

基于 Mamba 与频域融合的图片去模糊网络

王茂华^{1,2,3}, 高 银^{2,3}, 李 俊^{1,2,3}

(1. 福建师范大学 计算机与网络空间安全学院, 福州 350108;

2. 中国科学院 福建物质结构研究所, 福州 350108;

3. 中国科学院 泉州装备制造研究所, 福建 泉州 362200)

摘要: 针对相机抖动或物体运动引起的图像模糊问题, 提出了一种 Mamba 与频域融合的网络 MFNet; 该网络采用翻转解码器架构, 将视觉 Transformer 的非因果建模能力与 Mamba 模型框架结合, 通过融合频域信息提升图像去模糊性能; 设计了一种非因果像素交互方法, 利用注意力状态空间方程有效建模未扫描序列中语义相似像素, 并通过傅里叶变换模块缓解长距离信息衰减问题; 实验结果表明, MFNet 在 GoPro 数据集上的性能超越现有主流方法, PSNR 为 33.43 dB, FLOPs 为 66.7 G, 恢复精度更高且计算开销更低, 能够有效去除图像模糊并恢复细节。

关键词: 图像去模糊; Mamba; 频域; 注意力状态空间方程; 傅里叶变换

Image Deblurring Network Based on Mamba and Frequency Domain Fusion

WANG Maohua^{1,2,3}, GAO Yin^{2,3}, LI Jun^{1,2,3}

(1. School of computer and Cyberspace Security, Fujian Normal University, Fuzhou 350108, China;

2. Fujian Institute of Research on the Structure of Matter, CAS, Fuzhou 350108, China;

3. Quanzhou Institute of Equipment Manufacturing, CAS, Quanzhou 362200, China)

Abstract: In response to image blurring caused by camera shake or object motion, a MFNet network combining Mamba and frequency-domain fusion is proposed. The network adopts a flip-decoder architecture, combines the non-causal modeling capabilities of Vision Transformer with Mamba modeling framework, and improves image deblurring performance by fusing frequency domain information. A non-causal pixel interaction method is designed to effectively model semantically similar pixels in unscanned sequences using the attention state space equation, and alleviate the issue of long-distance information decay by Fourier transform module. Experimental results show that MFNet outperforms existing mainstream methods on GoPro dataset, achieving a PSNR of 33.43 dB, FLOPs of 66.7 G, higher recovery accuracy, and lower computational overhead, while effectively removing image blur and recovering the details of the image.

Keywords: image deblurring; Mamba; frequency domain; attention state space equation; Fourier transform

0 引言

图像去模糊是图像处理与计算机视觉领域中的重要研究方向, 旨在通过先进的算法与技术消除因相机抖动或物体运动引起的图像模糊。这类模糊现象在摄影与视频拍摄中较为常见, 会导致图像细节的丢失和整体质量的下降。因此, 运动去模糊技术被广泛应用于恢复图像的清晰度, 去除模糊伪影, 提高信噪比, 从而增强视觉

质量^[1]。

随着深度学习技术的快速发展, 图像去模糊算法逐渐趋于完善, 许多优秀算法相继出现, 与传统方法不同, 深度学习技术依赖于海量数据集进行模型训练, 通过对网络结构和参数的持续优化, 最终实现清晰图像的高效还原。

早期的图像去模糊研究^[2-5]主要依赖卷积神经网络 (CNNs) 作为核心框架, 但随着视觉 Transformer

收稿日期: 2024-12-30; 修回日期: 2025-02-12。

基金项目: 国家自然科学基金(62001452); 福建省科技计划项目(2023T3040)。

作者简介: 王茂华(2000-), 男, 硕士研究生。

通讯作者: 李俊(1979-), 男, 博士, 教授。

引用格式: 王茂华, 高 银, 李 俊. 基于 Mamba 与频域融合的图片去模糊网络[J]. 计算机测量与控制, 2025, 33(6): 264-271.

(ViTs)^[6-7]的兴起, 其卓越的性能使其逐渐成为主流方法之一。文献 [8] 提出了一种基于盲反卷积的去模糊算法, 该算法通过单幅模糊图像估计模糊卷积核和清晰图像。在模糊图像提供足够测量值的情况下, 该算法能够生成接近真实的模糊核。然而, 其仅适用于均匀模糊场景, 对于非均匀模糊图像表现欠佳。文献 [9] 提出的尺度循环网络 (SRN) 则通过在各尺度上共享网络权重, 显著减少了训练参数, 同时设计的循环模块能够有效提取多尺度信息以实现图像复原, 但随着训练模型时间变长, 算法模型存在占用空间大等不足。另一项研究中, 文献 [10] 通过在不同尺度上使用具有不同深度结构的网络来增强通过在不同尺度上使用具有不同深度结构的网络来增强这一技术, 从而优化去模糊效果, 但是不同尺度之间的关联性无法得到充分利用, 恢复效果受到限制。此外, 文献 [11] 基于生成对抗网络, 设计了 DeblurGAN 模型, 采用自编码-解码结构作为生成器, 通过自监督学习生成高质量的清晰图像。但由于图像生成特性, 导致运行时间较长, 不适用于大数据量的图片修复。

Transformer 架构凭借其强大的全局建模能力成为视觉任务中的重要工具。与传统卷积运算不同, Transformer 通过计算各特征之间的全局关联来捕获上下文信息, 展现了其在多项高级视觉任务中的潜力。在图像去模糊领域, 基于 Transformer 的算法^[6]相较于传统 CNN 模型表现出更优异的性能, 但由于其全局关联的特性, 导致计算消耗大, 为了在高效计算与全局感知能力之间找到平衡, Mamba 模型被提出并应用于图像复原任务, 并取得了显著成果。其中, MambaIR^[12]是将 Mamba 引入图像复原领域的开创性研究之一, 主要解决了局部像素信息遗忘和通道冗余这两个关键问题。自此之后, Mamba 模型逐渐在多个图像复原任务中得到广泛应用。而 MambaLLIE 则对状态空间方程进行了改进, 使其更加注重局部信息的处理, 从而更适合低光照图像增强任务。此外, Mamba 模型还在图像去雾^[13]、图像去模糊^[14]中展现了优异性能。

近期, 选择性状态空间模型 (Mamba)^[14]作为一种新型替代主干, 被引入到图像复原任务中, 并展现了显著的潜力。然而, 现有基于 Mamba 的技术存在明显的局限性, 这主要与其因果状态空间建模的特性有关。

具体而言, Mamba 通过预定义扫描规则将二维图像转化为一维序列生成 token, 而模型中的每个像素仅依赖于扫描顺序中前序像素的信息进行建模, 即典型的因果特性。这种建模方式带来了以下问题: 首先, 由于像素无法捕获后续像素的信息, 导致图像中大量有用信息未被充分利用; 其次, 为减轻信息丢失, 现有方法通常采用多方向扫描策略, 但这显著增加了计算复杂度,

尤其在高分辨率输入下表现尤为明显。同时, 多方向扫描还引入了信息冗余, 研究表明, 这种冗余性在多方向扫描中十分显著。此外, 由于 Mamba 模型中 token 之间的交互会随距离的增加快速衰减, 查询像素难以有效利用前序序列中远距离但相关的像素信息。

除了上述问题, 频域空间中蕴含的潜在信息通常在空间域中难以获取。然而, 传统基于卷积操作的深度学习方^[15-16]仅专注于空间域的特征提取, 无法充分挖掘频域信息中的模糊特征, 这在一定程度上限制了去模糊性能的进一步提升。

针对上述问题, 本文提出了 Mamba 与频域融合网络 (MFNet), 该方法利用翻转解码器架构, 结合视觉 Transformer 的非因果建模能力与 Mamba 模型框架, 并通过融合频域信息, 大幅提升图像去模糊的效率与性能。具体而言, 本文从注意力机制与状态空间模型的结合入手, 设计了一种全新的非因果像素交互方法, 并引入了注意力状态空间方程 (ASSE)。通过在 Mamba 的状态空间方程中融入提示学习机制, 模型能够实现对未来扫描序列中语义相似像素的查询。具体方法包括设计用于表示具有相似特性像素集合的提示, 并通过残差加法将提示信息整合至状态空间方程的输出中, 从而突破因果限制, 并实现单次扫描的高效建模。

此外, 为进一步结合空间域与频域的信息, 本文引入了傅里叶变换模块。该模块首先对编码特征进行傅里叶变换, 利用自学习权重矩阵提取频域信息, 随后通过逆傅里叶变换将特征重新映射回空间域。这一机制显著缓解了 Mamba 模型中长距离信息衰减的问题, 并充分利用频域信息以提升去模糊效果。

实验结果表明, MFNet 在多个公开数据集上的表现超越了现有主流方法, 不仅恢复精度更高, 同时具备更低的计算开销, 展现了卓越的性能与实际应用价值。

1 Mamba 与频域融合网络设计

基于 Mamba 与频域融合的图像去模糊网络如图 1 所示, MFNet 整体框架有一个编码器, 二个解码器构成, 从而实现端到端的去模糊处理。其核心模块为注意力状态空间模块 (ASSM) 与频域计算模块 (FCM) 组成。FCM 中集成了注意力状态空间方程 (ASSE)。ASSM 与 FCM 作为主要模块进行多层堆叠, 用于去模糊任务。

网络结构采用一编码器, 二解码器的架构, 该架构是一个简单而且有效的解决方案, 利用翻转解码器^[17]来解耦图像特征, 在不增加计算消耗的情况下, 获取正交方向的模糊特征。如图 1 (a) 所示, 在具体模块中, ASSM 利用 Mamba 模型中的 ASSE 实现长短时序记忆, 激活图像中潜在的像素信息, 获取长期依赖, 更好

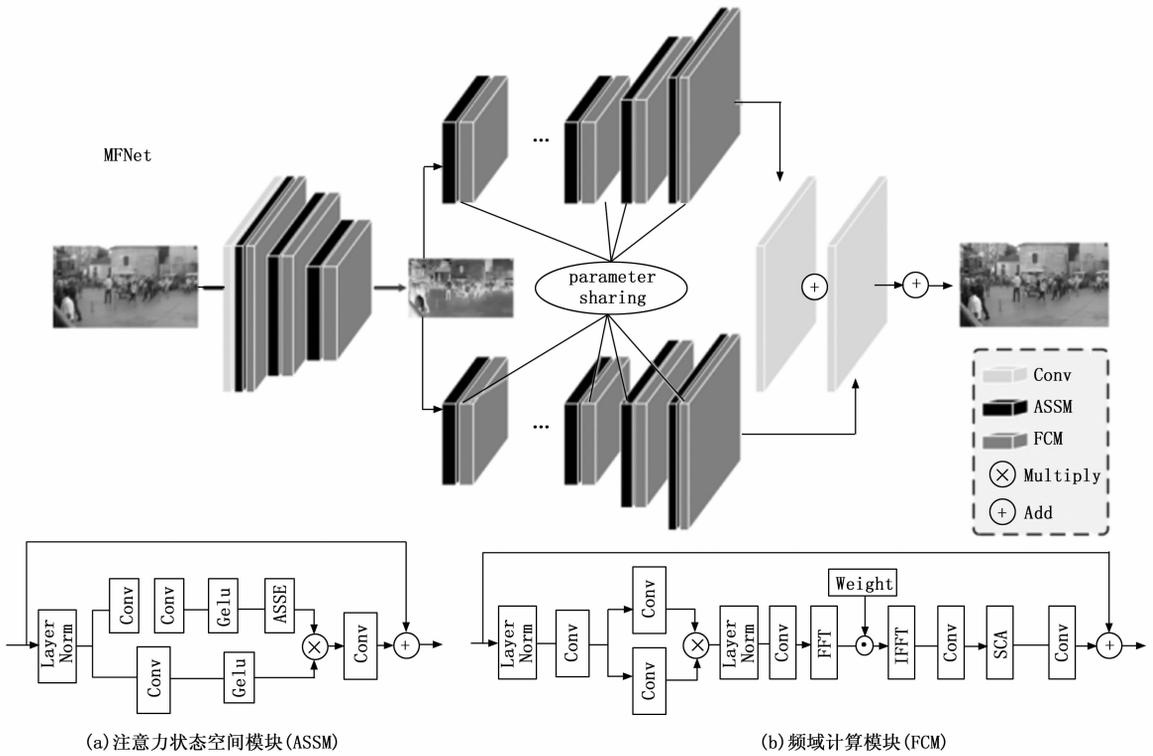


图 1 MFNet 网络架构

地恢复清晰图像。如图 1 (b) 所示, FCM 中包含快速傅里叶变换 (FFT) 与逆快速傅里叶变换 (IFFT), 利用频域中地高低频信息, 恢复清晰图像。此外, 利用简化通道注意力机制 (SCA) 实现通道级别的注意力计算, 无需激活函数, 实现非线性变换, 更好地获取图像潜在清晰信息。

1.1 翻转解码器架构

单编码器—单解码器在图像去模糊领域得到广泛应用, 将其作为其基础架构。其编码器主要用于提取图像特征。它通过一系列的卷积层和下采样操作, 逐步获取图像的不同层次的抽象特征; 而解码器通过一系列的上采样操作和卷积层, 在逐渐恢复图像空间分辨率的同时, 将编码器提取的深层特征与浅层特征进行融合, 最终生成还原图像。然而, 单编码器—单解码器也存在一定局限性。它欠缺对复杂模糊信息进行有效分解与处理的能力, 无法充分捕捉因复杂运动而在模糊图像中产生的不同方向和程度的模糊特征, 进而难以精准地还原出清晰图像。为此, 本文使用翻转解码器结构用于图像去模糊, 利用翻转解码器结构与其它模块融合, 从而巧妙实现对垂直方向和水平方向的模糊去除, 如图 2 所示。

具体来说, 通过翻转操作, 解码器能够实现对内容的正交化捕获, 例如将水平方向的注意力特征转换为垂直方向的注意力特征。在本研究中, 通过结合翻转策略



图 2 正交互补的模糊特征

和参数共享设计, 有效减少了可变卷积的使用数量, 大幅提升了网络的运行效率。首先利用编码器从输入数据中提取重要的特征信息。编码器所生成的特征中的模糊退化部分由两个解码器分别独立处理和分离, 如公式 (1) 所示:

$$Z = Enc(x) \tag{1}$$

其中: Enc 为编码器, Z 为编码后的特征信息。根据线性空间扩展的理论, 多个独立回归网络的线性重构可以扩展输出的覆盖范围。独立的回归网络能够学习彼此互补的特征, 进而显著提高整体性能。理想情况下, 这两个解码器的回归输出应相互独立, 以便实现更大的特征空间解耦能力。更具体地讲, 其中一个解码器主要提取沿主轴方向的模糊残差信息, 即退化特征中变化幅度最大的部分, 而另一个解码器则捕获其互补残差, 从而完成线性重构的过程。如公式 (2) 所示:

$$\hat{r} = Dec_{hor}(Z, \theta_{hor}) + Dec_{vec}(Z, \theta_{vec}) \tag{2}$$

其中: \hat{r} 为线性重构的结果。 Dec_{hor} , Dec_{vec} 分别为水

平方向和垂直方向的解码器。 θ_{hor} , θ_{vec} 分别为水平方向与垂直方向的网络学习参数。

1.2 注意力状态空间模块

现有的状态空间复原方法主要基于 Mamba 架构发展而来。从形式上看, Mamba 采用离散状态空间方程来建模 token 之间的交互关系:

$$\begin{aligned} h_i &= \bar{\mathbf{A}}h_{i-1} + \bar{\mathbf{B}}x_i \\ y_i &= \mathbf{C}h_i + \mathbf{D}x_i \end{aligned} \quad (3)$$

其中: $\mathbf{A} = \exp(\Delta\mathbf{A})$ 为控制矩阵, $\bar{\mathbf{B}} = (\Delta\mathbf{A})^{-1}(\exp(\mathbf{A}) - \mathbf{I}) \cdot \Delta\mathbf{B} \approx \Delta\mathbf{B}$ 作为输入矩阵, \mathbf{C} 为输出矩阵。公式 (3) 表明, 第 i 个 token 完全依赖于前 $i-1$ 个 token, 这表明状态空间模型具有因果性质。虽然这种因果性质对于自然语言处理 (NLP) 等自回归任务很有帮助, 但在图像复原任务中却并不适用。现有基于 Mamba 的方法通常采用特定的扫描策略, 将二维图像展开为一维序列, 以便通过 Mamba 进行顺序建模。然而, 第 i 个像素只能看到整个图像中有限的 $i-1$ 个像素, 无法全局利用具有相似特性的像素。因此, 现有方法通常采用多方向扫描来扩展感受野, 但这不可避免地伴随着计算复杂度的增加。为此, 本文提出了注意力状态空间模块, 以新颖的注意力状态空间方程解决感受野低且计算复杂的问题。

提出的注意力状态空间模块, 它作为网络的核心模块, 通过 Mamba 实现非因果建模。如图 1 (a) 所示, 注意力状态空间模块中, 首先使用 LayerNorm 对输入进行层归一化, 将输入张量 X 进行标准化, 并沿通道均等分为两部分 Y_1 和 Y_2 , 分别执行卷积操作。在第一个分支中, 对 Y_1 进行两层 3×3 的卷积以提取细节模糊特征, 经过 Gelu 激活函数, 最后通过 ASSE 块获取状态空间方程中的长程依赖, 如公式 (4) 所示:

$$Y_{\text{ASSE}} = \text{ASSE}\{\text{Gelu}[\text{Conv}(Y_1)]\} \quad (4)$$

对于第二个分支, 对 Y_2 使用 3×3 的获取清晰图像, 使用激活函数激活得到, 如公式 (5) 所示:

$$Y_3 = \text{Gelu}[\text{Conv}(Y_2)] \quad (5)$$

最后将 Y_{ASSE} 与 Y_3 进行矩阵点乘, 并将结果输入 1×1 的卷积层降低通道数后与输入相加生成恢复图像, 表达式如公式 (6) 所示:

$$Y = X + \text{Conv.}(Y_{\text{ASSE}} \odot Y_3) \quad (6)$$

注意力状态空间方程: 注意力状态空间方程的提出目标是修改输出矩阵 $\mathbf{C} \in \mathbb{R}^{L \times d}$ 其中 $L = HW$ 为展平后的图像序列长度, d 为 Mamba 的隐藏状态数量, 从而能够在全图范围内查询相关像素。为此, 本文提出了注意力状态空间方程, 该方程由 Mamba 的原始状态空间方程发展而来, 但具有非因果特性。如图 3 所示, 提出的 ASSE 将提示向量 (prompts) 引入 \mathbf{C} 中, 这些提示向量可以学习表示一组具有相似语义的像素, 从而补充

未见像素的缺失信息。具体来说, 首先构建提示池 $\mathbf{P} \in \mathbb{R}^{T \times d}$, 其中, T 是提示池中提示的数量。为了提高参数效率, 对 \mathbf{P} 进行低秩分解, 如公式 (7) 所示:

$$\mathbf{P} = \mathbf{M}\mathbf{N}, \mathbf{M} \in \mathbb{R}^{T \times r}, \mathbf{N} \in \mathbb{R}^{r \times d} \quad (7)$$

其中: \mathbf{N} 在不同的网络块中共享, \mathbf{M} 用于特定的网络块, r 为矩阵内秩, 满足 $r \leq \min\{T, d\}$ 。

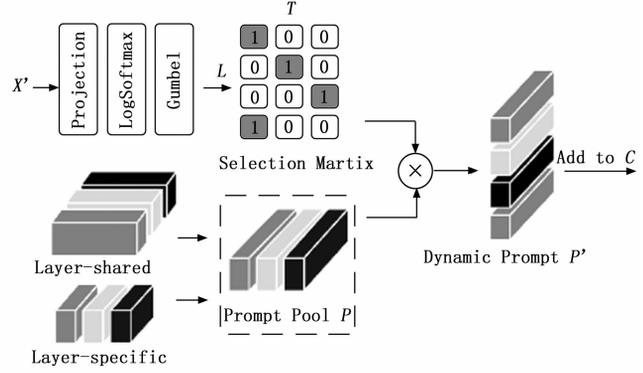


图 3 注意力空间状态方程

低秩分解的主要思想是允许不同的网络块共享相似的特征空间 (即 \mathbf{N} 是共享的), 同时通过特定的组合系数 (\mathbf{M}) 使得每个块具有独立的特性。

随后, 从提示池 \mathbf{P} 中选择 L 个实例特定的提示向量 $\mathbf{P}' \in \mathbb{R}^{T \times d}$ 。这些提示向量将被加入 \mathbf{C} 中以补充未扫描像素的信息。具体来说, 给定展平后的输入特征 $x' \in \mathbb{R}^{T \times d}$, 首先通过一个线性层将通道维度从 \mathbf{C} 投影到 \mathbf{T} , 然后通过 LogSoftmax 计算每个提示向量被 $x'_i, i = 1, 2, \dots, L$ 采样的概率。然后, 对这些对数概率应用 Gumbel-Softmax, 以实现可微分的提示选择, 并生成单热矩阵 $\mathbf{P} \in \mathbb{R}^{L \times T}$ 最终, 实例特定的提示向量通过矩阵乘法生成 $\mathbf{P}' = \mathbf{R}\mathbf{P}$ 。最后, 通过残差加法将 \mathbf{P}' 融入 \mathbf{C} 中, 从而得到最终的注意力状态空间方程:

$$\begin{aligned} h_i &= \bar{\mathbf{A}}h_{i-1} + \bar{\mathbf{B}}x_i \\ y_i &= (\mathbf{C} + \mathbf{P}')h_i + \mathbf{D}x_i \end{aligned} \quad (8)$$

通过引入提示向量, ASSE 实现了类似注意力机制的能力, 可以在全图范围内查询像素。通过引入能够表示相似像素集合的提示向量, 提出的 ASSE 有效缓解了未扫描像素的受限感知问题。此外, ASSE 仅需单方向扫描即可实现操作, 避免了现有方法中多方向扫描所带来的高计算成本和冗余。

1.3 频域计算模块

频域计算模块对每一个尺度采用傅里叶变换获取频域信息, 其中使用自学习权重矩阵, 获取频域中潜在的清晰特征, 通过不同的矩阵权重获取不同频域信息的加权结果, 如图 4 所示。随后通过简化通道注意力机制关注所需的高频信息。最后, 通过解码器得到图像潜像, 实现图像去模糊。

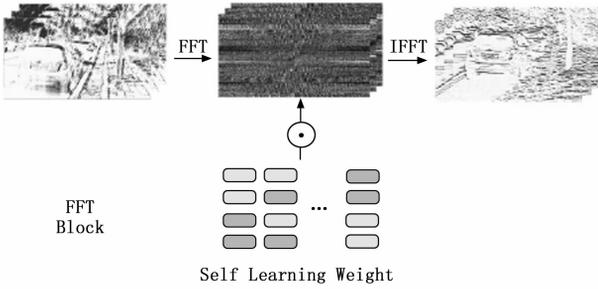


图 4 傅里叶变换块

频域计算模块首先对输入数据进行 LayerNorm 规范化处理, 并通过一个 3×3 卷积层提取初步的模糊特征信息。随后, 将特征沿着通道维度划分为两部分, 每部分的通道数为 $C/2$, 用于分离细化模糊特征。接着, 对每部分特征分别进行 3×3 的 Conv 操作, 以捕获更多细节化的模糊特性。再对两个 3×3 的 Conv 进行了门控机制设计, 从而实现了非线性特征映射的增强, 进一步丰富了模糊信息表示能力。再输入 LayerNorm 对映射后地数据进行标准化, 再进行 3×3 的卷积获取模糊特征, 送入傅里叶变换块, 获取频域信息。

最后, 进入简化通道注意力机制生成模糊残差信息。如图 5 所示, 对于输入特征, 先进行门控设计, 得到通道数为 $C/2$ 的模糊特征, 再分别执行 3×3 的 Conv, 最后利用通道注意力生成的模糊残差, 与输入进行相加的表达式如公式 (7) 所示, 并与输入进行相加操作, 整体过程可用公式 (9) 表示:

$$Out = Y + CA(X_2) \tag{9}$$

式中, CA 表示通道注意力计算, Out 表示所复原的图像。本模块通过获取频域信息和筛选高层语义信息, 模型的去模糊效果得到显著提升。

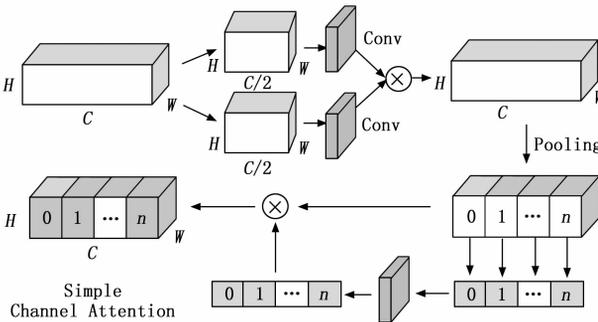


图 5 简化通道注意力机制

1.4 网络损失函数

实验中, 模型优化以峰值信噪比 (PSNR) 为主要评估指标^[18], 作为损失函数的一部分。PSNR 通过量化重建图像与原始图像之间的差异, 精确衡量恢复质量, 对重建性能具有较高的敏感度。为了降低模型在训

练过程中对异常值的过度拟合风险, 引入了基于 L_2 范数的损失函数。 L_2 范数能够有效平滑异常值对损失计算的影响, 从而减轻过拟合现象, 增强模型的泛化能力。结合这一策略, MFNet 模型采用公式 (10) 进行端到端的优化流程:

$$L_{psnr} = -10 \log_{10} \left\{ \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W [Y_{gr}(i, j) - Y_{pred}(i, j)]^2 \right\}$$

其中: H 和 W 表示图像的高度和宽度, i, j 用来表示像素坐标。 Y_{gr} 和 Y_{pred} 分别代表真值清晰图像和网络恢复图像。通过这种损失函数的设计, 能够有效评估模型在图像复原中的效果, 确保模型不仅在训练集上表现良好, 且在测试集上具有较强的泛化能力。

2 实验结果与分析

实验环境采用软件环境为 Ubuntu 20.04.5, Pytorch1.13, python3.8。硬件环境为 NVIDIA A100 显卡 4 张、CPU 为 Intel (R) Core (TM) i7-4790 CPU @ 3.60 GHz 及多光谱视觉实验平台。

2.1 数据集及训练方案

为了验证提出方法的有效性, 在两个图像去模糊公开数据集上进行了评估: GoPro^[19]数据集, 该数据集由连续帧平均生成的模糊数据集组成, 所有图像都有清晰的对应图像作为真值。其中 2 103 对图像用于训练, 1 111 对图像用于测试 (分辨率为 $1\ 280 \times 720$), 涵盖自然、室内、街道等多种场景, 被广泛用于图像去模糊模型的训练与评估; HIDE^[20], 包含真实场景中人物运动导致的模糊图像及对应的清晰图像对, 总计 8 422 对高质量图像对, 其中 2 025 对远景图像与近景图像 (分辨率为 $1\ 280 \times 720$), 全部用于测试。

训练过程中采用多种数据增强策略, 包括旋转、翻转以及随机裁剪等操作, 以提升模型的泛化能力。训练阶段使用经过增强的分辨率为 256×256 的图像对, 而在测试阶段输入图像的分辨率设置为 $1\ 280 \times 720$ 。模型训练时采用 Adam 优化器, 其参数设置为 $\beta_1 = 0.9$ 和 $\beta_2 = 0.9$ 。初始学习率设置为 10^{-3} , 并通过余弦退火调度策略逐步衰减至 10^{-6} , 以确保优化过程的平稳性。

2.2 客观对比实验

在客观分析实验中, 本章采用 PSNR 和结构相似性指数 (SSIM) 等客观指标对方法进行评价。具体来说, 峰值信噪比 PSNR、结构相似性指数 SSIM 和每秒浮点运算次数 (FLOP, floating-point operations per second) 被用作核心评价标准。通常, PSNR 值越高, 说明处理后的图像更接近原始图像, 质量更高; SSIM 值越高, 表示图像在结构信息保留方面表现更佳; 而 FLOPs 值越低, 则表明网络计算效率更高, 整体性能更优。

为了验证所提方法的优势, 本研究通过对比实验,

与多种当前最先进的方法进行客观指标的对比分析。表 1 列出了所有比较方法在 GoPro 测试集上的客观分数 PSNR、SSIM 和 FLOPs; DeepDeblur^[19]、MIMO-UNet^[21]、MAXIM^[22]、HINet^[23]、MPRNet^[24]、Restormer^[25]、Stripformer^[25]、BANet^[26]、HSDNet^[27]、DeepRFT+。与这些主流方法相比, 本方法提出的 MFNet 取得了最好的结果。PSNR 为 33.43 dB, SSIM 为 0.967, FLOPs 为 66.7 G。这些实验结果突出了本方法在处理运动模糊的有效性。

表 1 不同算法在 GoPro 数据集上的性能指标分析

算法	PSNR/dB	SSIM	FLOPs/G
DeepDeblur	30.40	0.807	110.0
MIMO-UNet	31.73	0.951	67.0
HINet	32.71	0.959	170.7
MPRNet	32.66	0.959	760.0
Restormer	32.92	0.963	140.0
MAXIM	32.86	0.961	169.5
Stripformer	33.08	0.962	170.0
BANet	32.54	0.957	264.0
HSDNet	31.68	0.950	88.0
DeepRFT+	33.23	0.963	187.0
MFNet	33.43	0.967	66.7

如表 2 所示, 本章节方法在 HIDE 数据集与当前主流方法进行比较, 本方法提出的 MFNet 分别在 PSNR 和 SSIM 方面均取得了最佳结果。PSNR 为 31.83 dB, SSIM 为 0.950。

表 2 不同算法在 HIDE 数据集上的性能指标分析

算法	PSNR/dB	SSIM	FLOPs/G
MIMO-UNet	29.99	0.930	67.0
HINet	30.32	0.932	170.7
MPRNet	30.96	0.939	760.0
Restormer	31.22	0.942	140.0
Stripformer	31.03	0.962	170.0
BANet	30.16	0.930	264.0
DeepRFT+	31.42	0.944	187.0
MFNet	31.83	0.950	66.7

2.3 消融实验

为评估不同模块对图像去模糊性能的影响, 本方法在 GoPro 数据集上进行了消融实验。实验中对 ASSM 和 FCM 模块进行了逐一移除的分析, 同时探讨了网络层数对性能的影响。表 3 显示了实验结果。在 MFNet-16 中引入 ASSM 模块后 PSNR 提升了 0.99 dB; 加入 FCM 模块后, PSNR 增加了 2.62 dB; 当同时引入 ASSM 和 FCM 时, PSNR 提升幅度达到了 3.24 dB, 充分证明了这些模块的有效性。

此外, 实验还分析了 ASSM 和 FCM 的层数对模型性能的影响, 结果表明, 当网络达到 24 层时, 去模糊效果最佳。基于此, 针对 MFNet-24 进行了进一步的消融实验。当移除 ASSM 模块后, 模型的 PSNR 降至 33.13 dB, SSIM 为 0.95; 移除 FCM 模块时, PSNR 下降至 33.05 dB, SSIM 同样下降为 0.95。如果同时移除两个模块, 模型性能进一步显著降低。这些实验结果再次表明, ASSM 和 FCM 模块对 MFNet 的整体性能至关重要, 验证了其在去模糊任务中的必要性和有效性。

表 3 消融实验结果

网络	ASSM 与 FCM 层数	ASSM	FCM	PSNR/dB	SSIM
Baseline	16	×	×	29.11	0.82
MFNet-16	16	×	√	31.73	0.90
MFNet-16	16	√	×	30.10	0.91
MFNet-16	16	√	√	32.35	0.93
MFNet-20	20	√	√	33.21	0.96
MFNet-24	24	×	×	32.37	0.93
MFNet-24	24	×	√	33.13	0.96
MFNet-24	24	√	×	33.05	0.95
MFNet-24	24	√	√	33.43	0.97

2.4 主观对比实验

在主观评价方面, 本方法首先在 GoPro 数据集上, 将所提出的方法与其他 4 种经典去模糊方法进行了对比分析。图 6 展示了不同方法在处理由垂直抖动和水平抖动引起的模糊效果时的表现差异。在图 6 (a) 中, 由于相机垂直抖动, 车牌信息变得模糊不清。相比其他方法, 本文方法更准确地还原了车牌上的数字, 同时成功消除了伪影, 展现出更高的细节恢复能力。在图 6 (b) 图中, 模糊由目标快速水平运动引起, 导致图中的文字和数字失去清晰度。其他方法在复原过程中存在模糊和虚化问题, 而本文方法则在细节和结构保留方面表现优异, 成功重现了文字的清晰边界和结构特征。在图 6 (c) 中展示了复杂背景下的人脸模糊情况。所提方法能够更完整地恢复人脸细节, 同时有效去除了头发区域的残影。在图 6 (d) 中, 远处文字模糊的问题得到了显著改善。其他方法在处理过程中未能同时保持背景和目标的清晰度, 而本文方法不仅消除了模糊, 还恢复出干净清晰的图像。

图 7 则展示了不同方法在 HIDE 数据集上的去模糊效果对比。在图 7 (a) 中, 所提方法成功去除了字体模糊, 不仅准确复原了文字细节, 还有效减少了因抖动和失真带来的干扰。此外, 红色背景在复原过程中未受黑色部分的影响, 颜色和边缘得到了良好的保留。在图 7 (b) 中, 所提方法针对近距离人脸的模糊问题, 展

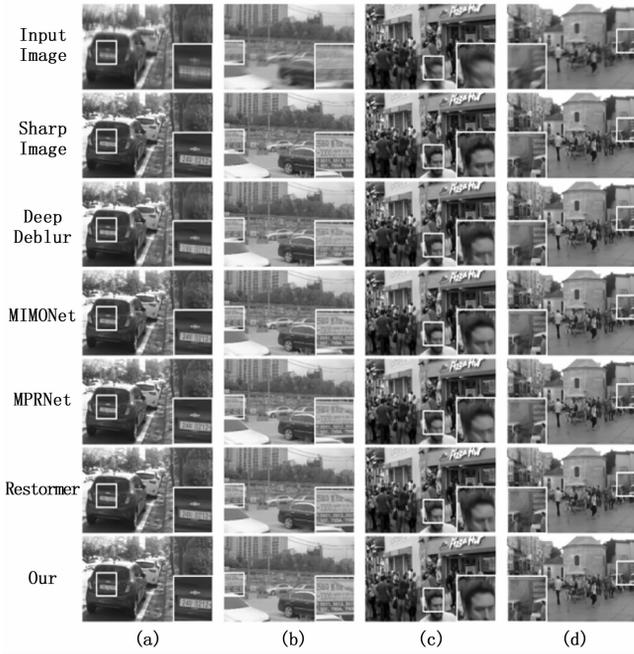


图 6 GoPro 数据集中去模糊可视化结果对比

现了出色的还原能力，不仅显著恢复了人脸细节，还在纹理和结构重建方面优于其他方法。尤其是在细节处理和纹理复原方面，本文方法表现出了显著的优势，能够更加真实地还原人脸特征。



图 7 HIDE 数据集中去模糊可视化结果对比

3 结束语

图像去模糊技术作为重要的研究方向，具有广泛的应用前景。本研究通过提出利用翻转解码器架构，结合 Mamba 模型和频域信息的 MFNet 方法，显著提升了图像去模糊的性能与效率。然而，仍需进一步优化非因果建模的计算复杂度，降低高分辨率输入场景下的资源消耗，并在实际应用中提升模型对复杂场景模糊的泛化能力。未来，MFNet 可以在摄影、视频处理及医疗图像等领域中获得更广泛的应用，为图像复原技术的发展提供新的思路。

参考文献:

- [1] ADJABI I, OUAHABI A, BENZAOUI A, et al. Past, present, and future of face recognition: a review [J]. *Electronics*, 2020, 9 (8): 1188.
- [2] XU X, PAN J, ZHANG Y J, et al. Motion blur kernel es-

timization via deep learning [J]. *IEEE Transactions on Image Processing*, 2017, 27 (1): 194 - 205.

- [3] MAO Y, WAN Z, DAI Y, et al. Deep idempotent network for efficient single image blind deblurring [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, 33 (1): 172 - 185.
- [4] ZHANG H, DAI Y, LI H, et al. Deep stacked hierarchical multi-patch network for image deblurring [C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 5978 - 5986.
- [5] 孙成阳, 毛家发, 胡亚红, 等. 基于 RGB 通道自适应的维纳滤波图像去模糊研究 [J]. *计算机测量与控制*, 2019, 27 (11): 215 - 219.
- [6] WANG Z, CUN X, BAO J, et al. Uformer: a general u-shaped transformer for image restoration [C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022: 17683 - 17693.
- [7] ZAMIR S W, ARORA A, KHAN S, et al. Restormer: efficient transformer for high-resolution image restoration [C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022: 5728 - 5739.
- [8] LEVIN A, WEISS Y, DURAND F, et al. Understanding and evaluating blind deconvolution algorithms [C] // *2009 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2009: 1964 - 1971.
- [9] TAO X, GAO H, SHEN X, et al. Scale-recurrent network for deep image deblurring [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 8174 - 8182.
- [10] KIM K, LEE S, CHO S. Mssnet: multi-scale-stage network for single image deblurring [C] // *European Conference on Computer Vision*. Cham: Springer Nature Switzerland, 2022: 524 - 539.
- [11] KUPYN O, BUDZAN V, MYKHAILYCH M, et al. Deblurgan: blind motion deblurring using conditional adversarial networks [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 8183 - 8192.
- [12] GUO H, LI J, DAI T, et al. Mambair: a simple baseline for image restoration with state-space model [C] // *European Conference on Computer Vision*. Springer, Cham, 2025: 222 - 241.
- [13] LU H, ZHANG D, HUANG B. Mamba-based all-in-one image restoration via prompt [C] // *Proceedings of the 2024 International Conference on Image Processing, Intelligent Control and Computer Engineering*, 2024: 71 - 78.
- [14] GAO H, MA B, ZHANG Y, et al. Learning enriched features via selective state spaces model for efficient image

- deblurring [C] //Proceedings of the 32nd ACM International Conference on Multimedia, 2024: 710 - 718.
- [15] MAO X, LIU Y, LIU F, et al. Intriguing findings of frequency selection for image deblurring [C] //Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37 (2): 1905 - 1913.
- [16] CHEN L, CHU X, ZHANG X, et al. Simple baselines for image restoration [C] //European conference on computer vision. Cham: Springer Nature Switzerland, 2022: 17 - 33.
- [17] JI S W, LEE J, KIM S W, et al. XYDeblur: divide and conquer for single image deblurring [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 17421 - 17430.
- [18] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13 (4): 600 - 612.
- [19] NAH S, HYUN KIM T, MU LEE 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.
- [20] SHEN Z, WANG W, LU X, et al. Human-aware motion deblurring [C] //Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 5572 - 5581.
- [21] CHO S J, JI S W, HONG J P, et al. Rethinking coarse-to-fine approach in single image deblurring [C] //Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 4641 - 4650.
- [22] TU Z, TALEBI H, ZHANG H, et al. Maxim: Multi-axis MLP for image processing [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 5769 - 5780.
- [23] CHEN L, LU X, ZHANG J, et al. Hinet: Half instance normalization network for image restoration [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 182 - 192.
- [24] ZAMIR S W, ARORA A, KHAN S, et al. Multi-stage progressive image restoration [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 14821 - 14831.
- [25] TSAI F J, PENG Y T, LIN Y Y, et al. Stripformer: strip transformer for fast image deblurring [C] //European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2022: 146 - 162.
- [26] TSAI F J, PENG Y T, TSAI C C, et al. BANet: A blur-aware attention network for dynamic scene deblurring [J]. IEEE Transactions on Image Processing, 2022, 31: 6789 - 6799.
- [27] CUI Y, REN W, CAO X, et al. Image restoration via frequency selection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 46 (2): 1093 - 1108.
- [28] 王茂华, 李 芳, 李 辉, 等. 基于 Mamba 与频域融合图像去模糊网络参数研究 [J]. 中国铁道科学, 2023, 44 (6): 153 - 161.
- [14] 张广驰, 乐文英, 庞海舰, 等. IRS 辅助认知无线移动通信网络的发射功率最小化算法 [J]. 华南理工大学学报 (自然科学版), 2023, 51 (3): 110 - 123.
- [15] 高茂成, 梅少辉, 王 楠, 等. 面向蜂群无人机自组网接入协议的多优先级动态阈值算法 [J]. 西安交通大学学报, 2023, 57 (6): 65 - 73.
- [16] 李 军, 陈士超. 无人机蜂群关键技术发展综述 [J]. 兵工学报, 2023, 44 (9): 2533 - 2545.
- [17] 肖振宇, 刘 珂, 朱立鹏. 无人机机间毫米波阵列通信技术 [J]. 通信学报, 2022, 43 (10): 196 - 209.
- [18] 宋春奇, 杨 静, 李文亮, 等. 带有前置 EDFA 卫星相干激光通信终端的信噪比分析 [J]. 半导体光电, 2023, 44 (5): 723 - 728.
- [19] 段雪峰, 李 健, 裴 鑫, 等. 基于 Arduino 的微波功率测量系统设计 [J]. 微波学报, 2023, 39 (s1): 228 - 231.
- [20] 段园园, 吉 晓, 阴万宏, 等. 一种大功率宽波段激光束散角的校准测试方法 [J]. 应用光学, 2023, 44 (2): 450 - 455.

(上接第 263 页)

- [7] 黄麒力, 胡林林, 马国武, 等. 基于量热法的大功率毫米波功率测量及校准系统设计 [J]. 强激光与粒子束, 2022, 34 (4): 71 - 74.
- [8] 张先超, 赵 耀, 叶海军, 等. 无线网络多用户干扰下智能发射功率控制算法 [J]. 通信学报, 2022, 43 (2): 15 - 21.
- [9] 沈 坤, 南晨晨, 杜保强, 等. 基于多 FPGA 的三相逆变器实时仿真系统研究 [J]. 计算机仿真, 2024, 41 (1): 129 - 134.
- [10] 宁晓哈, 雷维嘉. 能量收集多天线发送机的功率控制和天线选择 [J]. 西安电子科技大学学报, 2022, 49 (3): 83 - 92.
- [11] 陈 勇, 吴金兰, 刘焕淋, 等. 室内多输入多输出可见光通信系统接入点与功率分配优化 [J]. 光学学报, 2024, 44 (3): 28 - 37.
- [12] 孙钢灿, 吴新李, 郝万明, 等. 基于时延线阵列的毫米波 NOMA 系统混合预编码设计和功率分配 [J]. 通信学报, 2022, 43 (6): 179 - 188.
- [13] 蔺 伟, 王 芳, 李 辉. 铁路 5G-R 移动终端发射功