

doi: 10.14132/j.cnki.1673-5439.2025.05.007

单幅图像去模糊的多尺度特征提取和融合网络

武婷婷, 万少杰

(南京邮电大学 理学院, 江苏 南京 210023)

摘要:近年来,多层网络在图像去模糊领域取得了较大进展,但其性能受限于特征提取和残差连接。为解决这些问题,提出了一种多尺度融合网络(Multi-Scale Feature Extraction and Fusion Network, MSFN)用于图像去模糊,通过多尺度输入与输出,增强了对图像特征的提取能力。MSFN 利用其特征自适应细节增强(Adaptive Detail Enhancement, ADE)模块和跨尺度特征融合(Cross-Scale Feature Fusion, CSFF)模块,在不同网络深度上捕获不同尺度的特征,优化了特征提取过程,并有效融合了多尺度信息。实验结果表明,所提出的算法在定量分析上表现出色,并且在主观视觉效果上也得到了显著提升,这些结果充分证明了所提网络的卓越性能。

关键词:图像去模糊;深度学习;多尺度;细节增强;特征融合

中图分类号:TP391.41;TN911 文献标志码:A 文章编号:1673-5439(2025)05-0057-09

Multi-scale feature extraction and fusion network for single image deblurring

WU Tingting, WAN Shaojie

(College of Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: Significant advancements have been made in image deblurring through multi-layer networks, but their performance remains limited by challenges in feature extraction and residual connections. To address these issues, this paper proposes a multi-scale feature extraction and fusion network (MSFN) for image deblurring. The core idea of the network is to enhance image feature extraction through multi-scale inputs and outputs. Further, MSFN utilizes its feature adaptive detail enhancement (ADE) modules and cross-scale feature fusion (CSFF) modules to capture multi-scale features at different network depths, thereby optimizing the residual connection process and effectively integrating multi-scale information. Experimental results demonstrate that the proposed algorithm achieves superiority in quantitative analysis and significantly improves subjective visual effects, exhibiting an advanced performance.

Keywords: image deblurring; deep learning; multiple scale; detail enhancement; feature fusion

图像去模糊是一个经典的计算机视觉问题^[1-4],其目标是仅凭一张模糊的图像来恢复出原本清晰的图像。这个过程可用数学公式来描述,即模糊图像是由原始清晰图像与模糊核进行卷积操作后,再

加上一些随机噪声形成的图像,表示为

$$I = K * S + n \quad (1)$$

其中, I 表示模糊图像, K 表示模糊核, S 表示清晰图像, n 表示随机噪声。因为不同组的清晰图像 S 和

收稿日期:2024-12-02;修回日期:2025-04-07 本刊网址: <http://nyzr.njupt.edu.cn>

基金项目:国家自然科学基金(61971234)和江苏省研究生科研与实践创新计划项目(KYCX23_0960)资助项目

作者简介:武婷婷,女,博士,教授,博士生导师, wutt@njupt.edu.cn

引用本文:武婷婷,万少杰.单幅图像去模糊的多尺度特征提取和融合网络[J].南京邮电大学学报(自然科学版),2025,45(5):57-65.

Citation: WU Tingting, WAN Shaojie. Multi-scale feature extraction and fusion network for single image deblurring[J]. Journal of Nanjing University of Posts and Telecommunications(Natural Science Edition), 2025, 45(5): 57-65.

模糊核 K 能产生相同的模糊图像 I , 所以在已知模糊图像 I 的情况下, 式(1)求解清晰图像是一个不适定问题。

为得到清晰图像, 现有的方法提出各种各样的先验^[5-7], 使得去模糊问题变得适定。这类先验需要大量的经验和研究推测得到, 因此很难进行推广。为改善这一问题, 最近提出的方法使用卷积神经网络(Convolutional Neural Networks, CNN), 通过在大规模数据中捕获自然图像的信息来隐式学习更适用的先验。基于CNN的方法去模糊效果显著主要得益于其模型的设计, 现已有许多方法^[8-11]开发用于图像恢复的网络模块和功能单元, 包括残差学习、空洞卷积、注意力机制、密集连接、编码器-解码器和生成模块。然而, 这类方法的模型都是基于单层设计的, 虽然可以关注到图像的全局信息, 但随着网络层数的加深, 图像中复杂的结构可能会被忽略, 从而降低去模糊任务的性能。

因此人们将多层网络用于图像去模糊^[12-14], 不过这些方法性能上依旧存在瓶颈。首先, 多层网络只使用单一尺度的输入可能无法捕捉到所有重要的特征, 特别是在不同尺度上变化的特征。其次, 仅使用简单的卷积层来提取每一层的图像, 虽然节省模型的算力, 但通常只能捕获局部邻域内的信息, 会限制其提取特征的能力, 造成信息损失。最后, 直接地将各层通过残差相连接, 虽有助于解决深层网络训练中的退化问题, 却未能充分利用不同尺度上图像的相关信息, 网络性能反而下降。

为突破先前工作的局限性, 本文提出了一种多尺度特征提取和融合网络(Multi-Scale Feature Extraction and Fusion Network, MSFN)。该网络探索图像的多尺度输入与输出, 以利用跨尺度的图像信息进行单幅图像去模糊。接着, 开发一个有效的通道自适应细节增强(Adaptive Detail Enhancement, ADE)模块和跨尺度特征融合(Cross-Scale Feature Fusion, CSFF)模块, 将所提出模块嵌入到一个端到端的可训练网络中, 并且在通用数据集上取得良好的去模糊效果。

本文的主要贡献如下:

(1) 提出一种新颖的多尺度方法, 并结合创新性的CSFF模块, 将图像去模糊任务细化为3个子任务。通过多尺度的输入和输出, 网络能够在不同深度上拥有不同大小的感受野, 有效捕捉模糊图像中的边缘信息。同时, 将网络多个尺度上的特征进行有效融合, 避免简单特征连接导致的信息损失, 从而获得更全面的图像表示, 显著提升了复杂视觉任

务的性能。

(2) 进一步提出一个有效的ADE模块, 该模块在每层都充分提取图像的细节和纹理。通过优化特征的提取过程, ADE模块帮助网络在保持全局上下文的同时, 也不丢失部分特征, 避免图像的局部模糊。

1 相关工作

1.1 基于CNN的方法

随着深度学习的发展, 直接使用CNN评估清晰图像的方法层出不穷。Nah等^[15]开发出一种多尺度CNN, 将模糊图像由大尺寸到小尺寸逐一加入网络中训练, 在效果上相较之前的方法有较大改善。为更好探索图像的多尺度信息, Tao等^[16]提出一个更大规模的网络, Gao等^[9]提出共享参数和嵌套残差连接。但更深和更宽的网络容易在训练数据上过拟合, 从而降低模型在未知数据上的泛化能力。Zhang等^[17]开发一种深层多模块网络, 将前一阶段的特征连接起来以便于对下一阶段进行估计。Cui等^[18]开发一个多分支模块和内容感知模块, 局部动态地将特征分解为单独的频率子带, 然后使用通道注意力权重来强调特征中有用的部分。但在特征融合的过程中, 这类方法直接串联特征向量可能会导致在高层次特征传递过程中的关键视觉信息遭受损失, 进而影响模型对复杂计算机视觉问题的全面表征能力。

1.2 基于Transformer的方法

由于Transformer可以对全局上下文进行建模, 现已在许多高级视觉任务中取得重大进展。然而Transformer应用于图像会导致计算成本增加, 为缓解这种状况, Zamir等^[12]使用图像特征深度域中的点积缩放注意力机制, 将不同特征的信息沿通道维度进行传播。虽然能快速去模糊, 但没有充分探索图像中的空间信息。Wang等^[13]提出一种基于UNet的Transformer方法, 该方法使用非重叠窗口的自注意力机制来进行单幅图像去模糊。尽管使用分割策略降低其计算成本, 但多层神经网络仅依赖单一尺度的输入数据, 可能无法为特征提取模块提供足够的上下文信息, 影响网络对图像的整体理解。

2 网络结构

本文的目标是构建一种用于图像去模糊的MSFN模型, 该模型在增强图像边缘细节的同时, 有效防止了局部模糊现象。为打破去模糊的瓶颈, 模型中使用特征增强和融合模块。使得该模型相较

于单尺度网络,能够同时专注于图像的全局信息和局部信息。首先介绍 MSFN 的整体框架如图 1 所示,

接着描述新引入的核心模块:ADE 模块和 CSFF 模块,最后讨论损失函数的使用。

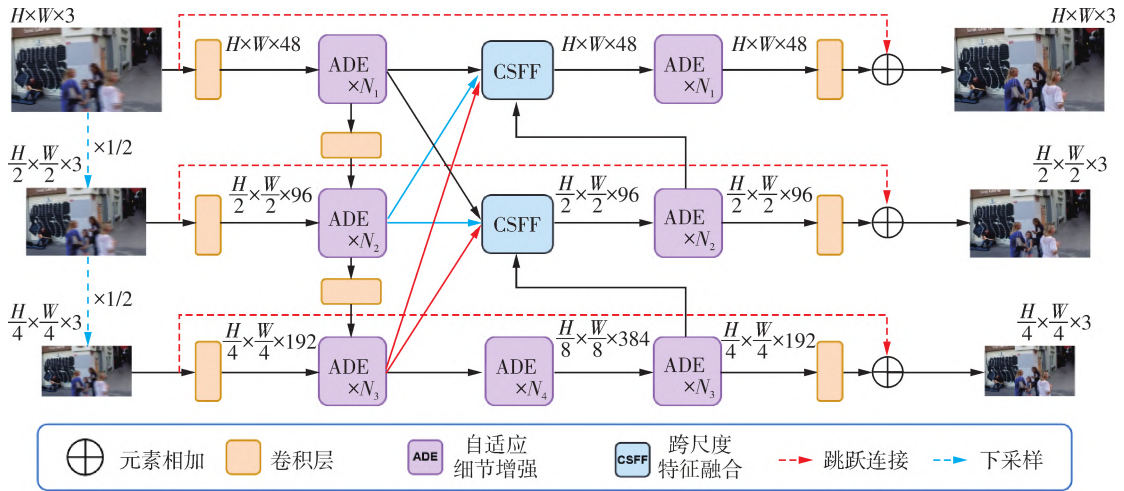


图 1 MSFN 结构

2.1 整体框架

给定一张模糊图像 $I \in \mathbf{R}^{H \times W \times C}$, 其中 H, W, C 分别表示图像的高度、宽度和通道维度。首先采用二阶段的下采样处理,生成 3 种尺寸的图像输入模型。由于模糊图像中的边缘不够清晰,多尺度输入有助于模型捕捉丰富的边缘信息,从而提升细节的识别能力。模型中每层使用 3×3 的卷积核提取图像特征,并引入 ADE 模块捕获局部纹理信息,增强对图像的理解能力。ADE 模块能够整合不同分辨率的特征,适应模糊图像的非均匀性,针对性地处理不同区域。在最低层通过三类 ADE,分别将提取到的特征进行编码、精炼、解码。上层则利用 CSFF 模块,不仅从左边 3 个 ADE 编码模块中学习到不同尺度的信

息,还融合下层 ADE 解码模块中的图像特征,以恢复模糊图像中丢失的高频细节信息。每层 ADE 解码后,通过 3×3 卷积与残差学习得到去模糊的图片。最终串联多个 ADE 和 CSFF 模块形成一个多尺度特征提取和融合的框架,达到去模糊效果。

2.2 自适应细节增强(ADE)模块

尽管多尺度网络仅使用 3×3 卷积层提取特征已取得一定效果,但面对模糊图像中的复杂空间结构,这种方法仍显不足。模糊图像中,物体以不同的速度或方向移动会导致不同区域上的局部模糊。对此,本文通过使用 ADE 模块来专注于图像局部主体结构进一步改善特征提取效果。ADE 模块分为两部分:细节融合和特征增强,如图 2 所示。

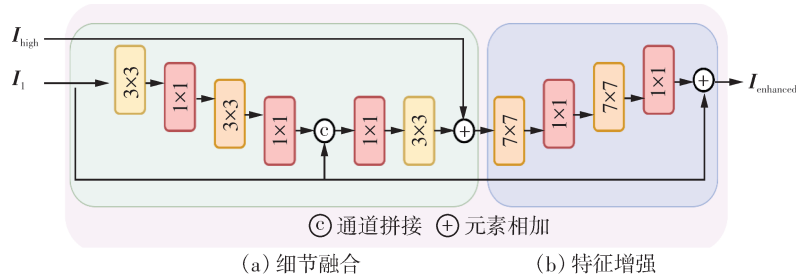


图 2 ADE 结构

在细节融合部分,给定高分辨率特征张量 $I_{high} \in \mathbf{R}^{H \times W \times C1}$ 和低分辨率特征张量 $I_1 \in \mathbf{R}^{H \times W \times C2}$ 。首先 I_1 通过两对 3×3 和 1×1 的卷积提取特征得到 I_2 , 然后使用通道拼接 I_1 和 I_2 得到处理后的低分辨率特征 I_3 。为学习到更丰富的细节纹理,对 I_3 进行

1×1 和 3×3 的卷积,与上层的高分辨率特征有效地融合得到 I_{fusion} 。当不存在上层特征时,则不需要细节融合。细节融合过程用数学公式表达为

$$I_2 = f_1\left(f_3\left(f_1\left(f_3(I_1)\right)\right)\right) \quad (2)$$

$$I_3 = f_c(I_1, I_2) \quad (3)$$

$$I_{\text{fusion}} = I_{\text{high}} + f_3(f_1(I_3)) \quad (4)$$

其中, $f_1(\cdot)$ 表示 1×1 卷积函数, $f_3(\cdot)$ 表示 3×3 卷积函数, $f_c(\cdot)$ 表示通道拼接函数。全局特征提供图像的整体信息, 而局部特征捕捉图像的细节。两者的融合为模型提供了全面的视角, 增强其对数据的理解。

Ding 等^[19] 提出, 通过采用大尺寸卷积核(如 31×31), 能够有效扩展模型的感受野, 从而使得模型能够捕获图像中的全局信息以及复杂的形状特征, 这一设计使得模型在图像分类、分割和识别等任务上取得了优异的表现。受到文献[19]的启发, 本文在特征增强部分加入两对 7×7 和 1×1 的卷积。最后使用残差连接得到增强后的特征 I_{enhanced} 。数学公式表达为

$$I_{\text{enhanced}} = I_1 + f_1\left(f_7\left(f_1\left(f_7\left(I_{\text{fusion}}\right)\right)\right)\right) \quad (5)$$

其中, $f_7(\cdot)$ 表示 7×7 卷积函数。本文采用 7×7 卷积核的设计, 相较于较小卷积核, 能够有效扩大模型感受野, 同时相比较大卷积核, 模型更专注于提取图像局部主体结构的特征。这样设计使得模型在处理局部模糊这一关键难点时更为有效, 从而契合去模糊的实际需求。

2.3 跨尺度特征融合(CSFF)模块

目前常用的多尺度融合方法通过整合单一尺度图像在不同层级的特征图, 来提升模型对不同尺度对象的感知能力。这种技术首先提取浅层和深层的特征图, 通过上采样将深层特征图放大至较低分辨率, 并与浅层特征图进行逐像素相加, 实现特征融合。但这类方法可能限制信息流动, 使网络难以捕捉复杂特征。此外, 模糊核扩散效应导致高频信息损失, 降低图像清晰度, 这进一步强调网络在处理模糊图像时需要更强的信息整合能力。因此在 MSFN 中, 本文将 3 个 ADE 编码模块和一个 ADE 解码模块之间引入一个 CSFF 模块。不同于多尺度融合方法, CSFF 模块通过融合来自多尺度图像的各层级特征图, 为模型提供了从局部到全局的多层次上下文信息。这种方法使模型能够更深入地理解图像内容, 并在复杂环境中提高去模糊能力, 具体实现方法如下。

图 1 中的两个 CSFF 模块在处理前 3 个编码模块输出的方式不相同。给定编码器 1、编码器 2 和编码器 3 的特征张量 $X_1 \in \mathbf{R}^{H \times W \times 48}$, $X_2 \in \mathbf{R}^{(H/2) \times (W/2) \times 96}$ 和 $X_3 \in \mathbf{R}^{(H/4) \times (W/4) \times 192}$, 解码器的特征张量 $Y_1 \in \mathbf{R}^{H_1 \times W_1 \times C}$ 。上层的 CSFF 模块需要将 X_2

上采样 2 倍, X_3 上采样 4 倍; 下层 CSFF 模块则需要将 X_1 下采样 2 倍, X_3 上采样 2 倍。图 3 展示了上层 CSFF 模块的结构。该模块首先将 3 个编码器得到的深层特征经过上述处理后拼接在一起, 再进行 1×1 和 3×3 卷积。在学习过程中, 通过随机归零部分权重降低节点间的相互依赖, 再与 X_1 进行残差连接得到融合特征 X_4 。解码器的张量 Y_1 通过 3×3 卷积和像素重组后, 再与 X_4 拼接, 最终进行 1×1 卷积得到不同尺度融合后的特征。CSFF 模块使用数学公式表达为

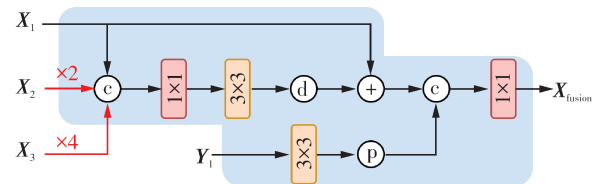
$$X_2 = f_{\text{up}}^1(X_2) \quad (6)$$

$$X_3 = f_{\text{up}}^2(X_3) \quad (7)$$

$$X_4 = X_1 + f_d\left(f_3\left(f_1\left(f_c(X_1, X_2, X_3)\right)\right)\right) \quad (8)$$

$$X_{\text{fusion}} = f_1\left(f_c\left(X_4, f_p\left(f_3(Y_1)\right)\right)\right) \quad (9)$$

其中, $f_{\text{up}}^i(\cdot)$ 表示上采样至原始尺寸的 2^i 倍, $f_d(\cdot)$ 表示随机失活函数, $f_p(\cdot)$ 表示像素重组函数。本文提出的 CSFF 相较残差连接具有显著优势: 首先, CSFF 通过融合不同尺度的特征, 有效减少过程中因重复上采样和下采样函数导致重要信息丢失的影响。其次, 该模块利用多尺度特征丰富每一层的表示, 为复杂视觉任务提供更全面的特征理解。最后, CSFF 的多尺度特性使模型能够灵活适应输入变化, 有效处理各种分辨率的模糊图像。



③ 通道拼接 ④ 随机失活 ⊕ 元素相加 ⑤ 像素重组 → 上采样

图 3 CSFF 结构

2.4 损失函数

本文提出一种多尺度网络架构, 旨在通过不同尺度的图像处理, 提高模型去模糊能力。该网络结构由 3 个尺度组成, 分别对应输入图像尺寸的 1 倍、1/2 倍和 1/4 倍。这种策略有助于平衡模型对细节和全局特征的处理, 并有效提取多级边缘信息。为训练网络, 本文构建了一个复合损失函数, 评估模型输出与清晰图像之间的差异, 并指导模型的优化过程。损失函数定义如下

$$L = \sum_{i=0}^2 \lambda_i \|f_{\text{down}}^i(S) - Z_i\|_2^2 \quad (10)$$

其中, S 表示清晰图像, Z_i 表示第 i 尺度的输出图像,

$f_{\text{down}}^i(\cdot)$ 函数表示下采样至 $1/2^i$ 倍。 λ_i 表示第 i 尺度的权重参数。通过这种权重分配,为不同尺度的特征提供差异化的学习信号,确保模型在训练过程中对每个尺度给予适当的重视,从而促进多尺度学习。

3 实验与评估

本节通过将MSFN模型与一些先进的方法进行比较,分别使用定性分析和定量分析的方式验证模型在图像去模糊中的有效性。在定量分析中,首先使用峰值信噪比^[20](Peak Signal-to-Noise Ratio, PSNR)和结构相似性(Structural Similarity Index Measure, SSIM)来客观地比较不同方法去模糊效果的差异。PSNR与SSIM定义如下

$$\text{PSNR} = 10 \lg \left(\frac{HW}{\|\mathbf{x} - \mathbf{y}\|_2^2} \right) \quad (11)$$

$$\text{SSIM} = \frac{(2\mu_x \mu_y + c1)(2\sigma_{xy} + c2)}{(\mu_x^2 + \mu_y^2 + c1)(\sigma_x^2 + \sigma_y^2 + c2)} \quad (12)$$

其中, \mathbf{x} 与 \mathbf{y} 表示两张数值在(0,1)之间的图像, μ_x 与 μ_y 表示 \mathbf{x} 与 \mathbf{y} 的平均亮度, σ_x^2 与 σ_y^2 表示 \mathbf{x} 与 \mathbf{y} 的亮度方差, σ_{xy} 表示亮度协方差, $c1$ 与 $c2$ 为常数。较高的PSNR和SSIM值表示图像的细节、纹理等视觉特征得到较好的保留,去模糊效果较好。

为深入评估MSFN在图像去模糊领域的性能,进一步分析其在目标检测任务中的应用,并量化对检测精度的实际提升。采用平均精度^[21](Average Precision, AP)作为衡量指标,来评估去模糊后的目标检测准确度。AP的计算方法具体如下

$$P_{\text{interp}}(R) = \max_{\tilde{R}: \tilde{R} \geq R} P(\tilde{R}) \quad (13)$$

$$\text{AP} = \frac{1}{11} \sum_{R \in \{0, 0.1, \dots, 0.9, 1\}} P_{\text{interp}}(R) \quad (14)$$

其中,召回率 R 可在 $\{0, 0.1, \dots, 0.9, 1\}$ 中设定阈值, $P_{\text{interp}}(R)$ 表示在大于等于阈值 R 的样本中选取精确率的最大值。AP指标综合考虑不同召回率下的精确率,以评估模型的整体性能。因此,对于同一种目标检测算法,去模糊效果越好,图像越清晰,目标检测越准确,从而AP值也越高。

3.1 数据集

本文使用两个数据集:GoPro数据集^[15]和HIDE数据集^[22]。GoPro数据集专为图像去模糊设计,包含2 103组训练图像和1 111组测试图像,每组包括一张模糊图像及其对应的清晰图像。HIDE数据集采用类似合成方法,专注于行人图像,包含6 397组训练图像和2 025组测试图像。模型仅利用GoPro数据

集进行网络训练,并在GoPro和HIDE的测试集上进行评估,以检验MSFN的去模糊效果和泛化能力。

3.2 实验设置

本文采用适应性矩估计算法(Adaptive Moment Estimation, Adam)训练MSFN模型,初始学习率为 1×10^{-4} ,通过余弦退火策略逐步降至 1×10^{-6} 。随着学习率的降低,批次大小减少至 $[4, 2, 1]$,图像尺寸增大至 $[128, 256, 512]$ 。数据增强包括随机裁剪、旋转和饱和度调整。损失函数中, $\lambda_1 = 1, \lambda_2 = 1/2, \lambda_3 = 1/4$ 。

3.3 去模糊结果与分析

首先将本文的去模糊结果与当前最先进的方法得到的结果进行比较,这些方法包括基于CNN和基于Transformer的算法。所有的方法都在GoPro训练集上训练,并且在GoPro和HIDE测试集上进行比较。通过直接引用已有实验数据的结果,对于缺少数据的方法,则根据原作者的参数设置进行测试,来确保实验的公正性。

表1展示在GoPro和HIDE测试集上的定量评估结果,本文提出的方法在PSNR和SSIM上均取得了最优值。在GoPro测试集上,与基于CNN的方法FSNet^[18]相比,本方法在PSNR上提升0.35 dB,在SSIM上提升0.004;与基于Transformer的Uformer^[13]

表1 测试集上的去模糊比较

模型	GoPro		HIDE	
	PSNR/dB	SSIM	PSNR/dB	SSIM
DeblurGAN ^[8]	28.70	0.858	24.51	0.871
DeepDeblur ^[15]	29.08	0.914	25.73	0.874
Zhang et al. ^[23]	29.19	0.931		
DeblurGAN-v2 ^[24]	29.55	0.934	26.61	0.875
SRN ^[16]	30.26	0.934	28.36	0.915
Shen et al. ^[22]	30.26	0.940	28.39	0.930
Gao et al. ^[9]	30.90	0.935	29.11	0.913
DBGAN ^[10]	31.10	0.942	28.94	0.915
MTRNN ^[25]	31.15	0.945	29.15	0.918
DMPHN ^[17]	31.20	0.940	29.09	0.924
Suin et al. ^[26]	31.85	0.948	29.98	0.930
SPAIR ^[27]	32.06	0.953	30.29	0.931
MIMO-UNet+ ^[28]	32.45	0.957	29.99	0.930
MPRNet ^[29]	32.66	0.959	30.96	0.939
Restormer ^[12]	32.92	0.961	31.22	0.942
Uformer ^[13]	33.05	0.962	30.89	0.940
MSSNet ^[30]	33.01	0.961	30.79	0.938
NAFNet ^[14]	33.08	0.963	31.22	0.943
FSNet ^[18]	33.29	0.963	31.05	0.941
MSFN(本文)	33.64	0.967	31.59	0.947

相比,PSNR提升0.59 dB,SSIM提升0.005。在HIDE测试集上,与NAFNet^[14]相比,PSNR提升0.37 dB,SSIM提升0.004;与Restormer^[12]方法相比,PSNR提升0.37 dB,SSIM提升0.005。这些结果表明,本方法在图像去模糊任务中具有显著的性能优势。

为验证MSFN去模糊算法的效果,本文在Go-Pro和HIDE测试集上选取代表性图像,并与顶尖去模糊算法进行对比。对比结果在图4和图5中展示,旨在深入分析算法在细节特征处理上的性能。

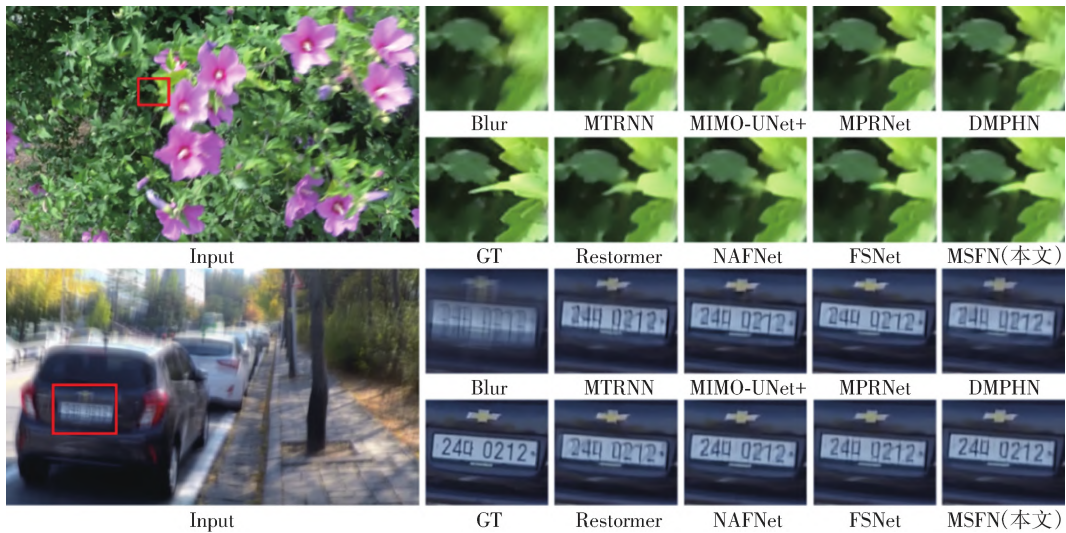


图4 GoPro数据集上的实验结果

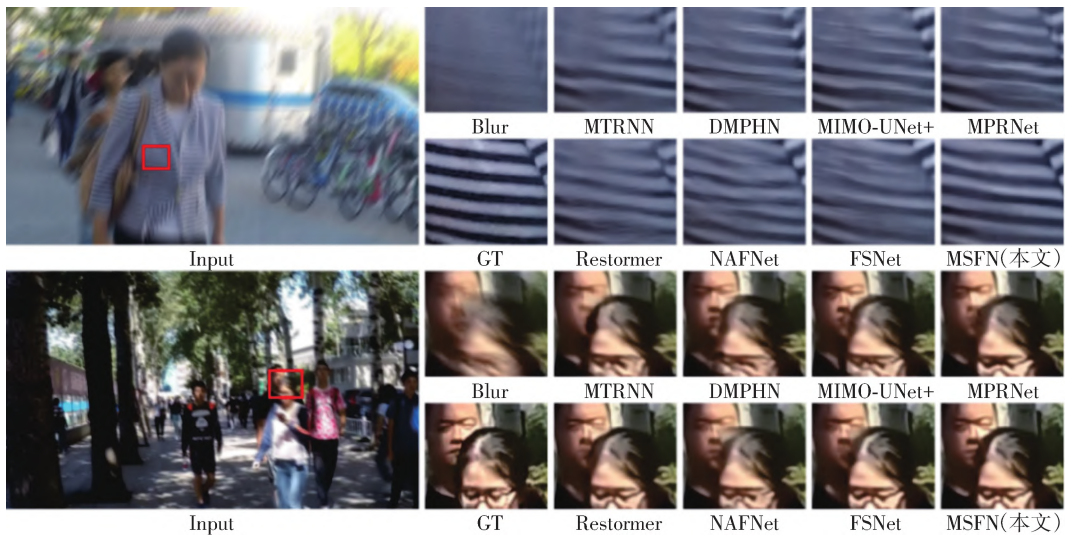


图5 HIDE数据集上的实验结果

图4展示了在GoPro数据集上,本研究提出的方法与其他方法之间的视觉比较。实验结果表明,MSFN模型能够更准确地恢复叶片的纹理细节,同时保持边缘的清晰度。在汽车图像的实验中,着重考察汽车牌照区域的去模糊性能。本文所提出的算法在汽车牌照的去模糊处理上显示出了卓越的性能,能够有效地提高牌照文字的识别性。另外在HIDE数据集的实验中,本算法在衣物纹理的去模糊效果上显著优于现有技术,并且在眼睛、鼻子和嘴巴等关键区域的清晰度和细节表现上提供卓

越的视觉质量。这进一步表明本文提出的多尺度网络在图像细节恢复方面具有显著优势。

3.4 目标识别性能增强

为全面评估MSFN在图像去模糊领域的性能,本节探讨去模糊技术对目标检测任务精度的影响。首先选用业界认可的先进算法YOLOv10进行目标检测。由于缺乏专门针对目标检测的模糊数据集,本文基于HIDE测试集进行扩展构建新数据集,并且利用AP评估去模糊效果。目标检测模型的高精度性将直接反映去模糊算法在提升图像清晰度方

面的有效性。

本研究从 HIDE 测试集中挑选 6 个最常见且具有代表性的类别构建数据集。表 2 展示使用不同去模糊技术后, YOLOv10 模型在这些类别上的 AP 及其总体平均值。与平均值排名第二的 NAFNet^[14]相比,本文提出的模型在平均值上提升 0.013。在 6 个类别上, AP 提升分别为 0.007、0.018、0.015、0.005、0.003 和 0.032。除在雨伞类别上略有不足外, MSFN 在所有类别上均取得最佳成绩, 这充分证明模型在去模糊方面的出色效果。

图 6 展示不同去模糊方法对 YOLOv10 目标检测性能的影响。结果显示, 去模糊技术能有效提高模型对行人的检测准确率, 尤其是 MSFN 模型, 显著提升检测精度和行人识别数量。这验证了本文

方法在图像清晰度恢复和细节增强方面的优势, 对增强目标检测系统的实用性和稳定性具有关键作用。

表 2 HIDE 数据集去模糊后目标检测不同类别的 AP 值

模型	行人	手提包	背包	自行车	雨伞	汽车	平均值
Blur	0.648	0.272	0.412	0.365	0.585	0.280	0.427
MTRNN ^[25]	0.891	0.676	0.744	0.683	0.837	0.601	0.739
DMPHN ^[17]	0.875	0.681	0.743	0.695	0.846	0.605	0.741
MIMO-UNet ^[28]	0.903	0.659	0.753	0.705	0.878	0.616	0.752
MPRNet ^[29]	0.917	0.703	0.780	0.712	0.901	0.606	0.770
Restormer ^[12]	0.919	0.718	0.779	0.735	0.896	0.621	0.778
NAFNet ^[14]	0.917	0.725	0.788	0.755	0.897	0.625	0.785
FSNet ^[18]	0.922	0.703	0.796	0.752	0.895	0.614	0.780
MSFN(本文)	0.924	0.743	0.803	0.760	0.900	0.657	0.798

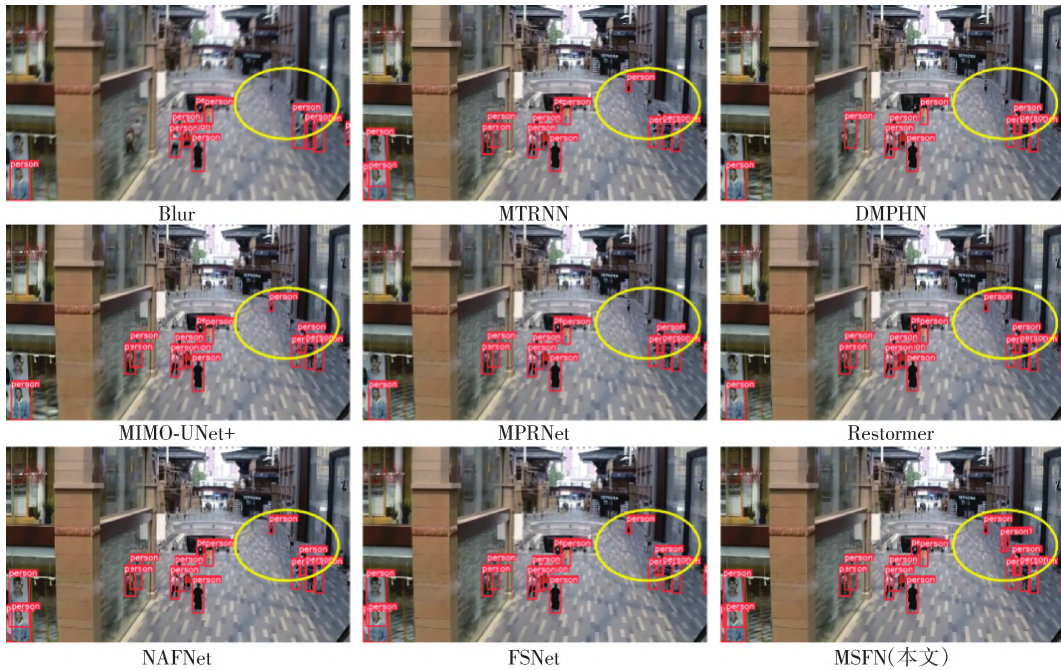


图 6 HIDE 数据集去模糊后目标检测的结果

4 结束语

本文针对现有网络在图像去模糊领域的局限性, 提出一种创新的 MSFN, 通过引入多尺度输入与输出机制, 显著提升模型特征的提取和利用效率, 从而更好地恢复图像中的边缘信息和局部模糊。MSFN 的核心贡献在于其 ADE 模块和 CSFF 模块, 这两个模块的结合不仅优化了特征提取过程, 还有效地融合多尺度信息, 从而在图像去模糊任务中实现显著性的提升。实验结果表明, MSFN 在定量和定性分析上的卓越表现, 这不仅能展示网络本身的高性能, 也验证每个模块的有效性。特别是 ADE 模块在细节纹理信息的提取上可以发挥重要作用, 而 CSFF

模块则通过有效融合不同尺度的特征, 为复杂视觉任务提供更全面的图像表示。但在实验中注意到在多尺度特征融合和上采样阶段可能会不自觉地放大图像中的噪声, 尤其是在处理高动态范围或复杂纹理的图像时, 这种噪声放大现象尤为明显。为解决这一问题, 未来的工作将探索创新的自适应滤波器, 能够根据图像内容动态调整参数, 以实现更有效的噪声抑制, 从而提高图像去模糊的整体性能。

参考文献:

[1] 李海波, 邵文泽. 图像盲去模糊综述: 从变分方法到深度模型以及延伸讨论[J]. 南京邮电大学学报(自然

- 科学版), 2020, 40(5): 84-94.
- LI Haibo, SHAO Wenzhe. Blind image deblurring: an overview from variational approaches to deep representation models and beyond[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2020, 40(5): 84-94. (in Chinese)
- [2] CUI Y N, KNOLL A. Exploring the potential of channel interactions for image restoration [J]. Knowledge-based Systems, 2023, 282: 1153-1169.
- [3] PHAM B D, TRAN P, TRAN A, et al. Blur2Blur: blur conversion for unsupervised image deblurring on unknown domains [C] //2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2024: 2804-2813.
- [4] 赵吴帆, 武文娜, 武婷婷. 各向异性的 L_0 正则化图像平滑方法[J]. 南京邮电大学学报(自然科学版), 2024, 44(4): 131-138.
- ZHAO Wufan, WU Wenna, WU Tingting. An anisotropic L_0 regularized image smoothing method [J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2024, 44(4): 131-138. (in Chinese)
- [5] XU L, ZHENG S C, JIA J Y. Unnatural L_0 sparse representation for natural image deblurring [C] //2013 IEEE Conference on Computer Vision and Pattern Recognition. 2013: 1107-1114.
- [6] PAN J S, SUN D Q, PFISTER H, et al. Blind image deblurring using dark channel prior [C] //IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1628-1636.
- [7] ZHANG M N, FANG Y Y, NI G X, et al. Pixel screening based intermediate correction for blind deblurring [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 5892-5900.
- [8] KUPYN O, BUDZAN V, MYKHAILYCH M, et al. DeblurGAN: blind motion deblurring using conditional adversarial networks [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018: 8183-8192.
- [9] GAO H Y, TAO X, SHEN X Y, et al. Dynamic scene deblurring with parameter selective sharing and nested skip connections [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 3848-3856.
- [10] ZHANG K H, LUO W H, ZHONG Y R, et al. Deblurring by realistic blurring [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2020: 2737-2746.
- [11] CUI Y N, REN W Q, CAO X C, et al. Revitalizing convolutional network for image restoration [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(12): 9423-9438.
- [12] ZAMIR S W, ARORA A, KHAN S, et al. Restormer: efficient transformer for high-resolution image restoration [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 5728-5739.
- [13] WANG Z D, CUN X D, BAO J M, et al. Uformer: a general u-shaped transformer for image restoration [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2022: 17683-17693.
- [14] CHEN L Y, CHU X J, ZHANG X Y, et al. Simple baselines for image restoration [C] //European Conference on Computer Vision. 2022: 17-33.
- [15] NAH S, HYUN K T, MU L K. Deep multi-scale convolutional neural network for dynamic scene deblurring [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 3883-3891.
- [16] TAO X, GAO H Y, SHEN X Y, et al. Scale-recurrent network for deep image deblurring [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018: 8174-8182.
- [17] ZHANG H G, DAI Y C, LI H D, et al. Deep stacked hierarchical multi-patch network for image deblurring [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 5978-5986.
- [18] CUI Y N, REN W Q, CAO X C, et al. Image restoration via frequency selection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(2): 1093-1108.
- [19] DING X H, ZHANG X Y, HAN J G, et al. Scaling up your kernels to 31×31 : revisiting large kernel design in cnns [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 11963-11975.
- [20] HORÉ A, ZIOU D. Image quality metrics: PSNR vs. SSIM [C] //20th International Conference on Pattern Recognition. 2010: 2366-2369.
- [21] KAUR R, SINGH S. A comprehensive review of object detection with deep learning [J]. Digital Signal Processing, 2023, 132: 103812.
- [22] SHEN Z Y, WANG W G, LU X K, et al. Human-aware motion deblurring [C] //IEEE International Conference on Computer Vision. 2019: 5572-5581.
- [23] ZHANG J W, PAN J S, REN J, et al. Dynamic scene deblurring using spatially variant recurrent neural networks

- [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018: 2521–2529.
- [24] KUPYN O, MARTYNIUK T, WU J R, et al. DeblurGAN-v2: deblurring (orders-of-magnitude) faster and better [C]//IEEE International Conference on Computer Vision. 2019: 8878–8887.
- [25] 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 [C] //European Conference on Computer Vision. 2020: 327–343.
- [26] SUIN M, PUROHIT K, RAJAGOPALAN A N. Spatially-attentive patch-hierarchical network for adaptive motion deblurring[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2020: 3606–3615.
- [27] PUROHIT K, SUIN M, RAJAGOPALAN A N, et al. Spatially-adaptive image restoration using distortion-guided networks [C] //IEEE/CVF International Conference on Computer Vision. 2021: 2289–2299.
- [28] CHO S J, JI S W, HONG J P, et al. Rethinking coarse-to-fine approach in single image deblurring [C]//IEEE International Conference on Computer Vision. 2021: 4641–4650.
- [29] ZAMIR S W, ARORA A, KHAN S, et al. Multi-stage progressive image restoration [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 14821–14831.
- [30] KIM K, LEE S, CHO S. MSSNet: multi-scale-stage network for single image deblurring [C]//European Conference on Computer Vision (ECCV). 2023: 524–539.

(责任编辑:李小溪)