

# 当快速傅里叶变换遇到图像恢复变压器

Xingyu Jiang<sup>ID</sup>, Xiuwei Zhang<sup>ID</sup>, Ning Gao<sup>ID</sup>, and Yue Deng<sup>ID\*</sup>

北京航空航天大学宇航学院, 北京, Chinaydeng@buaa.edu.cn

抽象的。自然图像可能会遭受由不利的大气条件或独特的退化机制引起的各种退化现象。这种多样性使得为各种恢复任务设计通用框架变得具有挑战性。现有的图像恢复方法不是探索不同退化现象的共性，而是专注于在有限的恢复先验下修改网络架构。在这项工作中，我们首先像之前一样从频率角度回顾各种退化现象。基于此，我们提出了一种高效的图像恢复框架，称为SFHformer，它将快速傅立叶变换机制融入到Trans-former架构中。具体来说，我们设计了一种用于多尺度感受野建模的双域混合结构，其中空间域和频域分别侧重于局部建模和全局建模。此外，我们为每个频率分量设计了独特的位置编码和频率动态卷积，以提取丰富的频域特征。对 31 个恢复数据集进行了广泛的实验，涉及除雨、去雾、去模糊、去雪、去噪、超分辨率和水下/低光增强等 10 种恢复任务，证明我们的 SFHformer 超越了最新水平art 接近并实现了性能、参数大小和计算成本之间的有利权衡。该代码位于：  
<https://github.com/deng-ai-lab/SFHformer>。

关键词：图像恢复·频率特征·深度学习

## 1 简介

自然图像可能会遭受不同的退化过程（见图1a），这是由各种恶劣的天气条件（雾霾[6]、雪[15]、雨[82]和水[40]）、镜头散焦（散焦模糊[2]）引起的。])、相对运动（运动模糊[55]）和曝光不足（低光[84]），这极大地影响了各种下游视觉任务[68,105]。因此，图像恢复作为一种解决方案，旨在将低质量退化图像重建为高质量清晰图像。

---

\* 通讯作者

近几十年来被广泛讨论。然而，由于图像恢复的不稳定性固有属性，设计有效的图像恢复模型来有效处理各种退化过程具有挑战性。到目前为止，现有的恢复方法可以很好地分为两类：传统的基于先验的方法[7,28,95]和深度数据驱动的方法[37,48,98]。基于先验的模型将图像恢复视为优化问题并利用物理假设来约束解决方案空间，但在具有复杂结构的挑战性场景中可能会失败。另一方面，数据驱动模型，尤其是CNN，采用端到端的方式完成图像恢复并获得了显着的性能。最近，类似变压器的全局建模方法[74,83]已经出现，并在各种图像恢复任务上实现了最先进的(SOTA)性能。



图1：(a) 空间域的各种恢复任务。(b) PSNR与FLOPs。(c) 动机：频域视角和各种退化现象的先验。

尽管最近类似变压器的全局建模方法在图像恢复方面取得了实质性进展，但它们仍然存在一些缺点  
stantial progress for image restoration, they still suffer from several shortcom-

英格斯。首先，大多数现有的全局建模恢复方法在网络设计中几乎没有考虑各种退化过程的共同先验属性。由于恢复先验有限，这些方法可能会在某些特定的恢复任务上遇到困难。如图1b所示，SOTA方法Restormer[98]在去模糊和去雨任务上获得了相当的性能，但在低光增强任务上失败了。其次，现有的全局建模机制，例如自注意力[49]，需要大量的计算资源。虽然许多方法已经针对自注意力机制进行了优化[65, 98]，但它们的性能下降，并且无法在性能、参数大小和计算成本之间实现良好的平衡。

在上述两个问题的驱动下，我们首先从频率角度重新审视各种退化现象，并观察到频域特征本质上体现了可辨别的恢复先验，可以在网络设计中适当利用。接下来，我们从以下三个方面阐明各种退化现象的频率视角背后的基本原理（见图1c）。(1)各种退化过程的独特且紧凑的频率表示。图1c中的第一行四个子区域通过快速傅里叶变换（FFT）显示了清晰图像和退化图像的残差及其频率版本。在去雨和去模糊中，退化主要表现在高频域，其中雨纹表现为增加模式，而模糊则表现为减少模式。另一方面，在去雾和弱光下，劣化主要表现在低频域，其中雾度表现为增加模式，而弱光则表现为减少模式。此外，频域退化残差比空间域退化残差表现出更紧凑的表示，这意味着从频域角度建模退化可以提供更高的效率。(2)恢复频率强。在图1c的后续两行中，我们基于(1)的发现交换了清晰图像和退化图像的紧凑退化表示区域（参见红色圆圈）。因此，最初清晰的图像变换为其退化图像，反之，退化图像变换为退化图像。图像展示清晰版本。因此，FFT operator 具有在很大程度上从频域角度解开真实情况和退化的能力，从而可以作为各种恢复任务的强大先验。(3)高效的全局感受野特性。值得一提的是，FFT运算本质上具有全局建模特性，即改变频域中的某个频率分量就会影响整个空间域中的像素点。此外，FFT 的计算复杂度为  $O(N \log N)$ ，优于 self-attention 的计算复杂度，而 self-attention 的计算复杂度为  $O(N^2)$ 。基于上述考虑并从[97]中汲取灵感，我们将 FFT 运算符合并到令牌混合器模块中，以促进高效的全局建模，取代普通的自注意力机制。这使我们能够通过提取复杂的频率特征来有效地模拟各种退化过程。

在这项工作中，我们提出了一种类似变压器的图像恢复主干，称为 SFHformer，用于恢复频率下的各种退化过程  
dubbed SFHformer, for restoring various degradation process under frequency

事先的。具体来说，SFHformer 由两个基本模块组成：局部全局感知混合器（LGPM）和多内核ConvFFN（MCFN）。在LGPM 内部，我们设计了一种新颖的空间-频率域混合结构来取代普通的自注意力，以降低计算复杂度，其中空间域分支专注于提取局部特征，频域分支专注于捕获全局关系。此外，为了更好地区分频域分支中的每个频率分量，我们引入隐式位置编码来动态分配唯一标识，并且受[39, 89]的启发，我们设计了频率动态卷积来捕获定制的频率特征。对于MCFN，我们引入了多尺度表示学习[76]，通过多核感受野聚合局部和全局特征。MCFN中的多内核机制可以从不同角度对上下文交互进行建模，并更好地保留细节。为了验证所提出模型的有效性，我们进行了全面的实验，并在 31 个基准数据集上展示了 SFHformer 在一系列图像恢复任务中的最先进性能，包括图像去雨、图像去雾、图像去雪、图像雨滴去除、运动去模糊、单-图像散焦去模糊、图像去噪、图像超分辨率、水下图像增强和弱光图像增强。我们将主要贡献总结为三个方面：

- 我们提供了频域视角和先验来审查各种退化过程，这可以用作图像恢复网络设计的新见解。
- 我们提出 SFHformer 作为各种图像恢复任务的新型高效骨干，它提取局部-全局通过空间频率混合域来提取特征，并引入多尺度表示学习来聚合它们。
- 我们证明 SFHformer 超越了现有的最先进方法，并在 31 上保持了性能、参数大小和计算成本之间的良好平衡。10 个不同恢复任务中的图像恢复基准。

## 2 相关作品

图像恢复。一直以来，旨在将降级图像恢复为清晰图像的图像恢复已经被讨论了很长时间。为了解决这个问题，现有的图像恢复方法可以分为两类：基于先验的方法[7,28,95]和数据驱动的方法[11,56,85,98]。大多数基于先验的方法都采用新颖的物理假设来约束解决方案空间。例如，他等人。[28]提出了图像去雾的暗通道先验。易等人。[95]提出了图像去雨的残留通道先验。这些基于先验的方法在特定场景中显示出良好的统计特性，但在物理假设不成立的现实世界图像中很容易失败。

physical assumptions do not hold.

最近，数据驱动的图像恢复方法，特别是 CNN，被提出来克服传统的基于先验的方法的缺点。受益于新颖的架构设计（例如残差[29]和密集块[46]）或先进的机制（例如注意力机制）[64]和多阶段机制[22]），这些基于CNN的方法以端到端的方式实现了令人印象深刻的性能。最近，Vision Transformer [49] 最初是为自然语言任务中的序列处理而开发的，在低级视觉社区中受到了欢迎。与 CNN 不同，Vision Transformer 从全局建模的角度捕获长程依赖关系。各种基于变压器的算法[44,65,98]已经验证了全局建模对于图像恢复的重要性，并在图像去雾[38]、图像去雨[91]、运动去模糊[55]和低光图像增强方面取得了优异的性能[ 84]。然而，这些采用自注意力机制来实现全局建模的Transformer 架构面临着巨大的计算成本。因此，在我们的工作中，仔细提出了频域操作作为进行有效全局建模的替代策略。

频率特征。许多算法[17,67,75,107]已经被开发出来通过聚合频域特征来解决各种计算机视觉问题。早期，频率特征用于高级视觉任务来表示非局部关系。例如，Chi 等人。[17]提出快速傅里叶卷积来直接提取频域特征，Rao等人。[67]提出了全局滤波器网络来控制图像分类中不同频率分量的特征。最近，随着更多研究[11,16,77,83]验证全局建模对于低级视觉任务也有效，频域特征在图像恢复任务中显示出巨大的潜力。例如，苏沃洛夫等人。[75]介绍了用于图像修复的快速傅立叶卷积，Cui 等人。[22]利用频率选择机制来选择信息最丰富的频率进行恢复。与现有的频域恢复方法不同，我们的工作亮点在以下几个方面：(1) 从频率表示和退化过程中获得更深层次的动机。许多现有模型主要基于其较高的全局建模效率而采用基于频率的模型，通常没有对恢复任务进行全面的分析。在我们的研究中，我们对各种退化过程中的频率表示进行了详细检查，并证明频率有效地充当了恢复的稳健先验。通过详细的可视化和分析，我们解释了“为什么频率特征对于帮助恢复非常有效？”，这为未来的研究奠定了基础。(2) 频率全局建模的更直接的实现。当前的模型通过各种方法实现频率全局建模：一些作为自注意力机制的补充[60]，一些将卷积定理纳入自注意力[35]，而另一些则通过门控使用频率滤波器[41, 67]。受MetaFormer的启发，我们凭经验证明频率操作可以有效地替代自注意力，并引入“动态频率卷积”通过可学习的参数直接提取频率特征，而不是过滤。

quency features by learnable parameters, rather than filtering.

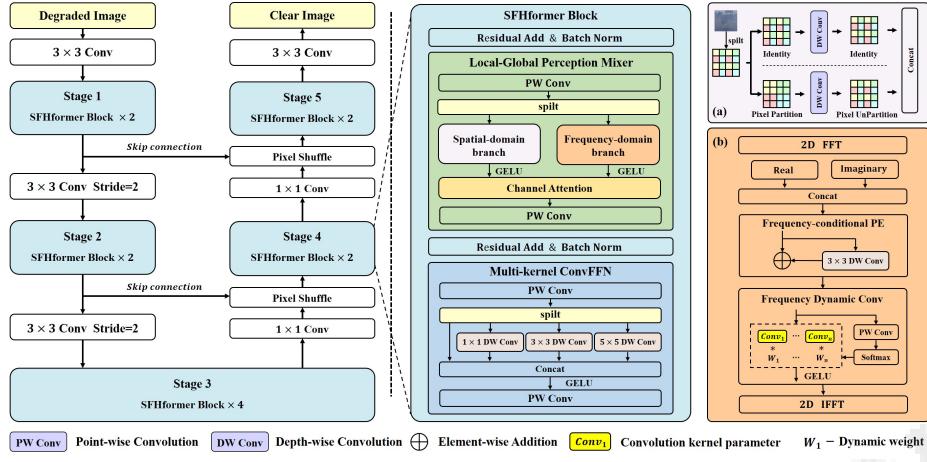


图2：拟议的SFHformer概述。(a) 空间域分支的细节和(b) 频域分支的细节。

### 3 方法

所提出的 SFHformer 采用五级的分层编码器-解码器结构：两级编码器（第一级和第二级）、瓶颈（第三级）和两级解码器（第四级和第五级）。在本节的剩余部分中，我们将首先介绍 SFHformer 的整体架构（见图2），然后描述所提出的 SFHformer 块的核心组件：(a) 局部全局感知混合器 (LGPM) 和 (b) 多内核ConvFFN (MCFN)。

#### 3.1 整体架构

如图2所示，给定退化图像  $x \in R^{H \times W \times 3}$ ，SFHformer 首先应用  $3 \times 3$  卷积来提取低级特征  $F_0 \in R^{H \times W \times C}$ ；其中  $H \times W$  表示空间维度， $C$  是通道数。接下来，浅层特征  $F_0$  依次通过5级分层编码器-解码器结构。在每个阶段之间，空间分辨率通过编码器中的下采样操作 ( $3 \times 3$ , 步长2卷积层) 逐渐降低，然后通过解码器中的上采样操作（像素洗牌）增加。在每个阶段内，输入特征  $F_0$  通过  $N$  SFH 个块来提取多尺度潜在特征  $\{F_{il}, i = 1, \dots, 5\} \in \{R^{n \times nC}, n = 1, 2, 4\}$ 。然后，为了保持用于恢复的结构和纹理特征，低级潜在特征  $F_{il}, i = 1, 2$  通过跳跃连接与高级潜在特征  $F_{il}, i = 3, 4$  连接。最后，将  $3 \times 3$  卷积层应用于最终的潜在特征  $F_{5l}$ ，生成残差图像  $s \in R^{H \times W \times 3}$ ，通过  $x = x + s$  获得清晰图像  $x$ 。在接下来的两小节中，我们将介绍两个基本模块的配置：LGPM 和 MCFN。

LGPM and MCFN.

### 3.2 局部-全局感知混合器

正如其名称所示，LGPM被配置为局部-全局建模结构，该结构由两个关键分支组成：用于局部感知的空间分支和用于全局感知的频率分支。如图2所示，LGPM首先应用PW卷积 $f_{pw}$ 将输入特征 $F_0 \in R^{H \times W \times C}$ 的通道维度加倍。接下来，增加的特征 $F_1 \in R^{H \times W \times 2C}$ 被分为两部分： $F_{sp} \in R^{H \times W \times C}$ 和 $F_{fr} \in R^{H \times W \times C}$ ，随后经过空间域分支和频域分支，分别。然后，我们采用通道注意操作来维护通道级特征聚合。最后，应用另一个PW Conv来获得简化特征 $F_2 \in R^{H \times W \times C}$ 。空间域分支和频域分支的实现细节将在以下段落中描述。空间域分支：空间域分支侧重于捕获像素级局部和区域相关性的空间特征。具体来说，如图2a所示，输入特征 $F_{sp}$ 首先被分为两个数据流： $F_{1sp}$ 和 $F_{2sp}$ ，其中 $F_{1sp}$ 保持不变，通过DW卷积 $f_{dw}$ 提取局部特征，另一个 $F_{2sp}$ 执行使用半全局感受野的像素分区来提取区域特征。在实践中，我们采用扩张卷积 $f_{dc}$ 来实现区域感受野以提高效率。最后，我们聚合两个数据流的特征，产生深层空间特征 $F_{dsp} \in R^{H \times W \times C}$ 。上述描述可以表示为：

$$F_{sp}^d = \text{Concat } f_{dw}(F_{sp}^1), f_{dc}(F_{sp}^2) \quad (1)$$

频域分支：频域分支专注于捕获频率特征，以便围绕整个图像进行全局建模。在这里，我们来看看傅里叶变换，它广泛用于分析图像的频率特性。给定图像 $y \in R^{H \times W \times C}$ ，快速傅立叶变换(FFT)  $F$ 将其转换到频率空间作为复分量  $F(y)$ ，表示为：

$$F(y)(u,v)=\frac{1}{\sqrt{HW}} \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} y(h,w) e^{-j2\pi(\frac{h}{H}u+\frac{w}{W}v)} \quad (2)$$

其中  $u$  和  $v$  是傅里叶空间的坐标。 $F^{-1}(y)$  表示快速傅立叶逆变换(IFT)。正如 FFT 公式 Eq.2 所示，频域中的各个分量对应于空间域中的特定像素集，因此需要对每个频率分量进行不同的处理。因此，为了有效地提取各种退化过程中的独特频率特征，我们引入了两个核心组件。首先，我们实现频率条件位置编码(FCPE)，它为每个频率分量分配一个不同的标识。其次，我们引入了频率动态卷积(FDC)，允许基于每个频率分量的输入进行灵活的建模。这些设计利用定制的处理方法来有效地利用每个频率分量的独特特性。

frequency component.

具体来说，如图2b所示，输入特征  $F_f r$  首先经过2D FFT  $F$  以获得复特征：频域中的实部  $F_R \in R H \times W \times C$  和虚部  $F_I \in R H \times W \times C$ 。接下来，与分别从实部和虚部提取特征不同，我们在通道维度上聚合实部  $F_R$  和虚部  $F_I$ ，产生联合特征  $F_J \in R H \times W \times 2C$ 。然后， $F_J$  逐步通过两个关键模块：FCPE 和FDC。在FCPE中，通过DWConv  $f_{dw}$  建模的位置编码以残差的形式合并到  $F_J$  中以获得特征  $F_P E$

$\epsilon R H \times W \times 2C$ 。FCPE可以写成如下：

$$F_J^{PE} = F_J + f_{dw}(F_J) \quad (3)$$

在FDC中，我们进行PW Conv  $f_{pw}$  和softmax运算来获取一系列可学习卷积参数 ( $conv_1, \dots, conv_n$ ) 的动态权重 ( $w_{u,v1}, \dots, w_{u,vn}$ )。因此，傅立叶空间中  $u, v$  坐标处的频率动态卷积  $f_{u,vf dc}$  可以写成如下：

$$\tilde{f}_{fd}^{u,v} = \sum_{i=1}^n w_i^{u,v} \cdot conv_i \quad (4)$$

最后，我们应用 2D IFFT 将调制的频率特征变换回空间域。

### 3.3 Multi-kernel ConvFFN

前馈网络 (FN) 是 Trans-formers 中广泛采用的有效模块，但之前的研究 [83, 98] 表明，标准 FN 利用本地连接进行图像恢复任务的能力有限。为了解决这一限制，我们引入多尺度表示学习[76]，它将高维特征分成多个部分，以从不同的感受野中提取局部关系。如图 2 所示，MCFNfirst 利用 PW 卷积  $f_{pw}$  将输入特征的维度  $F_1 \in R H \times W \times C$  加倍。然后，我们引入多核卷积  $f_{MC}$  将加倍特征  $F_2 \in R H \times W \times 2C$  分割到多个头中，并且在每个头中，我们应用各种内核大小的 DW 卷积来提取局部信息。最后，在连接所有多核特征后，应用另一个PW卷积  $f_{pw}$  以获得缩减特征  $F_3 \in R H \times W \times C$ 。整个操作可以表示为：

$$F_3 = f_{pw} \left( \sigma \cdot \left( f_{MC} \left( f_{pw}(F_1) \right) \right) \right) \quad (5)$$

其中  $\sigma$  是 GELU 激活函数，在多核卷积  $f_{MC}$  之后使用以引入非线性。

Convolution  $\tilde{f}_{MC}$  to introduce the non-linearity.

### 3.4 损失函数

受[20, 22]的启发，我们在优化流程中引入了与频域分支保持一致的双域损失，如式6所示，它由两部分组成：空间域损失和频域损失。

$$L = \|\hat{I} - G\|_1 + \lambda \|\mathcal{F}(\hat{I}) - \mathcal{F}(G)\|_1 \quad (6)$$

其中，采用L1损失来约束预测图像 $\hat{I}$ 与地面真实图像 $G$ 相似，并且将 $\lambda$ 设置为0.1以平衡双域学习。

## 4 实验与分析

为了评估我们的 SFHformer 有效性，我们对常见的图像恢复任务进行了广泛的实验，包括图像去雨 [25, 91]、图像去雾 [6, 38]、图像去雪 [15]、图像去模糊 [2, 55]、图像去噪 [1]，图像超分辨率[3]，水下图像增强[40]和低光图像增强[84,93]。在表格中，所评估方法的最佳和次佳质量分数被突出显示并带有下划线。

### 4.1 实验设置

数据集。我们采用RESIDE[38]、O-HAZE[6]、NH-HAZE[5]和DENSE-HAZE[4]进行去雾；用于除雨的 Rain200H [91]、Rain200L [91]、DDN-Data [25]、DID-Data[102] 和 SPA-Data [82]；Raindrop [63]用于去除雨滴；用于除雪的 CSD[15]、SRSS [14] 和 Snow100K [48]；UIEB [40] 和 LSUI [61] 用于水下增强；LOL-v1 [84]、LOL-v2 [93] 和 FiveK [9] 用于低照度增强；GoPro [55]、HIDE [73] 和 RealBlur [72] 用于运动去模糊；DPDD [2] 用于单图像散焦去模糊；SIDD [1] 用于图像去噪；DIV2k [3]、Set5 [8]、Set14 [101]、B100 [52]、Urban100 [30] 和 Manga109 [53] 用于高效图像超分辨率。实现细节。AdamW 优化器 [51] 的  $\beta_1$  和  $\beta_2$  分别等于 0.9 和 0.999，用于训练 SFHformer。初始学习率设置为 10<sup>-3</sup>。我们采用余弦退火策略 [50] 来训练模型，其中学习率从初始学习率逐渐降低到 10<sup>-6</sup>。所有实验均由 PyTorch [59] 1.7.1 和四个 NVIDIA 3090 GPU 实现。为了公平比较，所有模型的 FLOP 均在 256×256 输入分辨率上计算。每个恢复任务的更多实施细节可以在补充材料中找到。

### 4.2 实验结果

由于篇幅有限，我们选择有代表性的结果进行展示，更全面的结果（例如单图像散焦去模糊、图像去噪和高效图像超分辨率）和可视化可以在补充材料中详细找到。

found in supplementary material.

**Table 1:** Quantitative evaluations on the synthetic and real-world dehazing.

Method	ITS [38]		OTS [38]		O-HAZE [6]		NH-HAZE [5]		DENSE-HAZE [4]		Overhead #Param. FLOPs	
	SOTS-indoor		SOTS-outdoor		PSNR↑ SSIM↑		PSNR↑ SSIM↑		PSNR↑ SSIM↑			
	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑		
(TIP'16)DehazeNet [10]	19.82	0.8210	24.75	0.9271	17.57	0.77	16.62	0.52	13.84	0.43	0.009M 0.581G	
(ICCV'17)AOD-Net [37]	20.51	0.8164	24.14	0.9203	15.03	0.54	15.40	0.57	13.14	0.41	0.002M 0.115G	
(ICCV'19)GridDehazeNet [46]	32.16	0.9845	30.86	0.9827	22.11	0.71	13.80	0.54	-	-	0.956M 21.49G	
(CVPR'20)MSBDN [23]	33.67	0.9856	33.48	0.9824	24.36	0.75	19.23	0.71	15.37	0.49	31.35M 41.54G	
(AAAI'20)FFA-Net [64]	36.39	0.9894	33.57	0.9842	22.12	0.77	19.87	0.69	14.39	0.45	4.456M 287.8G	
(CVPR'22)DeHamer [27]	36.63	0.9881	35.18	0.9860	24.64	0.77	20.66	0.68	16.62	0.56	132.45M 48.93G	
(ECCV'22)PMNet [94]	38.41	0.9900	34.74	0.9850	24.64	0.83	20.42	0.73	16.79	0.51	18.90M 81.13G	
(TIP'23)Dehazeformer [74]	38.46	0.9940	34.29	0.9830	25.13	0.77	19.11	0.66	-	-	4.634M 48.64G	
(ICCV'23)FocalNet [20]	40.82	0.9960	37.71	0.9950	25.50	0.94	20.43	0.79	17.07	0.63	3.74M 30.63G	
(CVPR'23)C <sup>2</sup> PNet [106]	42.56	0.9954	36.68	0.9900	-	-	-	-	16.88	0.57	7.17M 460.9G	
(ICCV'23)MB-TaylorFormer [65]	42.64	0.9940	38.09	0.9910	25.31	0.78	-	-	16.44	0.57	7.43M 88.1G	
(Ours)SFHformer	<b>43.03</b>	<b>0.9966</b>	<b>38.83</b>	<b>0.9951</b>	<b>25.81</b>	<b>0.94</b>	<b>20.73</b>	<b>0.80</b>	<b>17.84</b>	<b>0.68</b>	3.87M 26.59G	



图 3: 图像去雾的定性结果。请放大以便更好地比较。

除雾。在表 1 中, 我们的方法在所有五个合成和真实去雾数据集上的 PSNR 和 SSIM 方面均实现了最佳性能。此外, 值得注意的是, 我们的方法优于之前 SOTA MB-TaylorFormer [65], 在 SOTS [38] 室内和室外设置上的 PSNR 分别为 0.39 dB 和 0.74 dB, 仅高出 52.1% #Param。和 30.2% 的失败率。如图 3 所示, 与最先进的方法相比, 我们的方法具有更好的色彩恢复和局部全局级别的雾度去除效果。在表 2 中, 我们的方法在所有五个合成和真实去雨数据集上在 PSNR 和 SSIM 方面实现了最佳性能。令人印象深刻的是, 我们的方法以 0.62 的成绩优于之前的 SOTA DRSformer [16]

**Table 2:** Quantitative evaluations on the synthetic and real-world deraining.

Method	Rain200L [91]		Rain200H [91]		DID-Data [102]		DDN-Data [25]		SPA-Data [82]		Overhead #Param. FLOPs	
	PSNR↑ SSIM↑		PSNR↑ SSIM↑		PSNR↑ SSIM↑		PSNR↑ SSIM↑		PSNR↑ SSIM↑			
	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑						
(ECCV'18)RESCAN [43]	36.09	0.9697	26.75	0.8353	33.38	0.9417	31.94	0.9345	38.11	0.9707	0.150M 32.12G	
(CVPR'19)PReNet [69]	37.80	0.9814	29.04	0.8991	33.17	0.9481	32.60	0.9459	40.16	0.9816	0.169M 66.25G	
(CVPR'20)MSPFN [33]	38.58	0.9827	29.36	0.9034	33.72	0.9550	32.99	0.9333	43.43	0.9843	20.89M 595.5G	
(CVPR'20)RCDNet [80]	39.17	0.9885	30.24	0.9048	34.08	0.9532	33.04	0.9472	43.36	0.9831	2.958M 194.5G	
(CVPR'21)MPRNet [100]	39.47	0.9825	30.67	0.9110	33.99	0.9590	33.10	0.9347	43.64	0.9844	3.637M 548.7G	
(ICCV'21)SPDNet [96]	40.50	0.9875	31.28	0.9207	34.57	0.9560	33.15	0.9457	43.20	0.9871	2.982M 96.29G	
(CVPR'22)Uformer [83]	40.20	0.9860	30.80	0.9105	35.02	0.9621	33.95	0.9545	46.13	0.9913	20.60M 41.09G	
(CVPR'22)Restormer [98]	40.99	0.9890	32.00	0.9329	35.29	0.9641	34.20	0.9571	47.98	0.9921	26.10M 141.0G	
(TPAMI'22)IDT [86]	40.74	0.9884	32.10	0.9344	34.89	0.9623	33.84	0.9549	47.35	0.9930	16.39M 58.44G	
(CVPR'23)DRSformer [16]	41.23	0.9894	32.18	0.9330	35.38	0.9647	34.36	0.9590	48.53	0.9924	33.70M 242.9G	
(Ours)SFHformer	<b>41.85</b>	<b>0.9908</b>	<b>32.33</b>	<b>0.9351</b>	<b>35.44</b>	<b>0.9655</b>	<b>34.38</b>	<b>0.9594</b>	<b>50.11</b>	<b>0.9942</b>	7.63M 50.59G	



Fig. 4: Qualitative Results on image deraining. Please zoom in for better comparison.

rain100L [91] 和 SPA-Data [82] 上的 PSNR dB 和 1.58 dB 仅降低了 22.6%#Param。和 20.8% 的失败率。如图4所示，我们的方法在颜色和纹理方面恢复了与真实情况最相似的视觉质量。

表 3：运动去模糊的定量评估。 表 4：低光增强的定量评估。

Method	GoPro [55]		HIDE [73]		Method	LOL-v1 [84]		LOL-v2-real [93]		LOL-v2-syn [93]		Overhead #Param. FLOPs
	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑		PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	
(CVPR'20)DBGAN [104]	31.10	0.942	28.94	0.915	(CVPR'19)DeepUPE [81]	14.38	0.446	13.27	0.452	15.08	0.623	1.02M 21.10G
(ECCV'20)MT-RNN [58]	31.15	0.945	29.15	0.918	(AAAI'20)RF [36]	15.23	0.452	14.05	0.450	15.97	0.632	21.54M 46.23G
(CVPR'19)DMPHN [103]	31.20	0.940	29.09	0.924	(CVPR'20)DeepLPF [54]	15.28	0.473	14.10	0.480	16.02	0.587	1.77M 5.86G
(ICCV'21)SPAHI [62]	32.06	0.953	30.29	0.931	(CVPR'21)IPT [12]	16.27	0.504	19.88	0.813	18.30	0.811	115.31M 6887G
(ICCV'21)MIMO-UNet+ [18]	32.45	0.957	29.99	0.936	(CVPR'21)UFFormer [83]	16.36	0.771	18.82	0.771	19.66	0.871	20.60M 41.09G
(CVPR'21)MPRNet [100]	32.66	0.959	30.96	0.936	(TIP'21)Sparse [93]	17.20	0.640	20.08	0.816	22.05	0.905	2.33M 53.26G
(CVPR'22)Restormer [98]	32.92	0.961	31.22	0.942	(TIP'21)EnGAN [34]	17.48	0.650	18.23	0.617	16.57	0.734	114.35M 61.01G
(ECCV'22)Stripformer [77]	33.08	0.962	31.03	0.940	(CVPR'21)RUAS [45]	18.23	0.720	18.37	0.723	16.55	0.652	0.003M 0.83G
(ECCV'22)MPRNet-local [19]	33.31	0.964	31.19	0.942	(CVPR'20)FIDEE [87]	18.27	0.665	16.85	0.678	15.20	0.612	8.62M 28.51G
(ECCV'22)Restormer-local [19]	33.57	0.966	31.49	0.945	(TIP'21)DRBN [92]	20.13	0.830	20.29	0.834	23.22	0.927	5.27M 48.61G
(ECCV'22)NAFNet [13]	33.71	0.967	31.31	0.943	(CVPR'22)Restormer [98]	22.43	0.823	19.94	0.827	21.41	0.830	26.10M 141.0G
(ICLR'23)SFNet [22]	33.27	0.963	31.10	0.941	(ECCV'20)MIRNet [99]	24.14	0.830	20.02	0.820	21.94	0.876	31.76M 785G
(ICML'23)IRNeXt [21]	33.16	0.962	-	(CVPR'22)SNR-Net [88]	24.61	0.842	21.48	0.849	24.14	0.928	4.01M 26.35G	
(ICCV'23)cDPMs-SA [70]	33.20	0.963	30.96	0.938	(ICCV'23)Retinexformer [11]	<b>25.16</b>	<b>0.845</b>	<b>22.80</b>	<b>0.840</b>	<b>25.67</b>	<b>0.930</b>	1.61M 15.57G
(CVPR'23)GRL [44]	33.93	0.968	<b>31.62</b>	<b>0.947</b>	(Ours)SFHformer	24.29	<b>0.862</b>	<b>23.78</b>	<b>0.872</b>	<b>25.80</b>	<b>0.937</b>	1.04M 7.75G
(Ours)SFHformer	<b>34.01</b>	<b>0.969</b>	<b>31.66</b>	<b>0.948</b>								

运动去模糊。表 3 显示了 GoPro [55] 和 HIDE [73] 上 SOTAmotion 去模糊方法的定量结果。在实践中，我们的模型仅在 GoPro 数据集上进行训练，并直接应用于 HIDE 数据集。与 GoPro [55] 上的 SOTA 模型 Restormer [98] 相比，我们的方法在 PSNR 方面实现了 1.09 dB 的增益。图 5 显示了评估模型的视觉比较，我们的方法生成了更清晰、视觉上更真实的结果，具有更多的高频细节。对于 RealBlur [72]，我们应用 [98] 设置，在 RealBlur 数据集上使用 GoPro 训练的模型。详细结果可在补充材料中找到。低光增强。如表 4 所示，我们的模型在 LOL-v1 [84] 和 LOL-v2 [93] 上的 PSNR 和 SSIM 方面超越了大多数 SOTA 方法。特别是，我们的方法在 LOL-v2-real [93] 上的 PSNR 为 0.98 dB，优于之前的 SOTA Retinexformer [11] 64.6% #Param。和 49.8% FLOPs。如图 6 所示，我们的模型实现了与 GT 最相似的对比度恢复，而其他方法往往表现出过度的亮度或黑暗。FiveK [9] 的详细结果可在补充材料中找到。



Fig. 5: Qualitative Results on image deblurring. Please zoom in for better comparison.



Fig. 6: Qualitative Results on low-light. Please zoom in for better comparison.

**Table 5:** Quantitative evaluations on underwater enhancement.

Method	L-400 [61]			U-90 [40]			#Param. FLOPs
	PSNR↑	SSIM↑	UCIQE↑	PSNR↑	SSIM↑	UCIQE↑	
(TIP'20)WaterNet [40]	23.38	0.9152	0.5729	16.31	0.7970	0.5777	1.022M 71.65G
(ICCV'21)PRWNet [31]	27.83	0.9268	0.5801	20.79	0.8231	0.5830	2.076M 14.71G
(AAAI'21)Shallow-UWNet [56]	20.56	0.7675	0.5517	18.28	0.8553	0.5517	0.2195M 20.14G
(TIP'23)U-shape Trans [61]	24.16	0.9184	0.5871	21.25	0.8432	0.5882	22.82M 24.35G
(ECCV'22)PUIE-Net [26]	21.28	0.8615	0.5887	21.38	0.8821	0.5887	0.8322M 21.84G
(TGGRS'22)URSCT-SESR [71]	29.31	0.9318	0.5890	22.72	0.9108	<b>0.6140</b>	11.26M 13.84G
(Ours)SFHformer	<b>30.18</b>	<b>0.9449</b>	<b>0.5943</b>	<b>23.54</b>	<b>0.9177</b>	0.6020	3.87M 26.59G

图7: 水下的定性结果。请放大以便更好地比较。

水下增强。如表5所示，我们的模型在 UIEB [40] 和 LSUI [61] 上的 PSNR、SSIM 和 UCIQE [90] 方面超越了大多数 SOTA 方法。特别是，我们的方法优于之前的 SOTA URSCT-SESR [71]，LSUI 和 UIEB 上的 PSNR 分别为 0.87 dB 和 0.82 dB，仅提高了 34.4% #Param。如图7所示，我们的方法重建了与地面实况最相似的水下恢复图像，并在深水区域实现了更好的恢复质量。

表6: 图像除雪的定量评估。

表7: 雨滴去除的定量评价。

Method	CSD [15]	SRRS [14]	Snow100K [48]	Overhead	Methods	Raindrop-A [63]	Raindrop-B [63]
	PSNR↑ SSIM↑	PSNR↑ SSIM↑	PSNR↑ SSIM↑	#Param. FLOPs		PSNR↑ SSIM↑	PSNR↑ SSIM↑
(TIP'18)DesnowNet [48]	20.13	0.81	20.38	0.84	30.50	0.94	15.60M -
(CVPR'18)CycleGAN [24]	20.98	0.80	20.21	0.74	26.81	0.89	7.840M 42.38G
(CVPR'22)ALL in One [42]	26.31	0.87	24.98	0.88	26.07	0.88	44.00M 12.26G
(ECCV'20)JSTASK [14]	27.96	0.88	25.82	0.89	23.12	0.86	65.00M -
(ICCV'21)HDCW-Net [15]	29.06	0.91	27.78	0.92	31.54	0.95	6.99M 9.78G
(CVPR'22)TransWeather [79]	31.76	0.93	28.29	0.92	31.82	0.93	21.90M 5.64G
(ECCV'22)NAFNet [13]	33.13	0.96	29.72	0.94	32.41	<b>0.95</b>	22.40M 12.12G
(ICCV'23)FocalNet [20]	37.18	<b>0.99</b>	31.34	<b>0.98</b>	33.53	<b>0.95</b>	3.74M 30.63G
(ICML'23)IRNetXt [21]	<b>37.29</b>	<b>0.99</b>	<b>31.91</b>	<b>0.98</b>	<b>33.61</b>	<b>0.95</b>	5.46M -
(Ours)SFHformer	<b>37.45</b>	<b>0.99</b>	<b>32.39</b>	<b>0.98</b>	<b>33.94</b>	<b>0.95</b>	3.87M 26.59G

除雪。如表6所示，我们的模型在三个广泛使用的除雪数据集上以较小的计算复杂度和参数大小实现了PSNR和SSIM的最佳性能。雨滴去除。如表7所示，我们的模型在PSNR和SSIM方面实现了最佳性能。具体来说，我们的方法在Raindrop-A上的PSNR优于之前的SOTA RainDropDiff [57]，PSNR为0.67 dB。

#### 4.3 消融研究

为了展示我们的SFHformer的有效性，我们对各个组件、损失函数和空间频域混合结构进行了各种消融研究。评估是在使用 $256 \times 256$ 图像块训练的Rain200L [91]数据集上进行的。

dataset trained on image patches of  $256 \times 256$ .

**Table 8:** Ablation Studies on Individual Components. Baseline is in the last col.

	FCPE		✓	✓	✓	✓	✓	✓	✓	✓
Individual Component	FDC	✓		✓	✓	✓	✓	✓	✓	✓
	MCFN	✓	✓			✓	✓	✓	✓	✓
	DFN [83]			✓						
	GDFN [98]				✓					
	MSFN [16]					✓				
	FFT Loss	✓	✓	✓	✓	✓	✓	✓	✓	✓
	L1 Loss	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Batch Norm	✓	✓	✓	✓	✓	✓			
	Layer Norm							✓		
Metric	PSNR↑	41.80	41.74	41.79	41.73	41.82	41.60	41.67	41.70	41.78
	SSIM↑	0.9906	0.9903	0.9904	0.9903	0.9907	0.9894	0.9901	0.9903	0.9906
										41.85
										0.9908

单独的组件。我们对三个提出的模块进行了消融研究：（1）频率条件位置编码（FCPE），（2）频率动态卷积（FDC）和（3）多内核ConvFFN（MCFN）。如表8前两行所示，与整体模型相比，删除FCPE会导致PSNR降低0.05 dB，而用staticPW Conv替换FDC会导致PSNR降低0.11 dB。此外，表8的第三行表明MCFN分别为DFN [83]、GDFN [98]和MSFN [16]实现了0.06 dB、0.12 dB和0.03 dB的PSNR增益。双域损耗的影响功能。我们采用双域损失来确保与设计模型的一致性。如表8的第四行所示，双损失比单L1/FFT损失获得了更优越的性能。批归一化与层归一化。在保持Transformer整体架构的同时，我们用FFT算子取代了自注意力机制，使我们的模型完全卷积。因此，出于标准化目的，我们不再将图像视为连续的补丁标记。由于这种结构转变，广泛应用于序列的层归一化不再适合我们的模型。相反，我们选择批量标准化。表8的第五行表明，与层规范相比，批量规范导致0.07 dB增益。

**Table 9:** Ablation Studies on Spatial-domain Branch.

structure choice		original	dual-branch with only DW Conv	single-branch with DW Conv	single-branch with dilated conv
Metric	PSNR↑	41.85	41.81	41.78	41.76
	SSIM↑	0.9908	0.9906	0.9906	0.9905

空间域分支的影响。如表9所示，我们对四种不同配置下的空间域分支进行了消融研究。从这些实验中，我们得出两个结论：首先，双分支结构优于单分支结构；其次，与单独使用单个深度卷积或单个扩张卷积相比，扩张卷积和深度卷积的组合产生了更好的结果。

表 10：混合结构的消融研究。

结构选择		空间频率混合	空间频率 spatial-frequency hybrid	only spatial	only frequency
Metric	PSNR↑		41.85	41.62	41.71
	SSIM↑		0.9908	0.9901	0.9902

空间频域混合结构的影响。如表10所示，我们设置了3种不同的结构：（1）混合结构，（2）单空间结构和（3）单频率结构。表 10 表明混合型实现了最佳性能。

#### 4.4 频域分支的定性分析

我们提供基于快速傅里叶变换的频域分支的定性分析。我们选择四个具有代表性的恢复任务，输入图像采样自 GoPro [55]（去模糊）、LOL-v2 [93]（低光）、rain200H [91]（去雨）和 ITS [38]（去雾）。这些特征是从解码器最后一个局部全局感知混合器中的第 5 阶段 SFHformer 块获得的。



图 8：我们的频域分支为四个图像恢复任务生成的学习特征及其 FFT 版本。结果是从最后一个解码器采样的。(a)退化图像; (b)习得的特征; (c)(b)的FFT结果; (d)真实情况。

为了证实所提出的频域分支建模各种恢复任务的退化模式的有效性，我们将学习到的特征及其相应的频率表示可视化，如图 8 所示。正如预期的那样，我们的模型熟练地捕获了各种退化的残余频率模式（参见红色圆圈）如图1c所示。

## 5 结论

在本文中，我们首先从频率角度重新思考各种退化过程，并提出一种高效的类似变压器的主干网SFH-former，用于图像恢复。具体来说，SFHformer由两个模块组成：局部全局感知混合器（LGPM）和多内核ConvFFN（MCFN）。在LGPM中，我们采用空间频域混合结构进行创新，其中空间域分支强调局部特征提取，而频域分支侧重于捕获全局关系。对于MCFN，结合多尺度表示学习，通过多核感受野聚合局部和全局特征。大量实验证明了我们的 SFHformer 在 31 个基准数据集上针对 10 个恢复任务的最先进性能。致谢：这项研究得到了中国国家自然科学基金委员会的支持（批准号：62031001，批准号：62325101）。

Foundation of China (Grant No.62031001, Grant No.62325101).

## References

1. Abdelhamed, A., Lin, S., Brown, M.S.: 智能手机摄像头的高质量去噪数据集。见：IEEE 计算机视觉和模式识别会议论文集。第 1692–1700 页 (2018) 92.
2. Abuolaim, A., Brown, M.S.: 使用双像素数据进行散焦去模糊。见：计算机视觉 - ECCV 2020: 第 16 届欧洲会议，英国格拉斯哥，2020 年 8 月 23-28 日，会议记录，第 X 部分，第 111-126 页。Springer (2020) 1, 93.
3. Agustsson, E., Timofte, R.: Ntire 2017 单图像超分辨率挑战：数据集和研究。见：IEEE 计算机视觉和模式识别研讨会会议记录。pp. 126–135 (2017) 94.
4. Ancuti, C.O., Ancuti, C., Sbert, M., Timofte, R.: 密集雾霾：密集雾霾和无雾图像的图像去雾基准。见：2019 年 IEEE 图像处理国际会议 (ICIP)。第 1014–1018 页。IEEE (2019) 9, 105.
5. Ancuti, C.O., Ancuti, C., Timofte, R.: Nh-haze: 具有非均匀有雾和无雾图像的图像去雾基准。见：IEEE/CVF 计算机视觉和模式识别研讨会论文集。pp. 444–445 (2020) 9, 106.
6. Ancuti, C.O., Ancuti, C., Timofte, R., De Vleeschouwer, C.: O-haze: 具有真实雾霾和无雾户外图像的除雾基准。见：IEEE 计算机视觉和模式识别研讨会会议记录。pp. 754–762 (2018) 1, 9, 107.
7. Berman, D., Treibitz, T., Avidan, S.: 使用雾线进行空气光估计。见：2017 IEEE 国际计算摄影会议 (ICCP)。第 1–9 页。IEEE (2017) 2, 48.
8. Bevilacqua, M., Roumy, A., Guillemot, C., Alberi-Morel, M.L.: 基于非负邻域嵌入的低复杂度单图像超分辨率 (2012) 99.
9. Bychkovsky, V., Paris, S., Chan, E., Durand, F.: 使用输入/输出图像对的数据库学习摄影全局色调调整。见：CVPR 2011。pp. 97–104。IEEE (2011) 9, 1110.
10. Cai, B., Xu, X., Jia, K., Qing, C., Tai, D.: Dehazenet: 用于单图像去雾的端到端系统。IEEE 图像处理交易 25(11), 5187–5198 (2016) 1011.
11. Cai, Y., Bian, H., Lin, J., Wang, H., Timofte, R., Zhuang, Y.: Retinexformer: 用于低光图像增强的基于 Retinex 的一级变压器。arXiv preprint arXiv:2303.06705 (2023) 4, 5, 1112。
12. 陈 H., 王 Y., 郭 T., 徐 C., 邓 Y., 刘 Z., 马 S., 徐 C., Xu, C., Gao, W.: 预训练的图像处理变压器。见：IEEE/CVF 计算机视觉和模式识别会议论文集。pp. 12299–12310 (2021) 1113.
13. Chen, L., Chu, X., Zhang, X., Sun, J.: 图像恢复的简单基线。见：欧洲计算机视觉会议。第 17–33 页。Springer (2022) 11, 1214.
14. Chen, W.T., Fang, H.Y., Ding, J.J., Tsai, C.C., Kuo, S.Y.: Jstasr: 基于改进的部分卷积和遮蔽效应去除的联合尺寸和透明度感知除雪算法。见：计算机视觉 - ECCV 2020: 第 16 届欧洲会议，英国格拉斯哥，2020 年 8 月 23-28 日，会议记录，第 XXI 部分 16。第 754–770 页。Springer (2020) 9, 1215.
15. Chen, W.T., Fang, H.Y., Hsieh, C.L., Tsai, C.C., Chen, I., Ding, J.J., Kuo, S.Y., et al.: 所有积雪已去除：使用单幅图像除雪算法等级制度 et al.: All snow removed: Single image desnowing algorithm using hierarchical

双树复小波表示和矛盾信道损失。见：IEEE/CVF 国际计算机视觉会议论文集。pp.4196–4205 (2021) 1, 9, 1216. Chen, X., Li, H., Li, M., Pan, J.: 学习稀疏变换网络以实现有效的图像去雨。见：IEEE/CVF 计算机视觉和模式识别会议论文集。pp. 5896–5905 (2023) 5, 10, 1317. Chi, L., Jiang, B., Mu, Y.: 快速傅立叶卷积。神经信息处理系统的进展 33, 4479–4488 (2020) 518. Cho, S.J., Ji, S.W., Hong, J.P., Jung, S.W., Ko, S.J.: 重新思考单图像去模糊中从粗到细的方法。见：IEEE/CVF 计算机视觉国际会议论文集。pp. 4641–4650 (2021) 1119. Chu, X., Chen, L., Chen, C., Lu, X.: 通过重新审视全局信息聚合来改进图像恢复。见：欧洲计算机视觉会议。pp. 53–71。Springer (2022) 1120. Cui, Y., Ren, W., Cao, X., Knoll, A.: 用于图像恢复的焦点网络。见：IEEE/CVF 国际计算机视觉会议论文集。pp. 13001–13011 (2023) 9, 10, 1221. Cui, Y., Ren, W., Yang, S., Cao, X., Knoll, A.: Irnext: 重新思考图像恢复的卷积网络设计 (2023) 11, 1222. Cui, Y., Tao, Y., Bing, Z., Ren, W., Gao, X., Cao, X., Huang, K., Knoll, A.: 用于图像恢复的选择性频率网络。见：第十一届学习表征国际会议 (2022) 5, 9, 1123. : 具有密集特征融合的多尺度增强去雾网络。见：IEEE/CVF 计算机视觉和模式识别会议论文集。页 2157–2167(2020) 1024. Engin, D., Genç, A., Kemal Ekenel, H.: 循环除雾：用于单图像除雾的增强型循环。见：IEEE 计算机视觉和模式识别研讨会会议记录。pp. 825–833 (2018) 1225. Fu, X., Huang, J., Zeng, D., Huang, Y., Ding, X., Paisley, J.: 通过深度细节网络从单张图像中去除雨水。见：IEEE 计算机视觉和模式识别会议论文集。pp. 3855–3863 (2017) 9, 1026. Fu, Z., Wang, W., Huang, Y., Ding, X., Ma, K.K.: 不确定性启发的水下图像增强。见：欧洲计算机视觉会议。第 465–482 页。Springer (2022) 1227. Guo, C.L., Yan, Q., Anwar, S., Cong, R., Ren, W., Li, C.: 具有传输感知 3d 位置嵌入的图像去雾变换器。见：IEEE/CVF 计算机视觉和模式识别会议论文集。pp. 5812–5820 (2022) 1028. He, K., Sun, J., Tang, X.: 使用暗通道先验的单图像雾霾去除。IEEE 模式分析和机器智能交易 33(12), 2341–2353 (2010) 2, 429. 何凯, 张旭, 任胜, 孙静: 图像识别的深度残差学习。见：IEEE 计算机视觉和模式识别会议论文集。pp. 770–778 (2016) 530. Huang, J.B., Singh, A., Ahuja, N.: 来自变换自我样本的单图像超分辨率。见：IEEE 计算机视觉和模式识别会议论文集。pp. 5197–5206 (2015) 931. Huo, F., Li, B., Zhu, X.: 基于高效小波增强学习的多阶段渐进细化网络，用于水下图像增强。在：诉讼程序

IEEE/CVF 计算机视觉国际会议的主席。 pp. 1944–1952(2021) 1232. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: 使用条件对抗网络进行图像到图像的转换。见： IEEE 计算机视觉和模式识别会议论文集。 pp. 1125–1134 (2017) 1233. 江凯、王志、易平、陈成、黄波、罗勇、马建、江建：用于单图像去雨的多尺度渐进融合网络。见： IEEE/CVF 计算机视觉和模式识别会议论文集。 pp.8346–8355 (2020) 1034. 蒋勇、龚X、刘丹、程勇、方成、沉X、杨建、周平、Wang, Z.: Enlightengan: 无需配对监督的深度光增强。 IEEE 图像处理交易 30, 2340–2349 (2021) 1135. Kong, L., Dong, J., Ge, J., Li, M., Pan , J.: 用于高质量图像去模糊的高效频域变压器。见： IEEE/CVF 计算机视觉和模式识别会议论文集。第 5886–5895 页 (2023)536. Kosugi, S., Yamasaki, T.: 采用强化学习控制的图像编辑软件的不成对图像增强。见： AAAI 人工智能会议记录。卷。 34, pp. 11296–11303 (2020) 1137. Li, B., Peng, X., Wang, Z., Xu, J., Feng, D.: Aod-net: 一体化去雾网络。 : IEEE 计算机视觉国际会议论文集。 pp.4770–4778 (2017) 2, 1038. Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., Wang, Z.: 单图像基准测试除雾及其他。 IEEE 图像处理汇刊 28(1),492–505 (2018) 5,9,10,1439. Li, C., Zhou, A., Yao, A.: 全维动态卷积。 arXiv preprintarXiv:2209.07947 (2022) 440. Li, C., Guo, C., Ren, W., Cong, R., Hou, J., Kwong, S., Tao, D.: 水下图像增强基准数据集和超过。 IEEE Transactions on Image Processing 29, 4376–4389 (2019) 1, 9, 1241. Li, F., Zhang, L., Liu, Z., Lei, J., Li, Z.: 多频表示增强具有视频超分辨率的特权信息。见： IEEE/CVF 国际计算机视觉会议论文集。 pp. 12814–12825(2023) 542. Li, R., Tan, R.T., Cheong, L.F.: 使用建筑搜索一次性去除恶劣天气。见： IEEE/CVF 计算机视觉和模式识别会议论文集。 pp. 3175–3185 (2020) 1243. Li, X., Wu, J., Lin, Z., Liu, H., Zha, H.: 用于单图像去雨的循环挤压和激励上下文聚合网络。见：欧洲计算机视觉会议 (ECCV) 会议记录。 pp. 254–269 (2018) 1044. Li, Y., Fan, Y., Xiang, X., Demandolx, D., Ranjan, R., Timofte, R., Van Gool,L.: 高效且明确的建模用于图像恢复的图像层次结构。见： IEEE/CVF 计算机视觉和模式识别会议论文集。 pp. 18278–18289 (2023) 5, 1145. Liu, R., Ma, L., Zhuang, J., Fan, X., Luo, Z.: Retinex 启发的展开与合作先验架构搜索低-光图像增强。见： IEEE/CVF 计算机视觉和模式识别会议论文集。 pp.10561–10570 (2021) 1146. Liu, X., Ma, Y., Shi, Z., Chen, J.: Griddehazenet: 基于注意力的多尺度网络进行图像去雾。见： IEEE/CVF 计算机视觉国际会议论文集。页 7314–7323 (2019) 5, 10

47. Liu, X., Suganuma, M., Sun, Z., Okatani, T.: 利用配对操作潜力进行图像恢复的双残差网络。见: IEEE/CVF 计算机视觉和模式识别会议论文集。页 7007–7016 (2019) 1248. Liu, Y.F., Jaw, D.W., Huang, S.C., Hwang, J.N.: Desnownet: 用于除雪的上下文感知深度网络。IEEE 图像处理学报 27(6), 3064–3073 (2018) 2, 9, 1249. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin 变换器: 使用移动窗口的分层视觉变换器。见: IEEE/CVF 计算机视觉国际会议论文集。pp. 10012–10022(2021) 3, 550. Loshchilov, I., Hutter, F.: Sgdr: 热重启的随机梯度下降.arXiv 预印本 arXiv:1608.03983 (2016) 951. Loshchilov, I., Hutter, F. : 解耦权重衰减正则化。arXiv preprint arXiv:1711.05101 (2017) 952. Martin, D., Fowlkes, C., Tal, D., Malik, J.: 人类分割自然图像数据库及其在评估分割算法和测量生态方面的应用统计数据。见: 第八届 IEEE 计算机视觉国际会议论文集。ICCV 2001.卷。2, 第 416-423 页。IEEE (2001) 953. Matsui, Y., Ito, K., Aramaki, Y., Fujimoto, A., Okawa, T., Yamasaki, T., Aizawa,K.: 使用 manga109 数据集进行基于草图的漫画检索。多媒体工具和应用程序 76, 21811–21838 (2017) 954. Moran, S., Marza, P., McDonagh, S., Parisot, S., Slabaugh, G.: Deeplpf: 用于图像增强的深度局部参数滤波器。见: IEEE/CVF 计算机视觉和模式识别会议论文集。第 12826–12835 页 (2020) 1155. Nah, S., Hyun Kim, T., Mu Lee, K.: 用于动态场景去模糊的深度多尺度卷积神经网络。见: IEEE 计算机视觉和模式识别会议论文集。第 3883–3891 (2017) 1, 5, 9, 11, 1456. Naik, A., Swarnakar, A., Mittal, K.: Shallow-uwnet: 水下图像增强的压缩模型 (学生摘要)。见: AAAI 人工智能会议记录。卷。35, 第 15853–15854 页 (2021) 4, 1257. Özdenizci, O., Legenstein, R.: 使用基于补丁的去噪扩散模型在恶劣天气条件下恢复视力。IEEE Transactions on Pattern Analysisand Machine Intelligence (2023) 1258. Park, D., Kang, D.U., Kim, J., Chun, S.Y.: 通过增量时间训练进行渐进式非均匀单图像去模糊的多时态循环神经网络。见: 欧洲计算机视觉会议。第 327–343 页。Springer (2020) 1159. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen,T., Lin, Z., Gimelshein, N., Antiga, L. 等人: Pytorch: 命令式风格的高性能深度学习库。神经信息处理系统的进展 32 (2019) 960. Patro, B.N., Namboodiri, V.P., Agneeswaran, V.S.: Spectformer: 频率和注意力是视觉转换器所需要的。arXiv preprint arXiv:2304.06446 (2023) 561. Peng, L., Zhu, C., Bian, L.: 用于水下图像增强的 U 形变换器。IEEE 图像处理汇刊 (2023) 9, 1262. Purohit, K., Suin, M., Rajagopalan, A., Boddeti, V.N.: 使用失真引导网络的空间自适应图像恢复。见: IEEE/CVF 国际计算机视觉会议论文集。第 2309–2319 页 (2021) 11

63.Qian, R., Tan, R.T., Yang, W., Su, J., Liu, J.: 用于从单个图像中去除雨滴的注意力生成对抗网络。见: IEEE 计算机视觉和模式识别会议论文集。第 2482–2491 页 (2018) 9,1264。秦X., 王Z., 白Y., 谢X., 贾H.: FFa-net: 用于单图像去雾的特征融合注意网络。见: AAAI 人工智能会议记录。卷。34, pp. 11908–11915 (2020) 5, 1065. Qiu, Y., Zhang, K., Wang, C., Luo, W., Li, H., Jin, Z.: Mb-taylorformer: Multi-通过泰勒公式扩展的分支高效变压器用于图像去雾。见: IEEE/CVF 国际计算机视觉会议论文集。pp.12802–12813 (2023) 3, 5, 1066. Quan, Y., Deng, S., Chen, Y., Ji, H.: 深度学习透过窗户看雨滴。见: IEEE/CVF 计算机视觉国际会议论文集。pp. 2463–2471 (2019) 1267. Rao, Y., Zhao, W., Zhu, Z., Lu, J., Zhou, J.: 用于图像分类的全局过滤网络。神经信息处理系统的进展 34, 980–993(2021) 568. Rawat, W., Wang, Z.: 用于图像分类的深度卷积神经网络: 综合综述。神经计算 29(9), 2352–2449 (2017) 169. Ren, D., Zuo, W., Hu, Q., Zhu, P., Meng, D.: 渐进式图像去雨网络: 一种更好的更简单的基线。见: IEEE/CVF 计算机视觉和模式识别会议论文集。pp. 3937–3946 (2019) 1070. Ren, M., Delbracio, M., Talebi, H., Gerig, G., Milanfar, P.: 用于图像去模糊的多尺度结构引导扩散。见: IEEE/CVF 国际计算机视觉会议论文集。pp. 10721–10733 (2023) 1171. Ren, T., Xu, H., Jiang, G., Yu, M., Zhang, X., Wang, B., Luo, T.: 增强型 swin-convs 变压器用于同时水下传感场景图像增强和超分辨率。IEEE Transactions on Geoscience and Remote Sensing 60,1–16 (2022) 1272. Rim, J., Lee, H., Won, J., Cho, S.: 用于学习和基准去模糊算法的真实世界模糊数据集。见: 计算机视觉 - ECCV 2020: 第 16 届欧洲会议, 英国格拉斯哥, 2020 年 8 月 23 - 28 日, 会议记录, 第 XXV 部分 16。第 184 - 201 页。Springer (2020) 9, 1173. Shen, Z., Wang, W., Lu, X., Shen, J., Ling, H., Xu, T., Shao, L.: 人类感知运动去模糊。见: IEEE/CVF 国际计算机视觉会议论文集。pp. 5572–5581 (2019) 9, 1174. Song, Y., He, Z.,qian, H., Du, X.: 用于单图像去雾的视觉变换器.IEEE Transactions on Image Treatment 32, 1927–1941 ( 2023) 2, 1075. Suvorov, R., Logacheva, E., Mashikhin, A., Remizova, A., Ashukha, A., Silvestrov,A., Kong, N., Goka, H., Park, K. , Lempitsky, V.: 具有傅立叶卷积的分辨率鲁棒大型掩模绘画。见: IEEE/CVF 计算机视觉应用冬季会议论文集。第 2149–2159 页 (2022) 576. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: 重新思考计算机视觉的初始架构。见: IEEE 计算机视觉和模式识别会议论文集。pp. 2818–2826 (2016) 4, 877. Tsai, F.J., Peng, Y.T., Lin, Y.Y., Tsai, C.C., Lin, C.W.: Stripformer: 用于快速图像去模糊的条带变换器。见: 欧洲计算机视觉会议. pp. 146–162。Springer (2022) 5, 1178. Tu, Z., Talebi, H., Zhang, H., Yang, F., Milanfar, P., Bovik, A., Li, Y.: Maxim: 多轴 mlp图像处理。见: IEEE/CVF 计算机视觉和模式识别会议论文集。第 5769–5780 页 (2022) 12

ence on Computer Vision and Pattern Recognition. pp. 5769–5780 (2022) 12

79. Valanarasu, J.M.J., Yasarla, R., Patel, V.M.: Transweather: 基于 Transformer 的恶劣天气条件退化图像恢复。见: IEEE/CVF 计算机视觉和模式识别会议论文集。 pp.2353–2363 (2022) 1280. Wang, H., Xie, Q., Zhao, Q., Meng, D.: 用于单图像除雨的模型驱动深度神经网络。见: IEEE/CVF 计算机视觉和模式识别会议论文集。 pp. 3103–3112 (2020) 1081. Wang, R., Zhang, Q., Fu, C.W., Shen, X., Zheng, W.S., Jia, J.: 使用深度照明估计的曝光不足照片增强。见: IEEE/CVF 计算机视觉和模式识别会议论文集。 pp. 6849–6857 (2019) 1182. Wang, T., Yang, X., Xu, K., Chen, S., Zhang, Q., Lau, R.W.: 空间注意力单图像除雨与高质量真实降雨数据集。见: IEEE/CVF 计算机视觉和模式识别会议论文集。 pp.12270–12279 (2019) 1, 9, 10, 1183. 王志., 存X., 包建., 周文., 刘建., 李红.: Uformer: 将军用于图像恢复的u形变压器。见: IEEE/CVF 计算机视觉和模式识别会议论文集。 pp. 17683–17693 (2022) 2, 5, 8, 10, 11, 1384. Wei, C., Wang, W., Yang, W., Liu, J.: 低光增强的深度视网膜分解。arXiv 预印本 arXiv:1808.04560 (2018) 1, 5, 9, 1185. 吴红、曲Y.、林S.、周J.、乔R.、张Z.、谢Y. , Ma, L.: 紧凑单图像去雾的对比学习。见: IEEE/CVF 计算机视觉和模式识别会议论文集。第 10551–10560 页 (2021)486. 肖建, 付新, 刘安, 吴芳, 查子建: 图像去雨变压器。 IEEE Transactions on Pattern Analysis and Machine Intelligence (2022) 10, 1287。 Xu, K., Yang, X., Yin, B., Lau, R.W.: 学习通过分解和增强恢复低光图像。见: IEEE/CVF 计算机视觉和模式识别会议论文集。 pp. 2281–2290 (2020) 1188. Xu, X., Wang, R., Fu, C.W., Jia, J.: Snr 感知的低光图像增强。见: IEEE/CVF 计算机视觉会议论文集和模式识别。第 17714–17724 (2022) 1189. Yang, B., Bender, G., Le, Q.V., Ngiam, J.: Condconv: 用于高效推理的条件参数化卷积。神经信息处理系统的进展 32 (2019) 490. Yang, M., Sowmya, A.: 水下彩色图像质量评估指标。IEEE Transactions on Image Processing 24(12), 6062–6071 (2015) 1291. Yang, W., Tan, R.T., Feng, J., Liu, J., Guo, Z., Yan, S.: 单幅图像的深度联合降雨检测和去除。见: IEEE 计算机视觉和模式识别会议论文集。第 1357–1366 页 (2017) 5, 9, 10, 11, 12, 1492。 Yang, W., Wang, S., Fang, Y., Wang, Y., Liu, J.: 基于波段表示的半监督低光图像增强: 弥合信号保真度和感知质量之间的差距。IEEE Transactions on Image Treatment 30, 3461–3473(2021) 1193. Yang, W., Wang, W., Huang, H., Wang, S., Liu, J.: 用于鲁棒低光的稀疏梯度正则化深度视网膜网络图像增强。IEEE 图像处理汇刊 30, 2072–2086 (2021) 9, 11, 14

94. Ye, T., Zhang, Y., Jiang, M., Chen, L., Liu, Y., Chen, S., Chen, E.: 图像去雾密度的感知和建模。见：欧洲计算机视觉会议。第 130–145 页。Springer (2022) 1095. Yi, Q., Li, J., Dai, Q., Fang, F., Zhang, G., Zeng, T.: 残渣通道先导的保结构除雨。见：IEEE/CVF 国际计算机视觉会议论文集。pp. 4238–4247 (2021) 2, 496. Yi, Q., Li, J., Dai, Q., Fang, F., Zhang, G., Zeng, T.: 残渣保构除雨渠道事先指导。见：IEEE/CVF 国际计算机视觉会议论文集。pp. 4238–4247 (2021) 1097. 余伟、罗明、周平、斯成、周勇、王新、冯建、严顺：Metaformer 实际上是您实现视觉所需要的。见：IEEE/CVF 计算机视觉和模式识别会议论文集。第 10819–10829 (2022) 398. Zamir, S.W.、Arora, A.、Khan, S.、Hayat, M.、Khan, F.S.、Yang, M.H.: Restormer：用于高分辨率图像恢复的高效转换器。见：IEEE/CVF 计算机视觉和模式识别会议论文集。页 5728–5739(2022) 2, 3, 4, 5, 8, 10, 11, 1399。Zamir, S.W., Arora, A., Khan, S., Hayat, M., Khan, F.S., Yang, M.H., Shao, L.: 学习丰富的特征以实现真实图像的恢复和增强。见：计算机视觉 - ECCV 2020：第 16 届欧洲会议，英国格拉斯哥，2020 年 8 月 23-28 日，会议记录，第 XXV 部分 16。第 492-511 页。Springer (2020) 11100. Zamir, S.W.、Arora, A.、Khan, S.、Hayat, M.、Khan, F.S.、Yang, M.H.、Shao, L.: 多阶段渐进式图像恢复。见：IEEE/CVF 计算机视觉和模式识别会议论文集。pp. 14821–14831 (2021) 10, 11101. Zeyde, R., Elad, M., Protter, M.: 使用稀疏表示的单图像放大。见：曲线与曲面：第七届国际会议，法国阿维尼翁，2010 年 6 月 24-30 日，修订后的精选论文 7。第 711-730 页。施普林格 (2012) 9102。张, H., 帕特尔, V.M.: 使用多流密集网络进行密度感知的单图像去雨。见：IEEE 计算机视觉和模式识别会议论文集。pp. 695–704 (2018) 9, 10103. Zhang, H., Dai, Y., Li, H., Koniusz, P.: 用于图像去模糊的深层堆叠分层多块网络。见：IEEE/CVF 计算机视觉和模式识别会议论文集。pp. 5978–5986 (2019) 11104. 张凯、罗文、钟宇、马丽、斯腾格、刘文、李宏：通过真实模糊进行去模糊。见：IEEE/CVF 计算机视觉和模式识别会议论文集。pp. 2737–2746 (2020) 11105. 赵 Z.Q.、郑 P.、徐 S.t.、吴 X.: 深度学习的目标检测：综述。IEEE Transactions on Neural Networks and Learning Systems 30(11), 3212–3232 (2019) 1106. Cheng, Y., Zhan, J., He, S., Dong, J., Du, Y.: 课程对比正则化用于物理感知的单图像去雾。见：IEEE/CVF 计算机视觉和模式识别会议论文集。第 5785–5794 页 (2023) 10107. 周明、黄杰、郭成林、李成：Fourmer：一种高效的图像恢复全局建模范式。见：国际机器学习会议。第 42589–42601 页。PMLR (2023) 5