

文章编号 1004-924X(2024)23-3479-11

基于注意力机制的宽波段小目标实时去模糊

朱德燕^{1,2*}, 敖咏琪^{1,2}, 徐家一^{1,2}, 李承城^{1,2}, 张钰凡^{1,2}

(1. 南京航空航天大学 航天学院, 江苏 南京 210016;
2. 南京航空航天大学 空间光电探测与感知工信部重点实验室, 江苏 南京 210016)

摘要: 基于天基平台的远距离运动目标探测和识别面临着波段宽、图像模糊、目标尺寸小的挑战。针对天基平台对去模糊算法高实时性、泛化性和重建质量的要求, 提出了一种基于注意力机制的多尺度多阶段卷积神经网络方法。利用优化模块数量的轻量化多尺度多阶段网络作为架构保障实时性, 结合基于频域的自注意力机制求解器及前馈神经网络(Frequency domain based self-attention solver and discriminative FFN, F-D)和卷积注意力模块(Concolutional Block Attention Module, CBAM)捕捉空域和频域上的关键信息。实验结果表明: 去模糊方法的复原速率优于 34 frame/s, 耗时仅为典型算法的 1/3, 在红外小目标数据集上的 PSNR 大于 32 dB, SSIM 大于 0.87; 在可见光小目标数据集上的 PSNR 大于 17 dB, SSIM 大于 0.93。该算法在宽波段具有泛化能力并可有效恢复小目标轮廓和形状信息。

关键词: 计算机视觉; 图像去模糊; 注意力机制; 小目标

中图分类号: TP394.1; TH691.9 **文献标识码:** A **doi:** 10.37188/OPE.20243223.3479

Real-time deblurring of wideband small target based on attention mechanism

ZHU Deyan^{1,2*}, AO Yongqi^{1,2}, XU Jiayi^{1,2}, LI Chengcheng^{1,2}, ZHANG Yufan^{1,2}

(1. College of Astronautics, Nanjing University of Aeronautics and Astronautics,
Nanjing 210016, China;
2. Key Laboratory of Space Optoelectronic Detection and Perception, Nanjing University of
Aeronautics and Astronautics, Nanjing 210016, China)
* Corresponding author, E-mail: zdy_nuaa@163.com

Abstract: The detection and recognition of long-distance moving targets on space-based platforms face challenges such as wide bandwidth, blurry images, and small target sizes. To address these issues, a multi-scale, multi-stage convolutional neural network incorporating an attention mechanism is proposed, meeting the demands for high real-time performance, generalization, and deblurring quality. The approach employs a lightweight multi-scale, multi-stage network with an optimized module count to ensure real-time processing, integrating a frequency-domain-based self-attention solver, a discriminative FFN (F-D), and a convolutional block attention module (CBAM) to extract critical spatial and frequency domain information. The experimental results show a deblurring restoration rate exceeding 34 frame/s, with time con-

收稿日期: 2024-09-02; 修订日期: 2024-10-03.

基金项目: 国家自然科学基金资助项目 (No. 62105145)

sumption reduced to one-third of conventional methods. On the infrared small target dataset, the PSNR exceeds 32 dB and the SSIM surpasses 0.87, while on the visible light small target dataset, the PSNR exceeds 17 dB and the SSIM surpasses 0.93. The algorithm demonstrates strong generalization across wide-band scenarios, effectively restoring the contours and shapes of small targets.

Key words: computer vision; image deblurring; attention mechanism; small target

1 引言

天基探测平台对运动目标进行成像时,由于传输距离远、相对运动快、衍射和散焦等问题,目标所成一般为小尺寸模糊图像,给目标的检测和识别带来了很大的困难。随着去模糊技术的进步^[1-2],将该技术运用到天基探测成像平台,在保证实时性的同时高质量复原宽波段的小目标轮廓,成为当前天基运动目标去模糊领域的热点与难点。

根据是否使用深度学习方法,去模糊技术分为两大类。基于传统优化方法的去模糊算法中,非盲去模糊^[3-4]只适用于模糊核已知情况,盲去模糊^[5-7]需要估计模糊核来实现图像的去模糊处理,数学模型较为复杂、计算效率较低,而且两种方法均存在严重的振铃效应。基于深度学习的去模糊方法大多为端到端,无需关注模糊核的具体形式,直接从图像中学习去模糊的映射关系,更能适应实际应用中的多类模糊情况。基于深度学习的去模糊方法根据网络架构可分为生成对抗网络(Generative Adversarial Network, GAN)^[8]、循环神经网络(Recurrent Neural Network, RNN)^[9]、TRANSFORMER^[10]、卷积神经网络(Convolutional Neural Network, CNN)^[11-12]。GAN通过生成器创造逼真数据和判别器区分真假数据的对抗过程来学习数据分布,但是其训练过程不稳定,对损失函数和训练策略提出了很高的设计要求。RNN通过单向或双向的时间流处理图像序列,结合隐藏状态积累全局时域信息,但序列每一步骤都需前一步骤的计算结果,限制了训练和推断速度。Transformer基于注意力机制进行信息的交互与传递,依赖输入序列长度,处理大规模数据时计算和存储成本高。多尺度多阶段卷积神经网络(MSSN)^[13]采用残差学习方案,简化学习目标以加速网络收敛同时解决网络加深带来的

梯度消失问题,以从粗到精的方式估计去模糊图像,可以有效融合大尺度的结构信息和小尺度的细节信息。因此,MSSN在保障算法实时性的同时,利用不同尺度下的信息实现了良好的去模糊效果。红外小目标图像缺乏丰富的纹理细节、像素数量稀少,有限的可用特征信息无法被传统的卷积操作提取到。上述研究工作主要运用在可见光波段的自然图像去模糊上,存在波段不匹配、实时性差和小目标轮廓难以高质量复原的问题。

针对该问题,本文提出了一种结合注意力机制的多尺度多阶段方法。将优化模块数量的轻量化多尺度多阶段网络作为架构保障实时性,结合基于频域的自注意力机制求解器及前馈神经网络(Frequency domain based self-attention solver and discriminative FFN, F-D)^[14]和卷积注意力模块(Concolutional Block Attention Module, CBAM)^[15]捕捉序列各元素间的关联性并判别关键高、低频特征信息,有效复原小目标的轮廓信息并提高了宽波段的泛化能力。

2 基本原理

2.1 基于天基平台的动目标模糊原理

用相机拍摄物体时,物体运动、相机抖动或者物体失焦会造成图像模糊,降低图像的视觉质量,这个过程即为图像退化。图像模糊一般被看作清晰图像卷积模糊核得到模糊图像的过程,其退化模型如下:

$$B = I \times K + N, \quad (1)$$

其中: B 为模糊图像, I 为待估计的清晰图像, K 为模糊核, N 为加性噪声, $\times$ 为卷积操作。天基平台的动目标成像过程主要由运动模糊、衍射模糊混合的模糊核合成模糊图像。

运动模糊包含相机抖动与目标高速运动^[16],

模糊核可表示为:

$$PSF(t) = \int_0^T A(\Delta_{\text{pix}}(t)) dt, \quad (2)$$

其中: Δ_{pix} 表示在时间 t 由于相机运动产生的像素位移; A 是空间积累操作符, 将位移转化为光照在相邻像素上进行分配; T 为曝光时间, 累积整个运动期间的光照变化。

光瞳会造成衍射模糊^[17], 光瞳函数可表示为:

$$P(x, y) = \text{rect}(x/a) \cdot \text{rect}(y/b), \quad (3)$$

其中 a 和 b 分别为光瞳的长与宽。对式(3)进行傅里叶变换到频域后, 取模的平方, 可得到光瞳的 PSF。该过程可表示为:

$$PSF(f_x, f_y) = a^2 b^2 \sin^2(\pi a f_x) \cdot \sin^2(\pi b f_y), \quad (4)$$

其中 f_x, f_y 为频域上的坐标。

高斯噪声代表探测器由于电子器件中随机热运动产生的噪声, 可表示为:

$$P(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (5)$$

其中: $P(\cdot)$ 为某个噪声值 x 出现的概率密度, σ 代表标准差, μ 表示噪声的平均值。

天基探测成像平台探测视场大, 小目标像素数少, 纹理特征弱, 受噪声影响大。因此, 在多尺度多阶段网络 MSSN 中, 各子网络设计在更细尺度上具有更深架构, 以准确反映模糊尺度及学习更深层次的特征。MSSN 通过跳跃连接实现输入的直接传递, 保证梯度可以无障碍地流动, 从而减轻梯度消失的问题。图像去模糊过程可表示为:

$$L = B + R, \quad (6)$$

其中: B 为输入的模糊图像, R 为网络预测的残差图像, L 为网络最后输出的去模糊图像。

2.2 网络结构

基于多尺度多阶段网络, 构建的小目标图像实时去模糊模型 MSSANet 结构如图 1 所示。MSSANet 由从粗到细的 S_1, S_2, S_3 3 个尺度组成。输入单幅模糊图像 B 后网络从粗到细地估计去模糊图像 L 。为了反映模糊尺度, MSSANet 在更细的尺度上有更深的结构, 即在 S_1, S_2, S_3 3 个尺度上分别拥有 1, 2, 3 个阶段, 每个阶段由单个编码器解码器 (Encoder-Decoder, E-D) 模块组成。

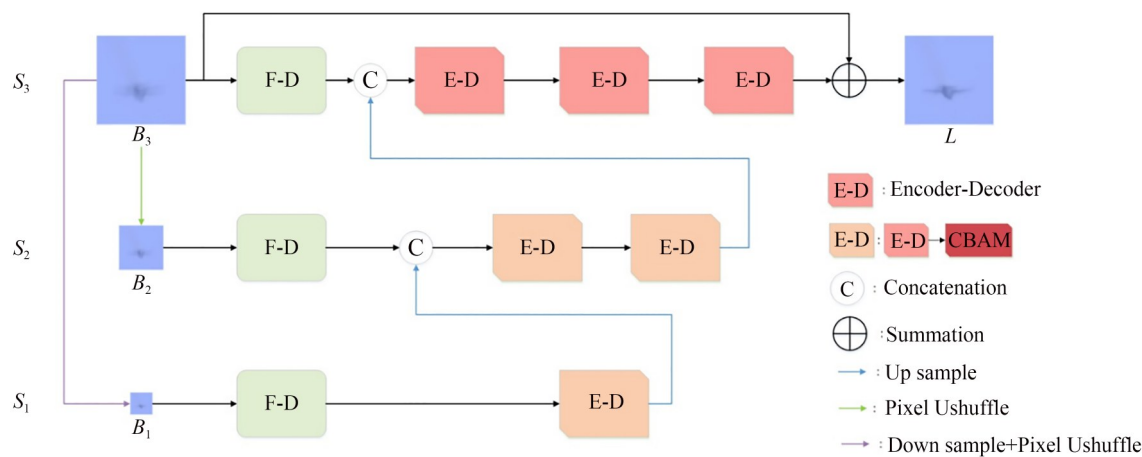


图 1 MSSANet 网络结构

Fig. 1 Network structure of MSSANet

在 S_3 尺度输入尺寸为 $W \times H \times 3$ 的模糊图像 B , 模糊图像在不同尺度提取的特征用 B_i 表示, i 为尺度标签, $B_3 = B$ 。在 S_2 尺度, 对 B_3 进行 Pixel-Unshuffle 操作得到尺寸为 $W/2 \times H/2 \times 12$ 的模糊图像 B_2 。在 S_1 尺度, 对 B_3 进行下采样得到尺寸为 $W/2 \times H/2 \times 12$ 的特征后, 进行 Pixel-

Unshuffle 操作得到尺寸为 $W/4 \times H/4 \times 48$ 的模糊图像 B_1 。

在 S_i 尺度, i 为尺度标签, 输入特征经过 F-D 模块后通过自注意力机制捕捉输入特征的长距离依赖关系, 判别并保留关键的高、低频信息。经过 F-D 模块处理的特征与 S_{i-1} 尺度传递的特

征进行拼接后进入E-D模块,关注图像在通道和空间维度的重要信息,以进行自适应特征细化后传递给 S_{i+1} 尺度。

2.3 F-D模块及CBAM模块结构

F-D模块及CBAM模块的具体结构如图2所示。图2(a)为F模块结构。输入空间分辨率为 $H \times W$,通道数为 C 的特征 X ,注意力机制对 X 进行线性变换得到特征 F_q, F_k, F_v ,再使用重塑函数获得查询 Q 、键 K 、值 V 。缩放点积注意力表示为:

$$V_{\text{att}} = \text{softmax}\left(\frac{QK^T}{\sqrt{CHW}}\right)V. \quad (7)$$

对于 $N \times N$ 的矩阵 Q 和 K ,在空域中计算矩阵乘法 QK^T 的时间复杂度为 $O(N^3)$,在频域中使用哈达玛积运算的时间复杂度为 $O(N^2)$ 。F

模块使用哈达玛积运算取代矩阵乘法降低了计算开销。网络通过 1×1 的逐点卷积和 3×3 的深度卷积从特征 X 中提取 F_q, F_k, F_v 。特征 F_q, F_k 在频域上的相关性为:

$$A = F^{-1}(F(F_q), \overline{F(F_k)}), \quad (8)$$

其中: F 代表傅里叶变换, F^{-1} 代表逆傅里叶变换, $\overline{F}$ 代表共轭转置运算。缩放点积注意力由下式计算:

$$V_{\text{att}} = L(A)F_v, \quad (9)$$

其中 L 表示层归一化操作。

最后输出特征为:

$$X_{\text{att}} = X + \text{Conv}1 \times 1(V_{\text{att}}), \quad (10)$$

其中 $\text{Conv}1 \times 1$ 表示卷积核大小为 1×1 的卷积操作。

图2(b)为D模块结构。D模块能够判别并

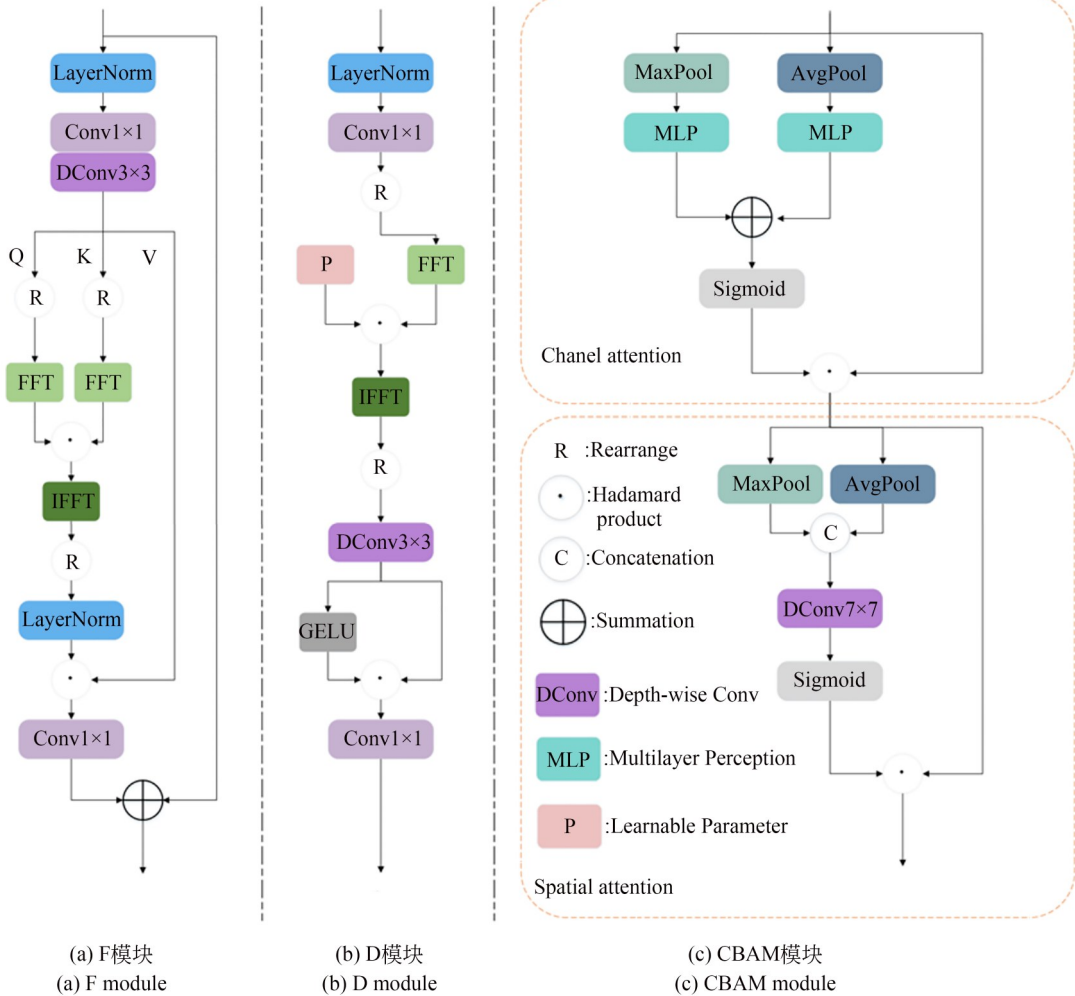


图2 F-D及CBAM的模块结构

Fig. 2 Block diagram of F-D and CBAM modules

保留关键的高、低频信息,输入特征向前传播的过程可表示为:

$$X_1 = F(\text{Conv}1 \times 1(L(X_{\text{att}}))), \quad (11)$$

$$X_2 = F^{-1}(P \odot X_1), \quad (12)$$

$$X_{\text{out}} = g(X_2) + X_{\text{att}}, \quad (13)$$

其中: $\odot$ 为哈达玛积运算, P 为网络中可学习的参数, g 为激活函数。

图 2(c)为 CBAM 模块结构。CBAM 模块包含通道注意力和空间注意力模块,增强了不同通道和位置的特征表示。

通道注意力机制中,输入的特征经过并行的最大池化层和平均池化层转化为两组 $C \times 1 \times 1$ 的特征图后,共享全连接层(Share MLP)将通道数压缩再扩张回原通道数。共享全连接层输出的两组特征进行逐元素相加后,应用 Sigmoid 激活函数产生通道注意力权重。利用注意力权重对原始特征图加权得到尺寸为 $C \times H \times W$ 的通道特征图。

空间注意力机制中,输入特征经过并行的最大池化层和平均池化层转化为两组 $1 \times H \times W$ 的特征图,沿着通道维度进行拼接操作得到 $2 \times H \times W$ 的特征图。特征图通过 7×7 的深度卷积降为 $1 \times H \times W$,应用 Sigmoid 激活函数产生空间注意力权重。利用注意力权重对原始特征图加权得到尺寸为 $C \times H \times W$ 的空间特征图。

2.4 模型参数设置与优化

残差块结构作为编码器-解码器模块的主干,如图 3 所示。

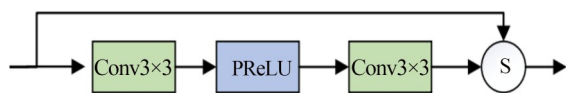


图 3 残差块结构

Fig. 3 Structure diagram of residual block

残差块的跳跃连接允许网络在不同层之间直接重用特征,梯度能够更有效地流动以加速网络训练。考虑到模型对时效性的要求,通过实验验证模块数量对算法运行时间及去模糊效果的影响,使用模型 $M_{i,j}$ 在红外仿真数据集上进行实验,其中 i 为注意力模块的数量, j 为残差块数量。模块数量与运行时间、去模糊效果的关系分别见

图 4 与表 1。

由图 4 可知,在注意力模块数量小于 4 时,算法无法恢复序列 1 的目标轮廓;残差块数量对目标复原效果无明显影响。由表 1 可知,模块数量

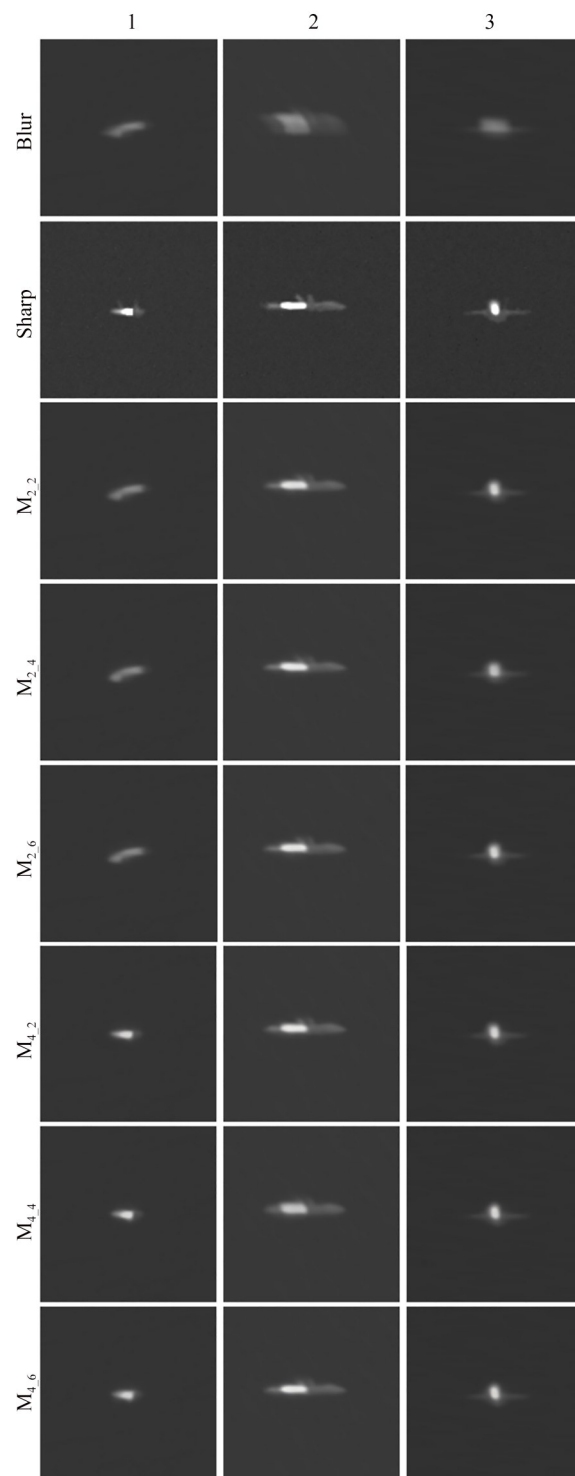


图 4 不同模块数量的模型运行去模糊结果

Fig. 4 Deblurring results of models with different numbers of modules

表 1 不同模块数量的模型运行时间

Tab. 1 Running time of models with different numbers of modules

Module	Time/s
$M_{2,2}$	0.027 0
$M_{2,4}$	0.027 8
$M_{2,6}$	0.034 7
$M_{4,2}$	0.028 4
$M_{4,4}$	0.035 4
$M_{4,6}$	0.042 2

的增加导致算法运行时间明显增长。因此在保证图像复原质量的前提下,算法中 F-D 模块中的注意力模块数量设置为 4,编码器-解码器模块每层的残差块数量设置为 2。

为扩展模型的适用波段,输入输出通道数设置为 3。当输入输出通道设置为 3,单通道红外图像复制为三通道图像不会导致信息损失;当输入输出通道设置为 1,输入的 RGB 图像需要转换为单通道图像,色彩相关的细节(色调、饱和度)会丢失。相比自然图像,小目标图像中的特征信息较少,因此,各卷积层的卷积核尺寸设置为 3×3 。

$$L_f = \frac{1}{N_1} \|F(L_1^1) - F(S_1)\|_1 + \sum_{j=1}^2 \frac{1}{N_2} \|F(L_2^j) - F(S_2)\|_2 + \sum_{j=1}^3 \frac{1}{N_3} \|F(L_3^j) - F(S_3)\|_3, \quad (17)$$

其中 F 为傅里叶变换。

3 实验结果与分析

3.1 数据集与评价指标

利用公开数据集 GoPro^[18]对模型进行训练,构建红外及可见光小目标图像数据集进行测试。数据集的详细信息如表 2 所示,其中相对尺寸为目标边界框宽高与图像宽高之比,小于 0.1 即为小目标。如图 5 所示,构建数据集的模糊核具有

学习率最初设置为 1×10^{-4} ,采用 Adam 优化器进行模型优化,设置 $\beta_1 = 0.9$, $\beta_2 = 0.999$,权重衰减系数为 5×10^{-4} ,batch size 设置为 8。

在训练 MSSANet 时,损失函数由内容损失 L_c 和频率重构损失 L_f 组成,具体公式如下:

$$L_{\text{final}} = L_c + \lambda L_f, \quad (14)$$

其中 $\lambda = 0.1$ 。内容损失函数采用 L1 损失,具体公式如下:

$$L_c = \frac{1}{N_1} \|L_1^1 - S_1\|_1 + \sum_{j=1}^2 \frac{1}{N_2} \|L_2^j - S_2\|_1 + \sum_{j=1}^3 \frac{1}{N_3} \|L_3^j - S_3\|_1, \quad (15)$$

其中: S 为真实的清晰图像, S_i 为 S 在第 i 个尺度下采样后的图像, L_i^j 为网络在第 i 尺度第 j 阶段生成的去模糊图像, N_i 为 i 尺度对应的归一化因子,计算公式如下:

$$N_i = (W/R) \times (H/R) \times 3, \quad (16)$$

其中: W 为图像宽度, H 为图像长度, R 为缩放因子。当 $i = 1$, $R = 1$; 当 $i = 2, 3$, $R = 2$ 。

频率重构损失可以最小化去模糊图像与清晰图像在频域的差异,从而恢复高频细节。具体公式如下:

多样化的尺寸和形态,有助于全面评估算法的复原能力。

实验使用的台式机包含了 1 块 Intel (R) Core (TM) i7-13700F CPU 和 1 块 NVIDIA GeForce RTX 4090 24G GPU,操作系统 Windows 10 下使用 Pytorch 框架运行网络。采用峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 与结构相似性 (Structural Similarity, SSIM) 两个典型参数指标评价小目标去模糊与轮廓恢复的效果。

(1) PSNR: 根据清晰图像与复原图像之间对

表 2 数据集信息

Tab. 2 Details of image datasets

Dataset	Blur type	Image size	Object	Relative size	Number
GoPro	Motion blur	$1\,280 \times 720$	Various	—	3 214
Infrared ray	Mixed blur	640×512	Plane	0.05—0.10	180
Visible light	Mixed blur	960×540	Plane	0.03—0.07	1 130

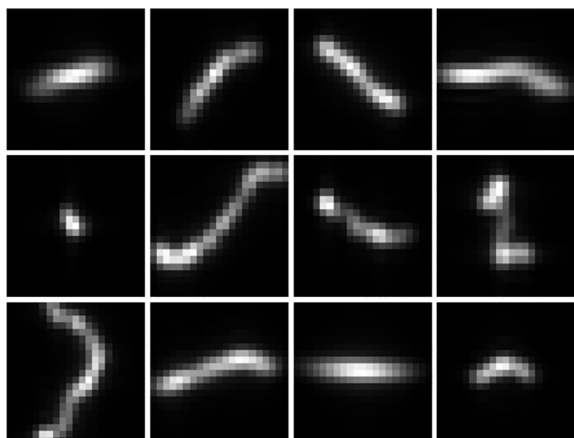


图5 模糊核可视化图像

Fig. 5 Blur kernel visualization image

应像素点的误差进行计算,PSNR越高,说明图像复原结果越好,公式如下:

$$PSNR = 10 \cdot \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right), \quad (18)$$

其中MSE为模型重建结果与真实标签之间的均方误差(Mean Square Error, MSE)。MSE的定义如下:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i, j) - \hat{I}(i, j)\|^2, \quad (19)$$

其中: $I(i, j)$ 与 $\hat{I}(i, j)$ 分别为真实标签图像和复原图像在 (i, j) 处的像素值, m 和 n 分别为图像长宽。

(2)SSIM:通过度量图像的亮度,对比度和结构3个特征来反映物体结构属性。SSIM值越接近1,说明图像的复原结果越好,公式如下:

$$SSIM(I, \hat{I}) = \frac{(2\mu_I\mu_{\hat{I}} + c_1)(2\sigma_{II} + c_2)}{(\mu_I^2 + \mu_{\hat{I}}^2 + c_1)(\sigma_I^2 + \sigma_{\hat{I}}^2 + c_2)}, \quad (20)$$

其中: μ_I 和 $\mu_{\hat{I}}$ 分别为 I 和 $\hat{I}$ 的均值, σ_I 和 $\sigma_{\hat{I}}$ 分别为 I 和 $\hat{I}$ 的标准差, σ_{II} 为协方差, c_1 和 c_2 是用来维持稳定的常数。

3.2 去模糊结果对比

3.2.1 红外仿真数据集去模糊

为了验证算法的有效性,选用典型算法 Restormer^[10], Ffftformer^[14], Swintormer^[19]和 MRDNet^[20]的去模糊结果与提出方法进行对比,在红外仿真数据集上的去模糊结果及运行时间分别如表3和图6所示。

表3 不同算法对红外小目标去模糊评价指标对比

Tab. 3 Comparison of evaluation indicators for deblurring infrared small targets using different algorithms

Method	PSNR/dB	SSIM	Time/s
MRDNet	31.544	0.875	0.056
Restormer	31.770	0.878	0.086
Ffftformer	31.585	0.877	0.088
Swintormer	31.641	0.877	0.125
MSSANet	32.083	0.877	0.029

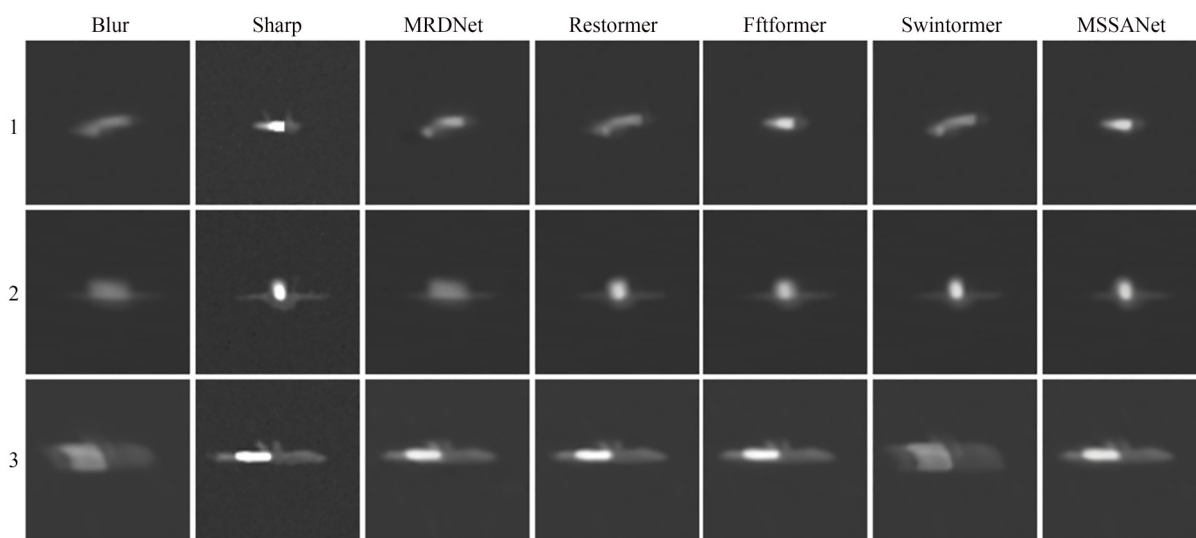


图6 部分红外仿真数据集实验对比结果

Fig. 6 Experimental comparison results of some infrared simulation datasets

由图 6 和表 3 可知:MSSANet 去模糊耗时最短且可以较好地复原飞机的形状与轮廓,在红外仿真数据集上去模糊结果的 PSNR 和 SSIM 不低于其他方法。序列 1 和序列 2 中的飞机经过模糊后轮廓完全丢失,而在去模糊图像中机身和尾焰的轮廓都得到了恢复;序列 3 中

的飞机经过模糊后出现了重影,在去模糊图像中成功消除了重影,恢复了目标的轮廓。

3. 2. 2 可见光数据集去模糊

为了验证算法在不同波段的泛化性,典型算法提出方法在可见光数据集上的去模糊结果及运行时间分别如图 7 和表 4 所示。

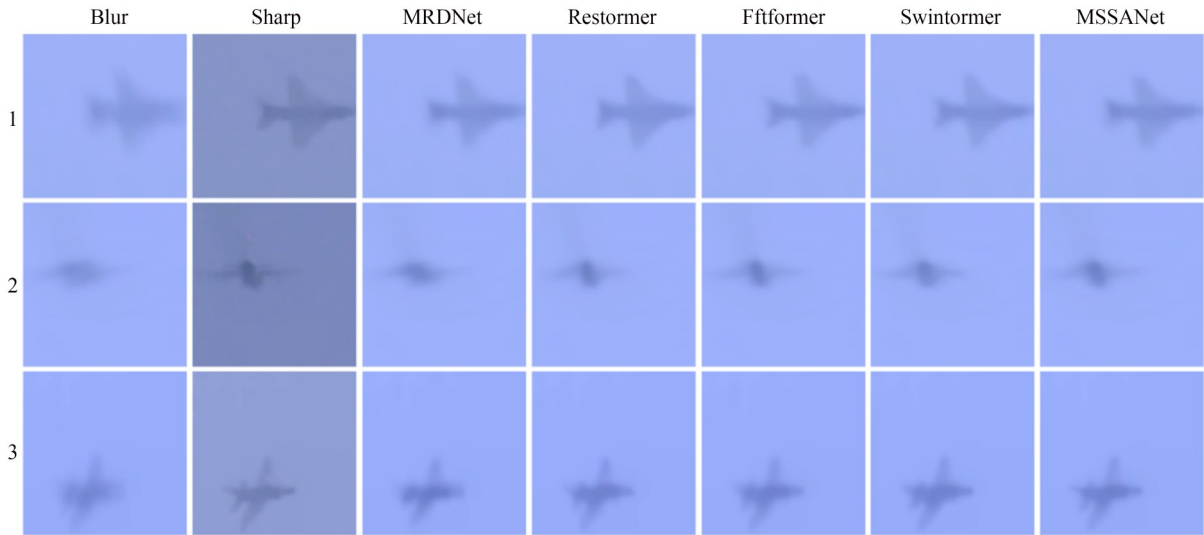


图 7 部分可见光数据集实验对比结果

Fig. 7 Experimental comparison results of some visible light datasets

表 4 不同算法对可见光小目标去模糊评价指标对比

Tab. 4 Comparison of evaluation indicators for deblurring visible small targets using different algorithms

Method	PSNR/dB	SSIM	Time/s
MRDNet	16.748	0.930	0.061
Restormer	16.851	0.929	0.084
Fftformer	16.970	0.931	0.087
Swintormer	16.983	0.931	0.104
MSSANet	17.118	0.930	0.028

由图 7 和表 4 可知:MSSANet 对目标轮廓的复原效果最佳且耗时最短,去模糊图像的 PSNR 均高于其他方法。序列 1 中的飞机轮廓边界不清晰,没有明显的边缘和形状,网络对图像进行去模糊后成功恢复了机身和机翼的轮廓;序列 2 和序列 3 中的飞机在模糊图像中机身轮廓完全丢失,而在去模糊图像中机身的轮廓都得到了恢复。

3. 3 消融实验

为验证 CBAM 模块和 F-D 模块对去模糊算法的有效性,在不改变训练集及训练参数的情况下对基础网络 MSSN、添加 CBAM 模块的 MS-SN+CBAM、添加 F-D 模块的 MSSN+F-D、添加 CBAM 模块与 F-D 模块的 MSSANet 进行训练,并在红外仿真数据集与可见光数据集上实验,结果如图 8~图 9 与表 5 所示。

由图 8~图 9 可知:在红外仿真数据集中,序列 1 和序列 2 中的飞机机身轮廓不清,尾焰严重变形,仅提出的 MSSANet 成功复原机身和尾焰的轮廓;序列 3 中的飞机由于模糊出现了重影,对基础网络添加 CBAM 模块或 F-D 模块均能消除重影,恢复机体轮廓。在可见光数据集中,序列 1 和序列 2 中飞机轮廓边界模糊,而序列 3 中飞机机翼形状难以辨认,各网络对图像进行去模糊后均能恢复机身和机翼的轮廓。

由表 5 可知:在红外仿真数据集和可见光数据集中,对基础网络单独添加 CBAM 模块或 F-D 模块均能在保障 SSIM 的同时提高 PSNR;MS-

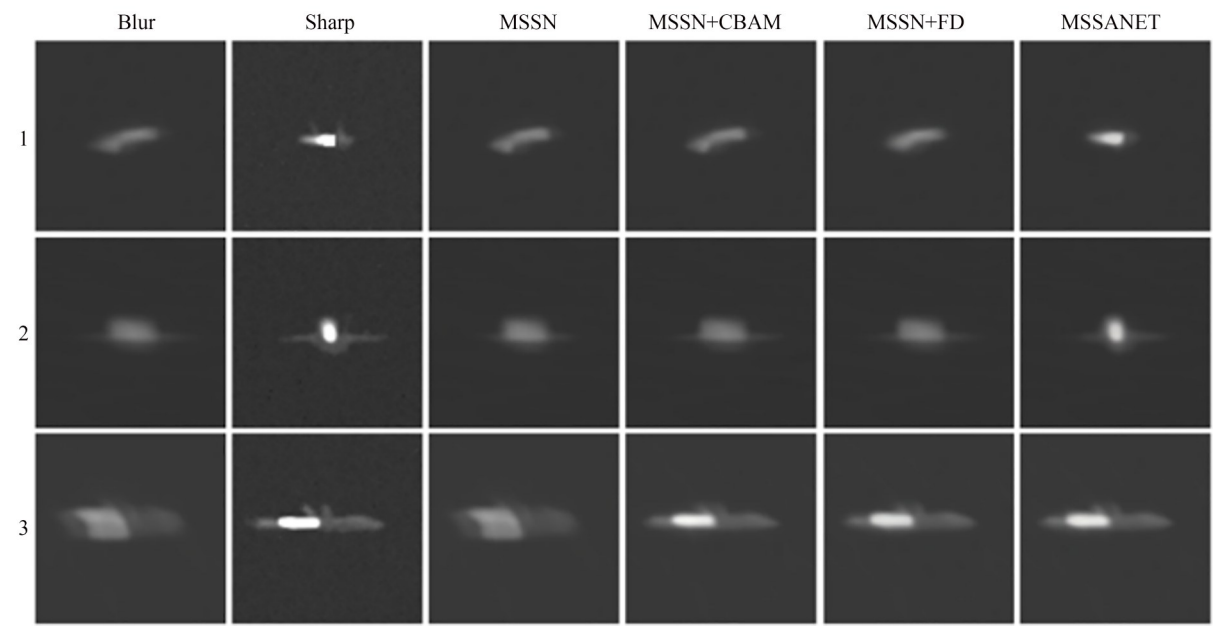


图 8 部分红外仿真数据集实验对比结果

Fig. 8 Experimental comparison results of some infrared simulation datasets



图 9 部分可见光数据集实验对比结果

Fig. 9 Experimental comparison results of some visible light datasets

表 5 消融实验结果

Tab. 5 Ablation experiment results

Algorithm	Infrared simulation dataset		Visible light dataset	
	PSNR/dB	SSIM	PSNR/dB	SSIM
MSSN	31.630	0.877	17.086	0.932
MSSN+CBAM	31.661	0.877	17.103	0.931
MSSN+F-D	31.757	0.877	17.093	0.931
MSSANet	32.083	0.878	17.118	0.931

SANet 融合了基础网络与 CBAM 模块和 F-D 模块,对 PSNR 的提升效果最为显著。

实验表明,基础网络结合 CBAM 机制模块和 F-D 模块均能提升模型的去模糊效果,联合使用能够更有效地提取图像特征,复原小目标轮廓。

4 结 论

本文提出了一种结合注意力机制的小目标图像实时去模糊方法,用于复原天基平台获取的模糊图像,采用残差连接和精简模块数量的轻量化

编码器解码器结构搭建多尺度多阶段网络,结合 F-D 模块和 CBAM 模块捕捉序列各元素间关联性并判别关键高、低频特征信息,通过实验验证了该方法的有效性。实验结果表明:去模糊方法的复原速率不低于 34 frame/s,远高于典型算法的 12 frame/s,在红外小目标数据集上的 PSNR 大于 32 dB,SSIM 大于 0.87;在可见光小目标数据集上的 PSNR 大于 17 dB,SSIM 大于 0.93。MSSANet 对于红外和可见光小目标均能去除模糊复原轮廓,对不同波段范围内目标具有广泛适应性。将天基探测成像平台与图像去模糊算法结合能够有效提升探测系统对小目标的检测与识别能力。

参考文献:

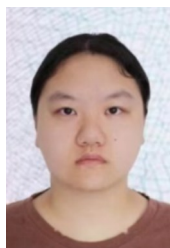
- [1] 王聪. 基于深度学习的图像复原与编辑[D]. 长春: 吉林大学, 2023.
WANG C. *Image Restoration and Editing Based on Deep Learning* [D]. Changchun: Jilin University, 2023. (in Chinese)
- [2] 李耀维. 基于深度学习的图像去模糊算法研究[D]. 上海: 同济大学, 2022.
LI Y W. *Research on Image Deblurring Algorithm Based on Deep Learning* [D]. Shanghai: Tongji University, 2022. (in Chinese)
- [3] WIENER N. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: with Engineering Applications* [M]. Cambridge, MA: MIT press, 1964.
- [4] RICHARDSON W H. Bayesian-based iterative method of image restoration[J]. *Journal of the Optical Society of America*, 1972, 62(1): 55.
- [5] SHAN Q, JIA J Y, AGARWALA A. High-quality motion deblurring from a single image[J]. *ACM Transactions on Graphics*, 2008, 27(3): 1-10.
- [6] SUN L B, CHO S, WANG J, *et al.* Edge-based blur kernel estimation using patch priors[C]. *IEEE International Conference on Computational Photography (ICCP)*. April 19-21, 2013. Cambridge, MA. IEEE, 2013: 1-8.
- [7] 曹学影, 王琳, 谭章燕, 等. 自适应双刃边法模糊红外图像复原[J]. *光学精密工程*, 2022, 30(5): 630-640.
CAO X Y, WANG L, TAN Q Y, *et al.* Restoration of blurred infrared image by adaptive double edge method[J]. *Opt. Precision Eng.*, 2022, 30(5): 630-640. (in Chinese)
- [8] 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*. June 18-23, 2018. Salt Lake City, UT. IEEE, 2018: 8183-8192.
- [9] PARK D, KANG D U, KIM J, *et al.* Multi-temporal Recurrent Neural Networks for Progressive Non-uniform Single Image Deblurring with Incremental Temporal Training[M]. *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2020: 327-343.
- [10] ZAMIR S W, ARORA A, KHAN S, *et al.* Restormer: efficient transformer for high-resolution image restoration[EB/OL]. 2021: 2111.09881. <https://arxiv.org/abs/2111.09881v2>.
- [11] NAH S, KIM T H, LEE K M. Deep multi-scale convolutional neural network for dynamic scene deblurring[C]. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017. Honolulu, HI. IEEE, 2017: 3883-3891.
- [12] 王成曦, 罗晨, 周江涛, 等. 深度特征维纳反卷积用于均匀离焦盲去模糊[J]. *光学精密工程*, 2023, 31(18): 2713-2722.
WANG CH X, LUO CH, ZHOU J H, *et al.* Uniform defocus blind deblurring based on deeper feature-based Wiener deconvolution [J]. *Opt. Precision Eng.*, 2023, 31(18): 2713-2722. (in Chinese)
- [13] KIM K, LEE S, CHO S. MSSNet: Multi-scale-stage Network for Single Image Deblurring[M].

- Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2023: 524-539.
- [14] KONG L S, DONG J X, GE J J, *et al.* Efficient frequency domain-based transformers for high-quality image deblurring[C]. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 17-24, 2023. Vancouver, BC, Canada. IEEE, 2023: 1-10.
- [15] WOO S, PARK J, LEE J Y, *et al.* CBAM: Convolutional Block Attention Module [M]. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018: 3-19.
- [16] GAVANT F, ALACOQUE L, DUPRET A, *et al.* A physiological camera shake model for image stabilization systems [C]. 2011 *IEEE SENSORS Proceedings*. October 28-31, 2011. Limerick, Ireland. IEEE, 2011: 1461-1464.
- [17] 宋欣晖. 高分辨率光瞳旋转光学成像系统特性与像质提升方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2020.
- SONG X H. *Study on the Characteristics of High-resolution Pupil Rotation Optical Imaging System and the Method of Image Quality Improvement* [D]. Harbin: Harbin Institute of Technology, 2020. (in Chinese)
- [18] NAH S, KIM T H, LEE K M. Deep multi-scale convolutional neural network for dynamic scene deblurring[J]. *CoRR*, 2016, abs/1612.02177.
- [19] CHEN K, LIU Y J. Efficient image deblurring networks based on diffusion models[EB/OL]. 2024: 2401.05907[. <https://arxiv.org/abs/2401.05907v2>].
- [20] ZHANG B, SUN J, SUN F M, *et al.* Image deblurring method based on self-attention and residual wavelet transform[J]. *Expert Systems with Applications*, 2024, 244: 123005.

作者简介:



朱德燕(1990—),女,重庆人,博士,副教授,硕士生导师,2011年于哈尔滨工业大学获得学士学位,2016年于中国科学院长春光机所获得博士学位,主要从事光电探测系统的设计与检测、计算成像与超分探测方面的研究。E-mail: zdy_nuaa@163.com



敖咏琪(1999—),女,江西樟树人,硕士研究生,2022年于南京航空航天大学获得学士学位,主要从事图像处理、深度学习的研究。E-mail: ayq_nuaa@163.com