图6中可视化结果预测图的边界十分清晰,进一步证明边缘信息的引入可以使得伪装目标检测的边缘细节更加清晰,显著提升了模型检测性能。

为了验证GFA的有效性,本文对其进行了不同变体的消融实验,本文设计了两种其他变体,并与本文最终的GFA进行比较,结果如表4所示,其中,模型a表示使用CA注意力代替MSCA注意力;模型b表示使用SA注意力代表MSCA注意力;模型c表示使用MSCA注意力。。

从表4中可以看出,使用MSCA注意力具有最优的效果。其中,模型a表示使用CA注意力替代MSCA注意力,虽然有一定的效果但仍与本文设计的模块具有一定的差距,在CAMO数据集上的$S_\alpha$比模型c降低了0.2%,$E_\varphi$比模型c低0.2%,在COD10K、NC4K数据集上也有明显的差距;模型b使用SA注意力各项指标也均有不同程度降幅,明显不如本文设计的模型c,经过以上消融实验验证了GFA的合理性和有效性。

**表4 基于CAMO、COD10K、NC4K对GFA的消融实验结果**

**Table 4 Ablation study results of GFA on CAOMO, COD10K and NC4K**

| 模型 | 方法 | CAMO | | | | | COD10K | | | | | NC4K | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $S_\alpha\uparrow$ | $F_\beta^w\uparrow$ | MAE↓ | $F_\beta$ | $E_\varphi\uparrow$ | $S_\alpha\uparrow$ | $F_\beta^w\uparrow$ | MAE↓ | $F_\beta$ | $E_\varphi\uparrow$ | $S_\alpha\uparrow$ | $F_\beta^w\uparrow$ | MAE↓ | $F_\beta$ | $E_\varphi\uparrow$ |
| a | W/CA | 0.865 | 0.823 | 0.047 | 0.853 | 0.927 | 0.867 | 0.785 | 0.022 | 0.811 | 0.933 | 0.879 | 0.829 | 0.034 | 0.855 | 0.932 |
| b | W/SA | 0.862 | 0.819 | 0.050 | 0.850 | 0.920 | 0.865 | 0.777 | 0.024 | 0.808 | 0.932 | 0.876 | 0.825 | 0.035 | 0.853 | 0.933 |
| c | W/MSCA | 0.867 | 0.825 | 0.046 | 0.854 | 0.929 | 0.867 | 0.784 | 0.022 | 0.813 | 0.935 | 0.879 | 0.830 | 0.034 | 0.857 | 0.936 |

## 3　结束语

本文针对多尺度信息利用不足和边缘特征信息匮乏的问题,设计了一种基于混合尺度感知与边缘特征交互的伪装目标检测网络(MSPEFI-Net),该网络充分利用混合尺度信息和边缘交互特征进行高效伪装目标检测。MSP模块用于分支特征共享和动态选择自适应扩大感受野,弥补混合尺度信息不足,提高了单层特征表示能力;EFI模块用于挖掘相邻层边缘细节特征,探索目标相关边缘语义特征信息,充分利用不同层的边缘特征信息;GFA模块用于聚合来自不同层的边缘语义信息与中心语义信息,实现全局上下文信息的聚合,从而实现简单、准确的伪装目标检测。本文在4个常用COD数据集上进行了广泛实验,实验结果表明,本文提出的MSPEFI-Net在大部分评估指标上优于12种最先进的算法,充分证明了其有效性。

**参考文献:**


[1] Fan D P, Ji G P, Sun G, et al. Camouflaged object detection[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020: 2777-2787.

[2] Sun Y, Chen G, Zhou T, et al. Context-aware cross-level fusion network for camouflaged object detection [C]//International Joint Conferences on Artificial Intelligence, Montreal, Canada, 2021: 1025-1031.

[3] Zhai Q, Li X, Yang F, et al. Mutual graph learning for camouflaged object detection[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021: 12997-13007.

[4] Pang Y, Zhao X, Xiang T Z, et al. Zoom in and out:a mixed-scale triplet network for camouflaged object detection[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, USA, 2022: 2160-2170.

[5] He C, Li K, Zhang Y, et al. Camouflaged object detection with feature decomposition and edge reconstruction[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, Canada, 2023: 22046-22055.

[6] Ji G P, Fan D P, Chou Y C, et al. Deep gradient learning for efficient camouflaged object detection[J]. Machine Intelligence Research, 2023, 20(1): 92-108.

[7] Song Y, Li X, Qi L. Camouflaged object detection with feature grafting and distractor aware[C]//2023 IEEE International Conference on Multimedia and Expo(ICME), Lisbon, Portugal, 2023: 2459-2464.

[8] Huang Z, Dai H, Xiang T Z, et al. Feature shrinkage pyramid for camouflaged object detection with transformers[C]//EEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, Canada, 2023: 5557-5566.

[9] Sun Y, Wang S, Chen C, et al. Boundary-guided camouflaged object detection[C]//International Joint Conference on Artificial Intelligence, Vienna, Austria, 2022:1335-1341.

[10] Dai Y, Gieseke F, Oehmcke S, et al. Attentional feature fusion[C]//IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, USA, 2021: 3560-3569.

[11] Fan D P, Ji G P, Qin X, et al. Cognitive vision inspired object segmentation metric and loss function[J]. Scientia Sinica Informationis, 2021, 51(9): 1475-1489.

[12] 王春华,李恩泽,肖敏. 多特征融合和孪生注意力网络的高分辨率遥感图像目标检测[J]. 吉林大学学报:工学版, 2024,54(1): 240-250.
Wang Chun-hua, Li En-ze, Xiao Min. Object detection in high-resolution remote sensing images based on multi-feature fusion and simeametric attention network[J]. Journal of Jilin University (Engineering and Technology Edition), 2024,54(1):240-250.

[13] Mei H, Ji G P, Wei Z, et al. Camouflaged object segmentation with distraction mining[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021:8772-8781.

[14] Fan D P, Ji G P, Cheng M M, et al. Concealed object detection[J]. IEEE transactions on pattern analysis and machine intelligence, 2021, 44(10): 6024-6042.

[15] Lyu Y, Zhang H, Li Y, et al. Uedg:uncertainty-edge dual guided camouflage object detection[J]. IEEE Transactions on Multimedia, 2023, 26: 4050-4060.

[16] Liang W, Wu J, Wu Y, et al. FINet: frequency injection network for lightweight camouflaged object detection[J]. IEEE Signal Processing Letters, 2024, 31: 526-530.