

改进多尺度 U 型网络的红外图像去模糊方法

张艳珠, 赵 赫, 刘义杰

(沈阳理工大学 自动化与电气工程学院, 沈阳 110159)

摘要: 针对热成像设备采集的红外图像纹理细节较粗导致去模糊效果不佳的问题, 在基于深度学习的可见光图像去模糊网络基础上, 提出了一种改进的适用于红外图像的单幅图像去模糊方法, 该方法以多尺度 U 型网络为基础进行改进。首先, 在编码模块中引入对偶注意力单元, 增强网络的特征表达能力; 其次, 将快速傅里叶变换嵌入到特征融合模块, 加强网络对于高低频信息的处理能力; 最后, 选择更优的激活函数和损失函数, 实现更好的信息流动并提升模型的鲁棒性。在红外数据集上进行测试, 并与原网络进行对比, 结果表明, 本文改进网络的图像细节恢复得更好, 峰值信噪比提升了 0.53 dB, 红外图像去模糊效果良好。

关键词: 红外图像; 深度学习; 图像去模糊; 对偶注意力单元; 快速傅里叶变换

中图分类号: TP391.41 **文献标志码:** A **DOI:** 10.3969/j.issn.1003-1251.2024.01.009

Infrared Image Deblurring Method Based on Improved Multi-scale U-shaped Network

ZHANG Yanzhu, ZHAO He, LIU Yijie

(Shenyang Ligong University, Shenyang 110159, China)

Abstract: For the poor deblurring effect due to coarse texture details of infrared images collected by thermal imaging equipment, an improved single-image deblurring method suitable for infrared images is proposed based on a deep learning-based visible light image deblurring network. The method is improved based on a multi-scale U-shaped network. First, a dual attention unit is introduced in the encoding module to enhance the feature representation ability of the network. Second, the fast Fourier transform is embedded into the feature fusion module to enhance the processing capability of the network for high and low frequency information. Finally, better activation function and loss function are selected to achieve better information flow and improve the robustness of the model. Testing on the infrared dataset and compared with the original network, the results show that the improved network in this study has achieved better restoration of image details, and the peak signal-to-noise ratio is 0.53 dB higher than the original network. The infrared image has a good deblurring effect.

Key words: infrared image; deep learning; image deblurring; dual attention unit; fast Fourier transform

近年来, 红外成像技术凭借其抗干扰能力强、可实现夜视成像等优点, 被广泛用于安防监控等

领域^[1-2]。由于拍摄期间设备的抖动或目标的快速移动会使图像存在模糊的伪影, 导致后续的识

别工作难度增大,从模糊图像中恢复出清晰图像是图像处理技术研究的重要内容。

图像去模糊方法大致可分为基于优化的去模糊方法和基于深度学习的去模糊方法。基于优化的方法将去模糊任务转变为逆滤波问题,如郝建坤等^[3]在模糊核已知情况下,采用去卷积方法恢复了清晰图像,华夏等^[4]在模糊核未知情况下,提出基于正则化的去模糊方法,仅通过模糊图像得到模糊核与清晰图像。基于优化的方法能够利用图像的先验信息较好地实现去模糊,但其只针对特定场景,在其他场景中表现不佳且处理十分耗时。基于深度学习的去模糊方法发展迅速,现已涌现出多种算法^[5-6]。Kupyn 等^[7]提出一种基于生成式对抗网络(GAN)的去模糊方法,并在其基础上进行了改进^[8],该方法能够较好地恢复图像的纹理细节,但稳定性不佳,模型训练过程中会出现参数震荡不收敛的情况。Wang 等^[9]和 Zamir 等^[10]均提出了基于 Transformer 的架构用于图像去模糊,受益于其多头自注意力机制,去模糊效果良好,但计算成本很高。Tao 等^[11]提出了一种共享参数的多尺度网络,能够有效处理各种类型的模糊,但处理速度较慢。基于深度学习的方法能够学习模糊图像和清晰图像之间的非线性映射,适合处理复杂的模糊图像,在可见光图像处理上显示了优异的性能。相较于可见光图像,红外图像纹理细节更粗,基于深度学习的去模糊方法对红外图像的细节信息恢复较差,应用效果不佳。

本文在多尺度 U 型网络(MIMO-UNet)^[12]的基础上提出一种适用于红外图像的单幅图像去模糊方法。在编码模块中引入对偶注意力单元(dual attention unit, DAU)^[13],使其在提取特征时抑制作用较弱的特征,并允许携带更多信息的特征继续传递;在特征融合模块中引入快速傅里叶变换(fast fourier transform, FFT),优化模块融合的多尺度特征信息;选择更优的激活函数和损失函数替换原网络使用的函数,实现更好的信息流动并提高模型泛化能力^[14]。

1 多尺度 U 型网络及其改进

1.1 多尺度 U 型网络

MIMO-UNet 主要由编码模块(encoder block, EB)、解码模块(decoder block, DB)、浅层卷积模块(shallow convolutional module, SCM)和非对称特征融合模块(asymmetric feature fusion,

AFF)组成,其结构如图 1 所示。

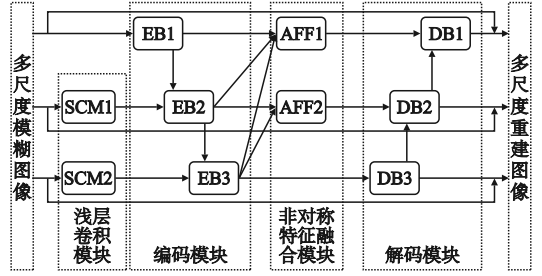


图 1 多尺度 U 型网络结构

Fig. 1 Multi-scale U-shaped network structure

编码模块将输入的多尺度图像信息进行特征提取,以捕获数据的关键特征;解码模块将特征信息转换回原始数据空间,以恢复多尺度清晰图像;浅层卷积模块将其他尺度的图像编码以便输入到网络;非对称特征融合模块合并跨编码器和解码器的多尺度信息流,以融合有较大感受野的深层特征和有丰富细节信息的浅层特征。

1.2 网络改进

正确学习清晰图像和模糊图像的差异能够有效提升网络的性能,两者之间最主要的差异是前者含有更多的细节信息即高频信息,后者含有更多的内容信息即低频信息。相较于可见光图像,红外图像纹理细节更粗、视觉效果更模糊,影响了网络对其高频信息的学习,虽然通过堆叠网络模块可以在一定程度上缓解该问题,但会增加计算复杂度。为此,在网络模块中引入 DAU 和 FFT,通过在空间域和频域上各单词首字母小写对优化的特征进行学习,提升网络的去模糊效果。本文改进网络的具体结构如图 2 所示。图中:Conv 表示卷积层;Dec 表示逆卷积层;RB 表示残差模块;AFFFTFF 表示非对称快速傅里叶变换特征融合模块;⊙表示拼接操作;⊕表示逐元素求和。

1.2.1 对偶注意力编码模块

DAU 能够在空间和通道两个维度上实现与特征张量进行信息共享,其结构如图 3 所示。图中:GAP 表示全局平均池化层;MAP 表示最大平均池化层;SA 表示空间注意力;CA 表示通道注意力;ReLU 为激活函数;Ⓢ表示 Sigmoid 激活函数。

DAU 主要由空间注意力(SA)和通道注意力(CA)两部分组成。SA 旨在利用特征图的空间依赖性生成空间注意力图,并用其重新校准输入的特征,自动捕获重要区域的特征。CA 则通过压缩和激励两个操作实现利用特征图通道间的关系:通过对全局空间信息进行压缩,在通道维度上进

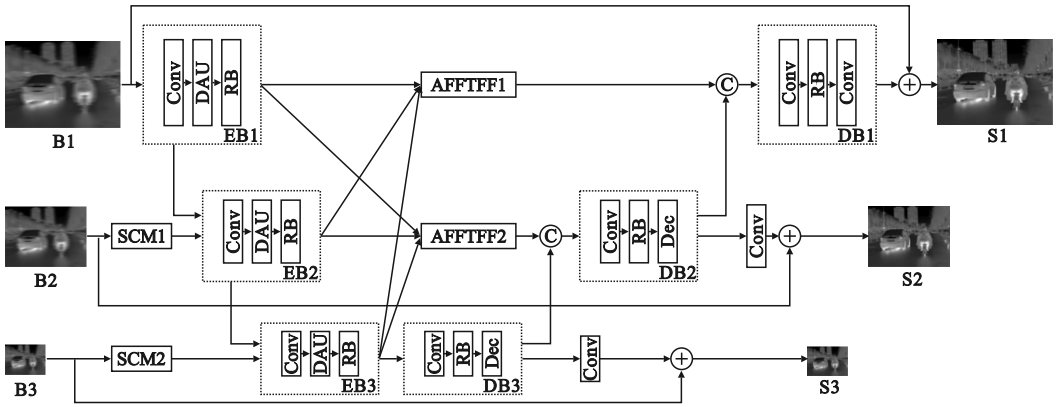


图 2 改进网络的结构

Fig. 2 Improved network structure

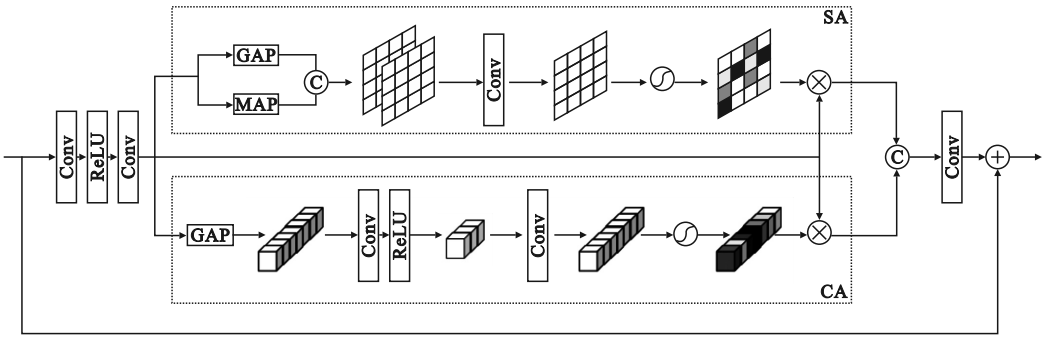


图 3 对偶注意力单元结构

Fig. 3 Dual attention unit structure

行特征学习,得到各个通道的重要性;通过激励对各个通道分配不同的权重。因此,DAU 可以抑制作用较弱的特征,允许带有更多信息的特征继续传递,引入 DAU 的编码模块可以更好地实现特征的学习和提取。

1.2.2 非对称快速傅里叶变换特征融合模块

较粗的纹理细节会影响卷积对特征的学习,将图像从空间域转换到处理较为容易的频率域能够在一定程度上解决该问题。FFT 能够将图像在空间域的细节信息和内容信息转成频率域的高频信息和低频信息,将 FFT 嵌入到特征融合模块中,可实现对图像高低频信息的学习。此外,输入的特征融合了多尺度图像编码后的特征,能够覆盖整个图像,更加轻松地捕捉模糊和清晰图像之间的差异,使得去模糊效果得到进一步提升。改进后的非对称快速傅里叶变换特征融合模块(AFFTFF)结构如图 4 所示。图中:Resize 表示缩放操作,用于调整特征图尺寸以便进行拼接;Inv FFT 表示快速傅里叶逆变换。

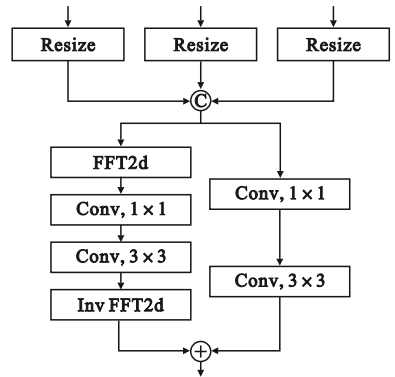


图 4 AFFTFF 结构

Fig. 4 AFFTFF structure

1.2.3 激活函数和损失函数

MIMO-UNet 的激活函数采用 ReLU 函数,ReLU 函数计算速度较快,但当函数值为 0 或为负值时,会出现参数不再更新的情况,导致模型的性能降低。Mish 激活函数能够弥补 ReLU 的不足。ReLU 和 Mish 激活函数的形式如式(1)和式(2)所示,两者对比如图 5 所示。

$$\text{ReLU}(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (1)$$

$$\text{Mish}(x) = x \tanh[\ln(e^x + 1)], x \in (-\infty, +\infty) \quad (2)$$

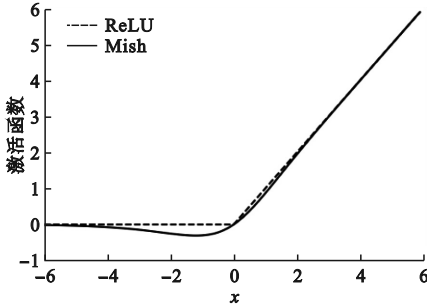


图5 激活函数对比图

Fig. 5 The comparison diagram of activation function

Mish 函数为平滑函数,能够避免 ReLU 函数的奇异点,其正值无边界避免了梯度饱和的情况,其负值较小能稳定网络的梯度流,使信息能够更好地在网络中流动。故本文使用 Mish 作为激活函数。

MIMO-UNet 的损失函数采用平均绝对误差 (MAE) 损失,又称 L1 损失。该损失函数不会出现梯度爆炸问题,但其在 $x=0$ 处不可导,且其不论损失值大小均具有相同梯度,可能会造成模型震荡不收敛,平滑的绝对误差损失 Smooth L1 则能解决该问题。MAE 损失(用 L_{MAE} 表示)和 Smooth L1 损失(用 L_{SmoothL1} 表示)计算公式如式(3)和式(4)所示,两者对比如图6所示。

$$L_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (3)$$

$$L_{\text{SmoothL1}} = \frac{1}{n} \sum_{i=1}^n \begin{cases} |\hat{y}_i - y_i| - \frac{1}{2}, & |\hat{y}_i - y_i| > 1 \\ \frac{1}{2}(\hat{y}_i - y_i)^2, & |\hat{y}_i - y_i| \leq 1 \end{cases} \quad (4)$$

式中: n 为像素总数量; \hat{y}_i 为估计值; y_i 为真实值。

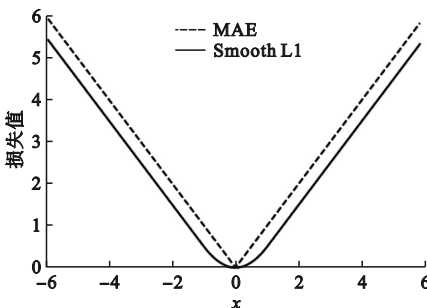


图6 损失函数对比图

Fig. 6 The comparison diagram of loss function

Smooth L1 为分段函数,当误差较大时,其梯度较大,但不会导致梯度爆炸问题;当误差较小时,其梯度足够小,更利于模型的收敛。因此,本文使用 Smooth L1 作为损失函数。

2 实验与结果分析

2.1 数据集处理

网络进行迭代训练需要使用成对的清晰-模糊图像,本文采用红外热像仪采集图像数据,并使用公开的红外图像数据集 FLIR 进行数据补充。由于同时获取清晰和模糊的图像较为困难,故将清晰图像进行模糊处理得到对应的模糊图像,如图7所示。产生模糊的主要原因有相机运动、物体运动以及相机和物体同时运动,使用 Kupyin 等^[7]提出的随机轨迹法模拟相机运动(图7第一行)、使用多帧平均法模拟物体运动(图7第二行)、使用运动退化模型模拟相机和物体同时运动(图7第三行),图7中从左到右分别对应清晰图像、轻度模糊图像和重度模糊图像。



图7 模拟模糊效果图

Fig. 7 The effect diagram of simulation blur

从 3 000 张红外图像中随机选择图像,并随机选择模糊方式和模糊程度,得到 3 000 组成对的清晰-模糊图像对,按 8:1:1 的比例分为训练集、验证集和测试集,图像的分辨率为 640×480 。

2.2 实验配置

具体实验配置如表 1 所示。

2.3 主观评价

为验证本文算法的有效性,使用 MIMO-UNet 和本文改进后的网络进行对比实验,测试结果如图8所示。

图8(a)和图8(b)分别为用于测试网络性能的清晰图像和模糊图像;图8(c)和图8(d)分别

表 1 实验配置

Table 1 Experimental configuration

软件或硬件名称	具体参数
操作系统	64 位 Ubuntu16.04LTS
CPU	Inter Xeon(R) E5-1603 v4, 2.1 GHz × 32
GPU	2 × Nvidia GeForce GTX 1080 Ti
内存	4.1 TB
编程语言	Python3.8
深度学习框架	Pytorch-GPU

为 MIMO-UNet 原网络和本文改进网络的测试结果,其中第一行为去模糊效果图,第二行为重建图像与清晰图像对比的残差图。由图 8 可以看出:MIMO-UNet 能够在一定程度上实现去模糊;本文改进的网络则能更好地恢复出图像的细节信息,如框中的黑色“書”字和背景以及框中的白色文字;改进网络的残差图包含的信息更少,其重建图像更加接近于清晰图像。因此,本文提出的改进网络能够有效实现红外图像去模糊。

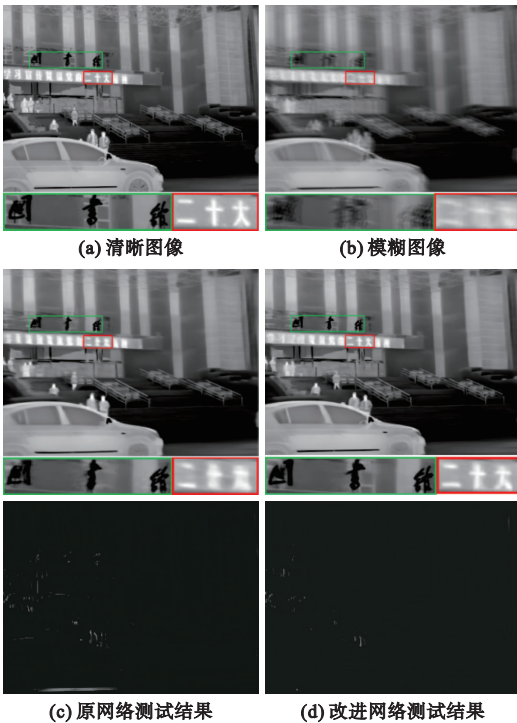


图 8 测试结果图

Fig. 8 The diagram of test results

2.4 客观评价

2.4.1 消融实验

为验证本文改进结构对网络性能的影响,在原网络基础上逐一加入改进的模块及函数,在相同数据集上对网络进行训练和测试,使用峰值信噪比 (PSNR) 和结构相似性 (SSIM) 两个常用的图像质量评估指标验证网络的去模糊性能,实验结果如表 2 所示。

表 2 改进网络结构的性能评估

Table 2 Improved network structure performance evaluation

DAU	FFT	Mish + Smooth LI	PSNR/dB	SSIM
—	—	—	31.79	0.86
—	—	✓	31.93	0.86
✓	—	—	32.13	0.87
—	✓	—	32.19	0.87
✓	✓	—	32.26	0.88
✓	✓	✓	32.32	0.88

由表 2 可以看出:加入 DAU 或加入 FFT 均能较大幅度地提升网络性能,采用更优的激活函数和损失函数能够在一定程度上提升网络性能;相较于原网络,本文改进网络的 PSNR 提升了 0.53 dB,由于红外图像本身视觉效果模糊,SSIM 提升较少,仅为 0.02。本文采用的改进模块及函数均有助于提升网络对红外图像的去模糊效果。

2.4.2 对比实验

在主观评价的基础上,采用 PSNR、SSIM、网络参数数量和运算时间 4 个指标对 MIMO-UNet、MIMO-UNet + 和本文改进网络的性能进行比较分析。实验结果如表 3 所示。

表 3 不同网络的性能评估

Table 3 Performance evaluation of different networks

网络	PSNR/dB	SSIM	参数量/MB	运算时间/s
MIMO-UNet ^[12]	31.79	0.86	6.80	0.016
MIMO-UNet + ^[12]	32.21	0.87	16.10	0.052
本文改进网络	32.32	0.88	7.29	0.018

由表 3 可以看出:MIMO-UNet 对于红外图像去模糊有一定的效果;MIMO-UNet + 通过进一步增加网络中的模块数量可提升去模糊效果,但也大大增加了计算量,使得运算时间成倍增加;本文改进后的网络在参数量少量增加、处理时间基本不变的情况下,更好地实现了红外图像去模糊。

3 结论

为更好地实现红外图像去模糊,本文基于 MIMO-UNet 进行改进,在编码模块中引入了 DAU,增强网络特征提取能力,在非对称特征融合模块中引入了 FFT,加强网络对于清晰图像和模糊图像之间差异的学习,同时优化了损失函数和激活函数。在采集的红外数据上进行训练和测试,消融实验结果表明,本文采用的改进模块及函数均有助于提升网络性能;对比实验结果表明,相

较于原网络,本文改进网络重建图像的细节更加清晰,PSNR 提升了 0.53 dB,SSIM 提升了 0.02。本文算法为红外图像去模糊任务提供了一种良好的解决方案。

参考文献(References):

- [1] 钱阳. 红外热成像技术在警用执法领域的应用[J]. 中国安防,2023(Z1):83-88.
QIAN Y. Application of infrared thermal imaging technology in the field of police law enforcement[J]. China Security & Protection,2023(Z1):83-88. (in Chinese)
- [2] 赵国顺. 红外热成像技术的应用延伸探析[J]. 中国安防,2023(Z1):94-97.
ZHAO G S. Analysis on the application extension of infrared thermal imaging technology[J]. China Security & Protection, 2023(Z1):94-97. (in Chinese)
- [3] 郝建坤,黄玮,刘军,等. 空间变化 PSF 非盲去卷积图像复原法综述[J]. 中国光学,2016,9(1):41-50.
HAO J K,HUANG W,LIU J,et al. Review of non-blind deconvolution image restoration based on spatially-varying PSF [J]. Chinese Optics,2016,9(1):41-50. (in Chinese)
- [4] 华夏,张悦芮,刘广,等. 视觉灵敏性滤波和自适应正则化图像盲去模糊[J]. 华中科技大学学报(自然科学版),2021,49(11):64-70.
HUA X,ZHANG Y R,LIU G,et al. Blind image deblurring based on visual sensitivity filtering and adaptive regularization [J]. Journal of Huazhong University of Science and Technology(Natural Science Edition),2021,49(11):64-70. (in Chinese)
- [5] SU J W,XU B Y,YIN H J. A survey of deep learning approaches to image restoration[J]. Neurocomputing,2022,487:46-65.
- [6] ZHANG K H,REN W Q,LUO W H,et al. Deep image deblurring:a survey[J]. International Journal of Computer Vision,2022,130(9):2103-2130.

- [7] KUPYN O,BUDZAN V,MYKHAILYCH M,et al. DeblurGAN: blind motion deblurring using conditional adversarial networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018:8183-8192.
- [8] KUPYN O,MARTYNIUK T,WU J R,et al. DeblurGAN-v2: deblurring (orders-of-magnitude) faster and better [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE,2020:8877-8886.
- [9] WANG Z D,CUN X D,BAO J M,et al. Uformer: a general U-shaped transformer for image restoration [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE,2022:17662-17672.
- [10] ZAMIR S W,ARORA A,KHAN S,et al. Restormer: efficient transformer for high-resolution image restoration [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE,2022:5718-5729.
- [11] TAO X,GAO H,SHEN X,et al. Scale-recurrent network for deep image deblurring [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE,2018:8174-8182.
- [12] CHO S J,JI S W,HONG J P,et al. Rethinking coarse-to-fine approach in single image deblurring [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, Canada: IEEE,2021:4621-4630.
- [13] ZAMIR S W,ARORA A,KHAN S,et al. Learning enriched features for real image restoration and enhancement [C]// Computer Vision: ECCV 2020 Workshops. Glasgow, UK: Springer,2020:492-511.
- [14] 王光博,陈亮. 改进关系网络的小样本图像分类方法[J]. 沈阳理工大学学报,2023,42(1):28-34.
WANG G B,CHEN L. Small sample image classification method based on improved relational network [J]. Journal of Shenyang Ligong University,2023,42(1):28-34. (in Chinese)

(责任编辑:宋颖韬)

(上接第54页)

- [11] 胡智翔,鲍胜利,徐传淇,等. 基于 DeepLabv3+ 的遥感建筑物语义分割算法[J]. 计算机应用,2021,41(S2):71-75.
HU Z X,BAO S L,XU C Q,et al. Semantic segmentation algorithm for remote sensing buildings based on DeepLabv3+ [J]. Journal of Computer Applications,2021,41(S2):71-75. (in Chinese)
- [12] 陈兵,贺晟,刘坚,等. 基于轻量化 DeepLab v3+ 网络的焊缝结构光图像分割[J]. 中国激光,2023,50(8):49-58.

CHEN B,HE S,LIU J,et al. Weld structured light image segmentation based on lightweight DeepLab v3+ network [J]. Chinese Journal of Lasers,2023,50(8):49-58. (in Chinese)

- [13] 徐聪,王丽. 基于改进 DeepLabv3+ 网络的图像语义分割方法[J]. 激光与光电子学进展,2021,58(16):1610008.
XU C,WANG L. Image semantic segmentation method based on improved DeepLabv3+ network [J]. Laser & Optoelectronics Progress,2021,58(16):1610008. (in Chinese)

(责任编辑:宋颖韬)