

# 针对红外弱小目标的多尺度级联融合分割网络

杨馨雨, 杨晓梅, 方璇

(四川大学 电气工程学院, 四川 成都 610065)

**摘要:** 红外弱小目标低信噪比、模糊形态等特点使得复杂背景下的针对此类目标的检测任务研究仍存在一定挑战。为弥补现有方法常见的目标信息丢失、与背景相互混淆等问题,提出了一种创新性的多尺度级联融合目标分割网络(MSCFNet)。具体而言, MSCFNet 通过不同尺度特征间的多次信息交互,最大限度保留小目标。同时,设计了一种特征增强模块,从全局和局部两个层面有效提取并整合信息,增强目标表征能力并抑制背景杂波。实验结果证明, MSCFNet 在多种复杂背景环境中能够有效地分割目标,并在两个公开红外弱小目标数据集上展现出更为优秀的分割性能。

**关键词:** 红外弱小目标分割; 红外图像; 深度学习; 特征增强

**中图分类号:** TP391.41; TN219 **文献标识码:** A **DOI:** 10.3969/j.issn.1001-5078.2025.03.018

## Multi-scale cascaded fusion network for infrared small target segmentation

YANG Xin-yu, YANG Xiao-mei, FANG Xuan

(College of Electrical Engineering, Sichuan University, Chengdu 610065, China)

**Abstract:** The low signal-to-noise ratio and fuzzy morphology of infrared (IR) small targets pose certain challenges in the research of such target segmentation tasks in complex backgrounds. To better separate small targets from clutter backgrounds, an innovative multi-scale cascaded fusion network (MSCFNet) is proposed. Specifically, MSCFNet preserves and utilizes small target information to the maximum extent through the multiple interaction between multi-scale features. At the same time, a feature enhancement module is designed to effectively extract and integrate target information from global semantic and local context, improving the discriminability of targets and complex backgrounds. The experimental results prove that MSCFNet can effectively segment IR small targets in various complex environments and exhibits better performance on two publicly available IR small target segmentation datasets.

**Keywords:** infrared small target; object segmentation; deep neural networks; feature enhancement

### 1 引言

红外成像系统凭借相较于可见光、雷达系统的高灵敏度、出色适应性和实时性等优势,在军事侦

察、工业检测等领域中展现出广阔的应用前景<sup>[1]</sup>。然而红外弱小目标存在低信噪比、弱对比度以及模糊形态等特点<sup>[2]</sup>。因此,在复杂背景环境中实现精

**作者简介:** 杨馨雨(1999-),女,硕士研究生,主研领域为红外弱小目标检测、图像处理、深度学习。  
E-mail: 3422821632qq.com

**通讯作者:** 杨晓梅(1973-),女,博士,副教授,主研领域为电能质量数据智能分析,深度学习、人工智能及相关专业。  
E-mail: yangxiaomei@scu.edu.cn

**收稿日期:** 2024-06-03; **修订日期:** 2024-07-30

准的红外弱小目标分割算法仍是一项具有挑战性的任务<sup>[3]</sup>。

早期红外弱小目标分割 (Infrared Small Target Segmentation, ISTS) 算法常采用模型驱动方法<sup>[4]</sup>。这些方法通常需要依据先验知识来设计特定的特征提取器和检测器,或者根据目标物体的物理特性手动调整算法参数。这导致模型驱动方法严重依赖于设计者的先验知识和专业经验,难以应对复杂多变的实际场景<sup>[5]</sup>。

近年来,随着计算机技术的快速发展,基于深度学习的数据驱动目标分割算法逐渐成为 ISTS 领域研究的主流趋势。数据驱动方法可以从大量红外图像中自主学习目标特征<sup>[4]</sup>,展现出较高的泛化能力。

Wang 等人<sup>[6]</sup>基于 GAN 提出了 MDvsFA (Miss Detection vs. False Alarm) 模型,将红外弱小目标分割问题拆解为漏检和虚警两个子任务,通过对抗性学习联合优化实现任务平衡,从而提升分割性能。Dai 等人<sup>[7]</sup>提出非对称上下文机制 (Asymmetric Contextual Modulation, ACM) 网络,利用自上而下的全局上下文反馈和自下而上的逐点通道注意力机制,实现高层与低层特征间的信息交换,提升特征表示能力;Zhang 等人<sup>[8]</sup>提出一种提出了称为注意力引导金字塔上下文网络 (Attention-Guided Pyramid Context Network) 的数据驱动方法,通过感知特定尺度的块内和块之间的像素相关性保留更多小目标细节。Wu 等人<sup>[9]</sup>提出了一种简单有效的 UIU-Net (U-Net in U-Net),实现目标特征的多层次、多尺度特征提取与学习。Wang 等人<sup>[10]</sup>结合区域候选网络 (Region Proposal Network, RPN)<sup>[11]</sup>和 Transformer 编码器<sup>[12]</sup>对预选框中的像素注意力进行建模,搭建了一种由粗略到细致的内部注意感知网络 (Interior Attention-Aware Network, IANet)。

尽管近年来的 ISTS 算法取得一定程度上的进展,但仍然存在以下局限性:

(1) 现有数据驱动算法为实现强大的特征建模能力,大多仅集中于处理高层的高级语义表示。然而,在提取高级语义特征时常伴随着局部细节的丢失,小尺寸目标很容易被复杂的环境所淹没。因此,纯粹以高层的高级语义为中心的表示不足以完成可靠的分割任务。

(2) 在实际应用场景 (建筑、树林、云雾等等)

中,多数网络主要依靠图像中每个像素的局部上下文信息对比度进行分割,难以有效保存和突出目标的细微细节。

为解决上述问题,本文提出了一种名为多尺度级联融合网络 (Multi-Scale Cascaded Fusion Network, MSCFNet) 的数据驱动算法。具体来说,针对红外小目标缺乏内在信息的特点,MSCFNet 创新性设计了一种语义和上下文信息增强 (Semantic-Context Enhancement, SCE) 模块。SCE 模块从全局语义和局部上下文信息两个方面保留弱小目标关键信息并抑制背景杂波。此外,我们在对高层特征上采样过程中增设多尺度特征之间的多次融合,使 SCE 模块维护的目标信息得到充分利用。同时,为进一步提升 MSCFNet 的分割性能,我们添加低层特征的级联通道,将来目标细节特征与来自高层的高级语义特征拼接起来,获得包含整体场景多层次理解的预测结果,从而应对实际场景中存在的极端模糊情况。

## 2 网络框架

本文提出的 MSCFNet 整体框架如图 1 所示,包括四个主要部分:特征提取、特征增强、多尺度特征融合以及分类头。

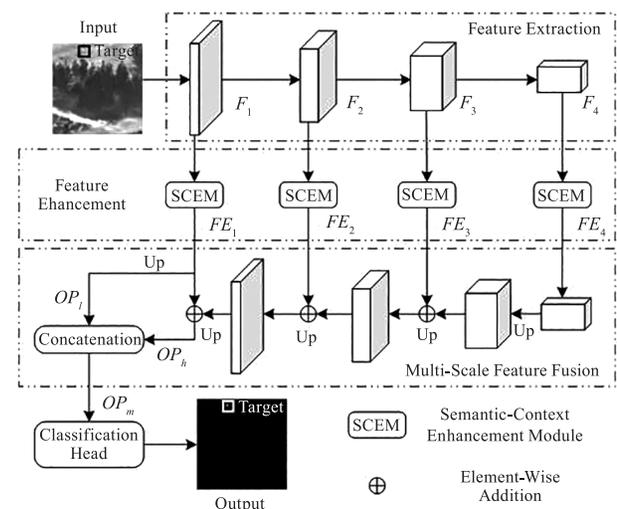


图 1 多尺度级联融合网络整体框架示意图

Fig. 1 Overall architecture of MSCFNet

### 2.1 特征提取

在特征提取阶段中,我们考虑到 ResNet 系列网络<sup>[13]</sup>高检测精度以及在拟合分布函数时的出色表现,采取 ResNet50 作为输入图像特征提取网络的基本骨架。并且由于小尺寸目标特征信息会随着网络层数的增加而损失,故我们只通过四次下采样操作

输出低层、两个中间层、高层,总计四个分层特征  $F_i (i = 1, 2, 3, 4)$ 。

设定输入图像为  $I \in R^{C \times H \times W}$ , 其中  $C, H$  和  $W$  分别表示  $I$  的通道数量、高度和宽度。在红外图像特征提取过程中,  $F_i$  的特征高度  $H_i$  和宽度  $W_i$  随着网络加深逐渐减小为  $H/2^i$  和  $W/2^i$ , 同时通道数  $C_i$  增加。

## 2.2 特征增强

MSCFNet 对于提取的多层特征  $F_i$  的增强操作由子模块 SCE 模块实现。该模块由三个分支组成, 如图 2 所示, 分别为全局语义分支、局部上下文分支以及残差连接分支, 三个部分共同作用充分提取并增强输入  $F_i$  中的弱小目标信息。

### (1) 全局语义分支

对于输入  $F_i \in R^{C_i \times H_i \times W_i}$ , 考虑到不同尺度下的特征层中弱小目标的感受野存在差异, 因此, 如图 2 所示, 全局语义分支由具有不同膨胀系数  $r$  的三层膨胀卷积<sup>[14]</sup>组成, 在不损失分辨率或覆盖范围的前提, 扩张网络感受野。

该分支输出的全局语义特征  $G_i \in R^{1/2 C_i \times H_i \times W_i}$  计算过程如下:

$$G_i = DC_{3 \times 3, r=3}(DC_{3 \times 3, r=2}(DC_{3 \times 3, r=1}(F_i))) \quad (1)$$

其中,  $DC_{3 \times 3, r(\cdot)}$  表示卷积核大小为  $3 \times 3$ ; 膨胀系数为  $r$  的膨胀卷积。

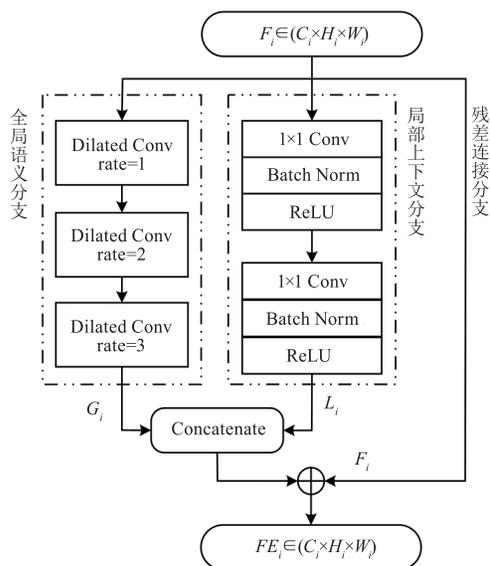


图 2 语义和上下文信息增强模块示意图

Fig. 2 Illustration of SCE module

### (2) 局部上下文分支

我们可以充分利用目标周围的局部区域上

下文信息用于提高模型的分割精度, 因此在 SCE 模块中增加局部上下文分支。

在过程中通过两次逐像素卷积提取目标局部信息并将  $F_i$  的通道数调整为  $1/2 C_i$ , 在减少计算量的同时, 关注局部区域内的细节, 捕捉目标边缘、轮廓等细节特征。局部上下文特征图  $L_i$  的计算如下所示:

$$L_i = \delta(BN(Conv_2(\delta(BN(Conv_1(F_i)))))) \quad (2)$$

式中,  $\delta, BN$  表示 ReLU 激活函数以及批量归一化操作;  $Conv_1$  和  $Conv_2$  表示卷积核大小为  $1 \times 1$  的像素级卷积。

### (3) 语义和上下文信息增强 (Semantic-Context Enhancement, SCE)

通过集成局部和全局两个层面的信息, 并结合第三条残差连接支路, 我们可以获得最终与每一个输入  $F_i$  维度相同的, 总计四组增强特征  $FE_i \in R^{C_i \times H_i \times W_i}$ , 如下式所示:

$$FE_i = [G_i, L_i] \oplus F_i \quad (3)$$

式中,  $[ \cdot ]$  代表沿通道维度的特征拼接操作;  $\oplus$  代表逐元素加法。

## 2.3 多尺度特征融合

前面提到过纯粹以高层的高级语义为中心的表示不足以完成可靠的分割任务, 因此我们在对高层增强特征  $FE_4$  上采样的过程中, 设计了对应的多尺度特征融合机制和低层特征级联通道。

具体来说, 如图 1 所示, 我们在  $FE_4$  的三个上采样结束节点增加融合路径, 将经 SCE 模块增强后的  $FE_1, FE_2, FE_3$  分别与具有相同维度的上采样特征图进行逐元素相加。通过多次融合操作, MSCFNet 可以更好地融合来自不同尺度的弱小目标信息, 获取最终的多尺度聚合特征  $Op_h \in R^{32 \times H_i \times W_i}$ , 表达式如下:

$$Op_h = up(up(up(up(FE_4) \oplus FE_3) \oplus FE_2) \oplus FE_1) \quad (4)$$

其中,  $up(\cdot)$  表示双线性插值上采样操作。

增设的级联通道将来自于浅层网络的特征  $FE_1$  直接上采样得到维度与  $Op_h$  相同, 并包含丰富目标细节特征的  $Op_l = up(FE_1)$ 。我们将二者沿通道维度拼接, 获得最终包含低层细节以及高级语义的多尺度聚合特征  $Op_m \in R^{64 \times H_i \times W_i}$ , 计算过程如下:

$$Op_m = [Op_h, Op_l] \quad (5)$$

最后,  $Op_m$  被输入到后续分类头中生成分析的红外图像  $I$  对应概率预测图。

### 3 实验

在本节中我们将验证 MSCFNet 及其组成成分的有效性,并详细介绍所有的实验设置,包括使用的两个主流用于检测 ISTS 算法性能的公开数据集、评估标准、超参数设置细节。

#### 3.1 实验设置

本文所提出的 MSCFNet 和基于数据驱动的 5 种比较网络均是基于 Pytorch 实现的 ISTS 算法。在对比实验中,除模型驱动算法以外,其余每个网络在两个数据集上进行训练次数分别为 20、100 的两次实验,消融实验则均设定为 100。优化器选择随机梯度下降 (SGD), 动量设定为 0.9, 权重衰减率为  $1.0 \times 10^{-4}$ 。初始学习率设置为 0.05, 采取 Poly 学习率调整策略并使用 Soft IoU<sup>[4]</sup> 损失函数进行损失值计算。

##### 3.1.1 数据集

我们使用 MDvsFA cGAN<sup>[6]</sup> 和 NUDT-SIRST<sup>[15]</sup> 两个公开红外数据集进行所有的试验,后续实验中分别缩写为 cGAN 和 NUDT。两个数据集中的样本尺寸、用于训练和测试的数量如表 1 所示。

表 1 数据集详细信息

Tab. 1 Details of datasets

数据集	图片大小	样本数量	
		训练	测试
cGAN <sup>[6]</sup>	128 × 128	8978	1000
NUDT <sup>[15]</sup>	256 × 256	1200	127

##### 3.1.2 评估指标

我们采取语义分割中常用的  $F_m$ 、平均交并比 (Mean Intersection over Union, mIoU) 作为所有方法的目标定位能力以及分割准确度的评估指标,具体计算式如下所示:

$$\begin{cases} pred = \frac{TP}{TP + FP} \\ rec = \frac{TP}{TP + FN} \end{cases} \quad (6)$$

$$\begin{cases} mIoU = \frac{TP}{TP + FP + FN} \\ F_m = \frac{2 \times pred \times rec}{pred + rec} \end{cases} \quad (7)$$

式中,  $pred$  代表被划分为正样本的像素中真正为正样本的比例;  $rec$  代表模型成功预测到的正例像素数与实际正例像素数之比; 二者的调和平均数值就是  $F_m$  指标;  $mIoU$  则用来衡量预测地面实况掩码和实际地面实况之间的分割相似度。

除上述两个常用指标外,我们还使用了接收者操作特征曲线 (Receiver Operating Characteristic Curve, ROC) 以及其对应的量化指标 (Area Under Curve, AUC) 作为度量标准。ROC 曲线用于衡量分类算法在不同阈值下的性能表现,越接近左上角坐标 (0,1) 代表模型性能越好。AUC 则是 ROC 曲线下方面积的量化数值,越接近于 1 越好。

#### 3.2 对比试验

在本节中,我们以上节介绍的指标参数为评估准则,对比 MSCFNet 与 7 种 ISTS 代表性算法的性能,其中包括 2 种传统模型驱动算法: IPI<sup>[16]</sup>、GST<sup>[17]</sup>; 5 种数据驱动算法: ENet<sup>[18]</sup>、BiSeNet<sup>[19]</sup>、DFANet<sup>[20]</sup>、ACM<sup>[7]</sup>、LW-IRST<sup>[21]</sup>。数据驱动算法训练过程中的具体设置参照 3.1 节。

##### 3.2.1 数值指标

表 2、3 展示了在 cGAN、NUDT 两个数据集上,所有算法模型的  $F_m$  (%)、mIoU (%)、AUC (%) 三种数值指标评估结果 (粗体字为最好结果,第二为下方添加下划线的数值)。在大多数情况下, MSCFNet 在两大类驱动算法: 2 种模型驱动与 6 种数据驱动,均展现出了显著的优势。

特别是在样本数量更少、背景更复杂的 NUDT 数据集的表现上,本文提出的 MSCFNet 相比其他七种算法模型展现出更优越的分割性能指标结果。特别值得注意的是,我们的网络可以在较少网络训练次数的情况下更高效解决 ISTS 任务。

为了进一步比较算法性能,图 3 中展示了 8 种算法时所对应 ROC 曲线 (其中,数据驱动算法训练次数为 20)。根据曲线走势来看,除图 3. (b) 中的 BiSeNet 曲线外,代表 MSCFNet 的曲线是所有算法中更为接近左上角坐标 (0,1)。这表明我们的方法可以在多变阈值下准确定位目标,展现更为精确的目标分割效果。

表 2 不同网络训练次数为 20 时的  $F_m(\%)$ 、mIoU( $\%$ )、AUC( $\%$ ) 数值  
 Tab. 2  $F_m(\%)$ 、mIoU( $\%$ )、AUC( $\%$ ) of different methods at training epochs of 20

Method	cGAN			NUDT		
	$F_m$	mIoU	AUC	$F_m$	mIoU	AUC
IPI	4.40	2.25	65.26	1.17	0.59	77.02
GST	17.09	9.34	56.38	18.41	10.14	56.29
DFANet	1.14	0.57	49.46	0.26	0.13	56.40
ENet	88.17	78.84	<u>94.34</u>	0.17	0.09	66.08
BiSeNet	77.25	62.93	93.79	0.14	0.07	51.93
ACM-UNet	<u>88.30</u>	<u>79.05</u>	93.29	<b>66.00</b>	<b>49.25</b>	<u>87.98</u>
LW-IRST	88.25	78.97	5.24	0.65	0.33	<b>88.79</b>
MSCFNet	<b>91.31</b>	<b>84.00</b>	<b>95.21</b>	<u>58.75</u>	<u>41.59</u>	77.49

表 3 不同网络训练次数为 100 时的  $F_m(\%)$ 、mIoU( $\%$ )、AUC( $\%$ ) 数值  
 Tab. 3  $F_m(\%)$ 、mIoU( $\%$ )、AUC( $\%$ ) of different methods at training epochs of 100

Method	cGAN			NUDT		
	$F_m$	mIoU	AUC	$F_m$	mIoU	AUC
IPI	4.40	2.25	65.26	1.17	0.59	77.02
GST	17.09	9.34	56.38	18.41	10.14	56.29
DFANet	2.92	1.48	67.45	0.83	0.42	55.06
ENet	89.49	80.99	94.64	76.62	62.10	90.49
BiSeNet	78.41	64.49	<b>96.28</b>	13.22	7.08	<b>99.67</b>
ACM-UNet	90.48	82.62	93.86	81.79	69.19	92.63
LW-IRST	<u>90.96</u>	<u>83.41</u>	94.88	<u>87.87</u>	<u>78.36</u>	50.05
MSCFNet	<b>92.72</b>	<b>86.42</b>	<u>95.78</u>	<b>90.83</b>	<b>83.20</b>	<u>95.75</u>

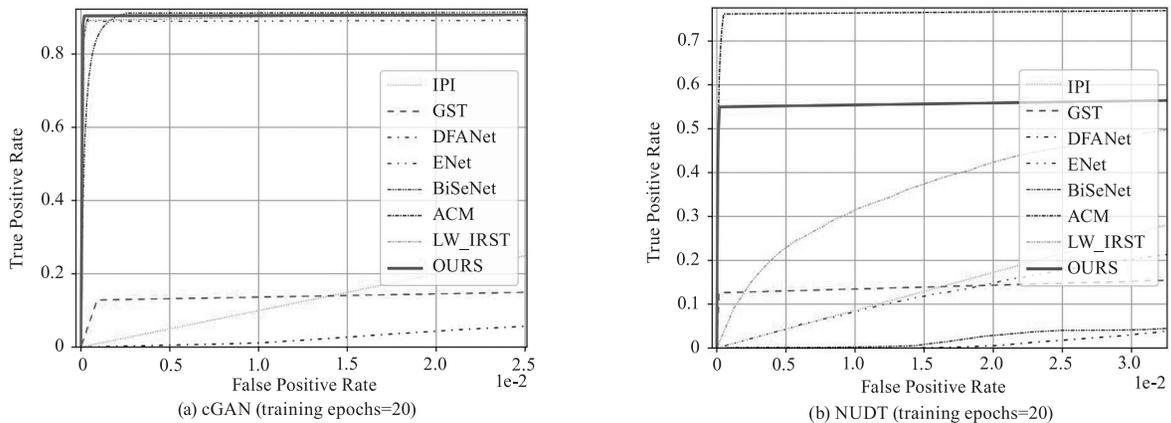


图 3 MSCFNet 和对比网络的 ROC 曲线图(粗实线代表 MSCFNet)

Fig. ROC curves of MSCFNet and other methods(The dashed line represents MSCFNet)

### 3.2.2 可视化指标

除了数值指标的量化比较,我们还利用了可视化结果更为直观地对比 MSCFNet 和其他对比算法之间的性能差异。图 4 和图 5 分别展示了 8 种算法框架在两个数据集含有的 6 种代表性场景真实分割结果(方框突出显示真实目标范围,并被放大到角落;圆圈表示误报),其中白色粗线方框突

出显示真实存在目标的范围,并被放大到角落;白色圆圈表示误报。

从可视化结果中我们可以观察到, MSCFNet 几乎不会出现错误预测情况,偶尔会存在一些遗漏目标的情况(即白色圆圈),但是通过增加训练次数至 100 次,即可实现对目标的准确预测。

综合数值与可视化两种指标来看,我们提出的

网络不仅可以在复杂场景中具有更好的分割性能,目标边缘轮廓分割相较于其他网络来说更精细、准确,这证实了 MSCFNet 能够有效突出目标并抑制背景杂波。

### 3.3 消融实验

在本节中我们通过三组消融试验在 cGAN 和 NUDT 两个数据集上证明,结果如表所示。每个消融实验对于网络的改动只存在一个,严格遵循控制变量法原理,具体描述如下:

1) 验证 SCE 模块的作用:我们去除了图 1 中网络的第二个特征增强模块,后续多尺度特征融合部分所有输入由增强特征  $FE_i$  变为原始分层特征  $F_i$ 。

2) 验证低层特征与高层特征拼接的作用:我们去除级联通道,使输入最后分类头的特征图由  $Op_m$  变为  $Op_h$ 。

3) 验证增加的多次融合线路的作用:在图 1 中的第三部分中我们去除了所有的融合路径,即进行拼接操作的高层特征由公式中的  $Op_h$  变为  $F_4$  经简单上采样操作之后的相同维度特征图。

结果如表 4 所示(粗体字为最好数值,第二为下方添加下划线的数值),通过对比三组消融实验与 MSCFNet 的数据结果,我们证明了三个部分的设置对于我们完整网络实现更精确和鲁棒的性能均有正面影响。

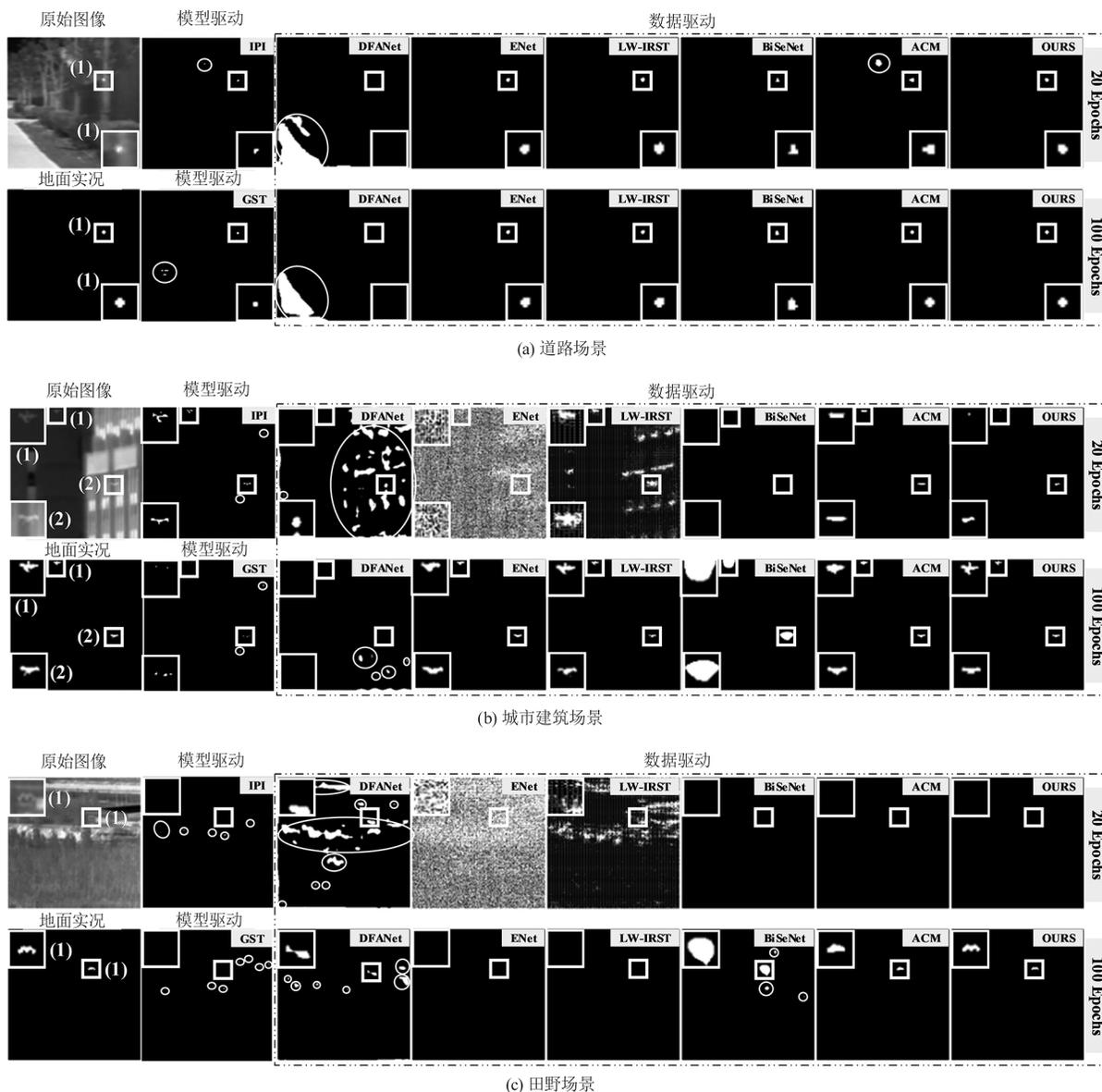


图 4 不同典型场景下的分割结果

Fig. 4 Segmentation results in typical scenes

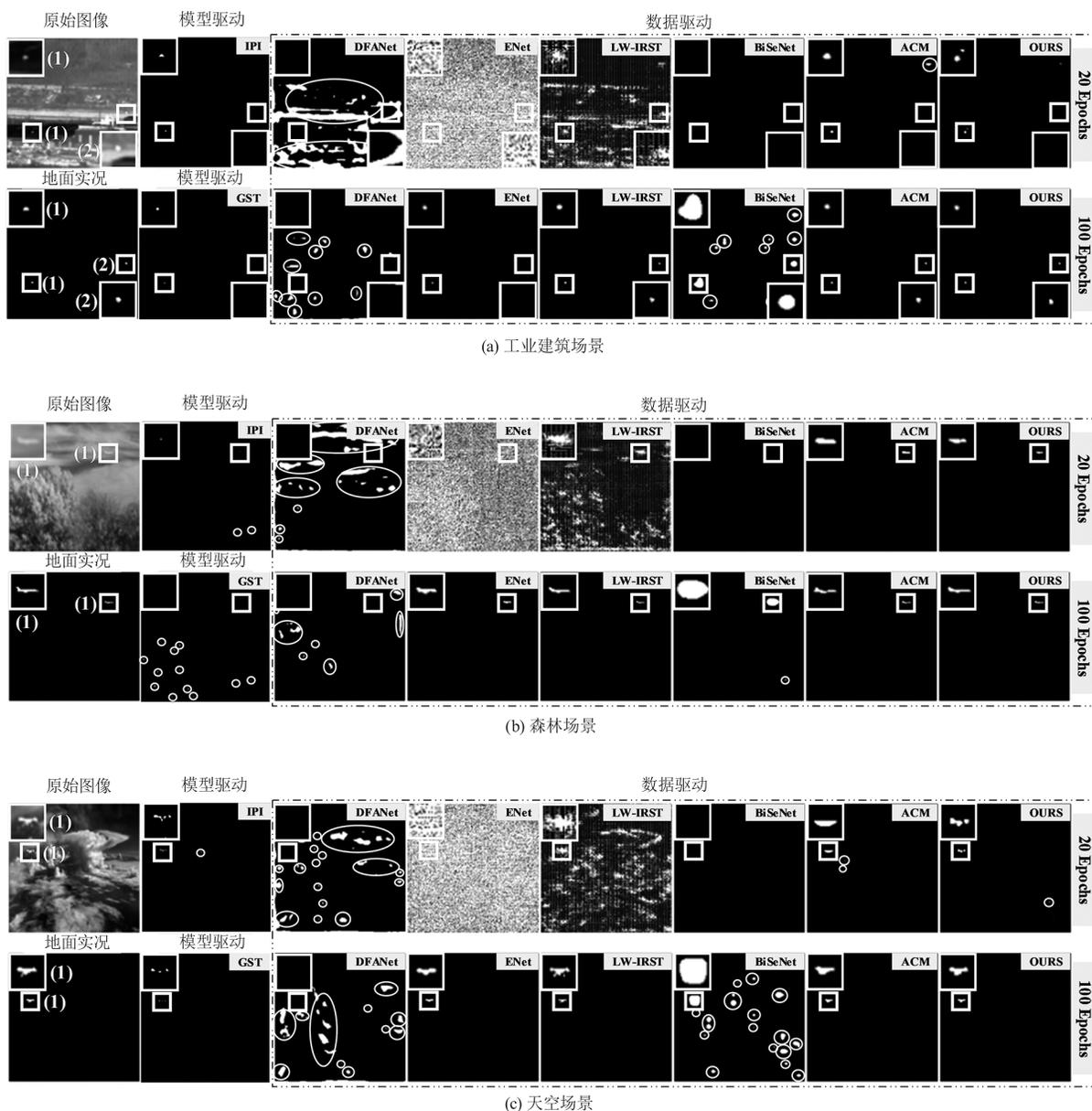


图 5 不同典型场景下的分割结果

Fig. 5 Segmentation results in typical scenes

表 4 训练次数为 100 情况下的消融实验结果

Tab. 4 The results of the ablation experiment

消融实验	cGAN		NUDT	
	$F_m/\%$	mIoU/%	$F_m/\%$	mIoU/%
①	92.44	85.94	87.36	77.55
②	92.05	85.27	00.00	00.00
③	84.08	72.53	76.75	62.27
MSCFNet	92.72	86.42	90.83	83.20

#### 4 结 语

本文针对复杂背景下的红外弱小目标分割任务,提出了一种多尺度级联融合分割网络(MSCF-

Net)。针对于深层网络中目标信息丢失、背景杂波干扰等问题,设计了 SCE 模块从局部上下文和全局语义信息两个层面上增强多层特征,有效提高了目标与背景像素的可分离性。此外,我们通过不同尺度特征之间融合与拼接,生成聚合低层细节以及高层语义的多尺度特征图,实现在上采样阶段中,最大程度保留并利用每一个增强特征中的目标信息。通过在两个公开红外弱小目标分割数据集上与其他代表性算法框架的对比试验证明了 MSCFNet 的优越性能与强鲁棒性;充分的消融实验结果证明了我们提出的方法各个组成成分的有效性。

## 参考文献:

- [1] Li Wenbo, Wang Qi, Gao Shang. A review of infrared small target detection algorithms based on deep learning [J]. *Laser & Infrared*, 2023, 53(10): 1476 – 1484. (in Chinese)  
李文博, 王琦, 高尚. 基于深度学习的红外小目标检测算法综述[J]. *激光与红外*, 2023, 53(10): 1476 – 1484.
- [2] Gu Yu, Zhang Hongyu, Peng Dongliang. Improved real-time infrared small target detection based on YOLOv5s [J]. *Laser & Infrared*, 2024, 54(2): 281 – 288. (in Chinese)  
谷雨, 张宏宇, 彭冬亮. 基于 YOLOv5s 的改进实时红外小目标检测[J]. *激光与红外*, 2024, 54(2): 281 – 288.
- [3] Zhao M, Li W, Li L, et al. Single-frame infrared small-target detection: a survey [J]. *IEEE Geoscience and Remote Sensing Magazine*, 2022, 10(2): 87 – 119.
- [4] Kou R, Wang C, Peng Z, et al. Infrared small target segmentation networks: a survey [J]. *Pattern Recognition*, 2023, 143: 109788.
- [5] Liu Zheng, Yang Dezhen, Li Jiangyong. A review of infrared single-frame dim and small target detection algorithms [J]. *Laser & Infrared*, 2022, 52(2): 154 – 162. (in Chinese)  
刘征, 杨德振, 李江勇, 等. 红外单帧弱小目标检测算法研究综述[J]. *激光与红外*, 2022, 52(2): 154 – 162.
- [6] Wang H, Zhou L, Wang L. Miss Detection vs. False alarm: adversarial learning for small object segmentation in infrared images [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 8508 – 8517.
- [7] Dai Y, Wu Y, Zhou F, et al. Asymmetric contextual modulation for infrared small target detection [C]//2021 IEEE Winter Conference on Applications of Computer Vision (WACV), 2021: 949 – 958.
- [8] Zhang T, Li L, Cao S, et al. Attention-guided pyramid context networks for detecting infrared small target under complex background [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2023, 59(4): 4250 – 4261.
- [9] Wu X, Hong D, Chanussot J. UIU-Net: U-Net in U-net for infrared small object detection [J]. *IEEE Transactions on Image Processing*, 2023, 32: 364 – 376.
- [10] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [J]. *arXiv*, 2017. arXiv: 1706.03762.
- [11] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137 – 1149.
- [12] Wang K, Du S, Liu C, et al. Interior attention-aware network for infrared small target detection [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1 – 13.
- [13] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770 – 778.
- [14] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions [C]//ICLR. 2016.
- [15] Li B, Xiao C, Wang L, et al. Dense nested attention network for infrared small target detection [J]. *IEEE Transactions on Image Processing*, 2023, 32: 1745 – 1758.
- [16] Gao C, Meng D, Yang Y, et al. Infrared patch-image model for small target detection in a single image [J]. *IEEE Transactions on Image Processing*, 2013, 22(12): 4996 – 5009.
- [17] Gao Ch Q, Tian J W, Wang P. Generalised-structure-tensor-based infrared small target detection [J]. *Electronics Letters*, 2008, 44(23): 1349.
- [18] Paszke A, Chaurasia A, Kim S, et al. ENet: a deep neural network architecture for real-time semantic segmentation [J]. *arXiv*, 2016. arXiv: 1606.02147.
- [19] Yu C, Wang J, Peng C, et al. BiSeNet: bilateral segmentation network for real-time semantic segmentation [J]. *arXiv*, 2018. arXiv: 1808.00897.
- [20] Li H, Xiong P, Fan H, et al. DFANet: deep feature aggregation for real-time semantic segmentation [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 9514 – 9523.
- [21] Kou R, Wang C, Yu Y, et al. LW-IRSTNet: lightweight infrared small target segmentation network and application deployment [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 1 – 13.