



# 计算机科学

COMPUTER SCIENCE

## 图像去模糊算法研究综述

陈康, 林建涵, 刘元杰

引用本文

陈康, 林建涵, 刘元杰. [图像去模糊算法研究综述](#)[J]. 计算机科学, 2025, 52(11): 98-112.

CHEN Kang, LIN Jianhan, LIU Yuanjie. [Survey on Image Deblurring Algorithms](#)[J]. Computer Science, 2025, 52(11): 98-112.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

### [一种基于深度分区聚合的神经网络后门样本过滤方法](#)

Neural Network Backdoor Sample Filtering Method Based on Deep Partition Aggregation

计算机科学, 2025, 52(11): 425-433. <https://doi.org/10.11896/jsjcx.240900007>

### [面向可见光与红外多模态目标检测的对抗攻防综述](#)

Survey of Adversarial Attack and Defense for RGB and Infrared Multimodal Object Detection

计算机科学, 2025, 52(11): 349-363. <https://doi.org/10.11896/jsjcx.241200151>

### [基于多尺度层次网络的人体重建神经辐射场](#)

Neural Radiance Field for Human Reconstruction Based on Multi-scale Hierarchical Network

计算机科学, 2025, 52(11): 175-183. <https://doi.org/10.11896/jsjcx.240900141>

### [基于颜色增强的多层次特征融合图像情感识别](#)

Multi-level Feature Fusion Image Emotion Recognition Based on Color Enhancement

计算机科学, 2025, 52(11): 157-165. <https://doi.org/10.11896/jsjcx.241000016>

### [基于细粒度注意力机制的人与物体交互检测](#)

Human-Object Interaction Detection Based on Fine-grained Attention Mechanism

计算机科学, 2025, 52(11): 141-149. <https://doi.org/10.11896/jsjcx.240900113>

# 图像去模糊算法研究综述

陈康 林建涵 刘元杰

中国农业大学信息与电气工程学院 北京 100083

(chenkang@cau.edu.cn)

**摘要** 图像去模糊是计算机视觉中的经典问题,旨在从模糊的输入图像或视频中恢复出清晰的视觉信息。模糊现象通常由相机对焦不准、相机抖动或目标快速运动等因素导致。传统去模糊方法通常将该任务建模为反卷积问题,视模糊图像为清晰图像与模糊核的卷积结果,但在处理复杂或非理想的模糊类型时存在局限性。近年来,基于深度学习的回归方法取得了突破。这类方法借助卷积神经网络(CNN)和 Transformer 等架构,通过学习模糊与清晰图像间的映射关系,实现了对复杂模糊情况的有效处理,且无需对模糊核进行显式建模。同时,生成式深度学习方法如生成对抗网络(GAN)和扩散模型在去模糊领域逐渐展示出显著潜力。生成式 AI 通过构建并学习图像的细节生成过程,不仅能够有效去除模糊,还可以生成具有细腻纹理的高质量图像,在复杂模糊场景中表现出优越性能。文中首先介绍了图像模糊的特性,阐述了去模糊的常见任务类别及评估指标,然后深入探讨了去模糊模型的基本架构与训练方法,并比较分析了代表性的前沿图像去模糊模型,最后进一步探讨了图像去模糊在未来可能的研究方向。

**关键词:** 图像去模糊;深度学习;反卷积;自回归模型;生成式人工智能

**中图分类号** TP391.41

## Survey on Image Deblurring Algorithms

CHEN Kang, LIN Jianhan and LIU Yuanjie

School of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

**Abstract** Image deblurring is a classic problem in computer vision, aiming to recover sharp visual information from blurry input images or videos. Blur is often caused by factors such as camera misfocus, camera shake, or fast-moving objects. Traditional deblurring methods typically model the task as a deconvolution problem, treating the blurry image as the convolution of a sharp image and a blur kernel. However, these methods face limitations when dealing with complex or non-ideal blur types. In recent years, deep learning-based regression methods have made significant breakthroughs. These approaches leverage architectures such as Convolutional Neural Networks(CNNs) and Transformers to learn the mapping between blurry and sharp images, enabling effective handling of complex blur scenarios without explicit modeling of the blur kernel. Additionally, generative deep learning methods, such as Generative Adversarial Networks(GANs) and Diffusion models, have shown considerable potential in the deblurring field. Generative AI, by modeling and learning the image detail generation process, not only effectively removes blur but also generates high-quality images with fine textures, demonstrating superior performance in challenging blur scenarios. This paper first introduces the characteristics of image blur and outlines common deblurring tasks and evaluation metrics. It then delves into the fundamental architectures and training methods of deblurring models, providing a comparative analysis of representative state-of-the-art deblurring models. Finally, the paper explores potential future research directions in the field of image deblurring.

**Keywords** Image deblurring, Deep learning, Deconvolution, Autoregression model, Generative artificial intelligence

## 1 引言

图像去模糊是低级计算机视觉中的经典任务,目的是将模糊图像恢复成高清图像,具有广泛的应用场景,如医疗、安防监控、图像识别和图像重建等<sup>[1-11]</sup>。因此,研究图像去模糊具有重要意义。

图像模糊可能由多种因素引起,如对焦不足、相机抖动或目标快速运动。根据造成模糊的原因,图像模糊常常被分为运动模糊<sup>[12]</sup>、离焦模糊<sup>[13]</sup>、高斯模糊<sup>[14]</sup>和混合模糊<sup>[15]</sup>。图像模糊的评价指标(IQA)一般分为主观指标<sup>[16]</sup>和客观指标<sup>[17]</sup>。主观指标常常是基于人类的判断,通常不需要参考图像,而客观评价方法根据是否需要参考图像,又分为全参考、

到稿日期:2024-12-06 返修日期:2025-08-03

基金项目:国家重点研发计划(2022YFC230400401);国家自然科学基金(32272930,61807032)

This work was supported by the National Key Research and Development Program of China(2022YFC230400401) and National Natural Science Foundation of China(32272930,61807032).

通信作者:刘元杰(yjliu@cau.edu.cn)

半参考和无参考。由于主观评价指标需要耗费大量人力资源,因此实际应用中多采取客观评价指标。

非深度学习的图像去模糊方法通常将该任务公式化为反卷积问题<sup>[18]</sup>,其中模糊图像被建模为空间不变或空间变化的模糊核卷积的结果。根据是否已知模糊卷积核,非深度学习的图像去模糊方法分为非盲去模糊问题和盲去模糊问题。一些早期非盲去模糊算法采用经典的图像去卷积算法(如 Lucy-Richardson 或维纳反卷积),有时结合 Tikhonov 正则化,以恢复清晰的图像<sup>[19-24]</sup>。一些盲图像去模糊方法<sup>[25-43]</sup>则假设模糊卷积核未知,旨在同时恢复清晰图像和模糊卷积核本身。由于该任务属于病态问题,因此需通过各种附加约束进行正则化。尽管这些非深度学习在某些情况下表现良好,但在更为复杂但常见的场景(如强运动模糊)中通常效果不佳。

随着深度学习技术的不断发展,基于端到端的深度学习算法相继被提出<sup>[44-48]</sup>。通过采用卷积神经网络(Convolutional Neural Networks, CNN)和 Transformer 等高效架构,这些方法能够学习模糊图像与清晰图像之间的映射关系,从而在复杂模糊场景下实现高效处理,且无需显式建模或估计模糊核。近年来,生成式人工智能(AI-Generated Content, AIGC)的最新进展对计算机视觉领域产生了深远影响<sup>[49-51]</sup>。生成式深度学习(如扩散模型<sup>[52]</sup>、生成对抗网络<sup>[49-55]</sup>)在图像去模糊任务中表现出显著优势,其一站式生成高清图像的能力极大地推动了去模糊技术的发展。然而,深度学习的迅猛发展也使得全面了解该领域的最新进展变得日益复杂和具有挑战性。为填补这一空白,本文旨在对近期的研究成果进行系统综述,旨在为新进入该领域的研究人员提供参考。具体而言,本文将聚焦于最新的图像去模糊方法,并对其关键技术和发展趋势进行深入探讨。本文的目标是:

1) 回顾图像去模糊的基本知识,包括问题定义、模糊成因、去模糊方法、质量评价指标,以及用于性能评估的基准数据集。

2) 探讨去模糊的最新进展,重点分析基于深度学习的方法对去模糊领域的影响,并提供对现有方法的分类体系。

3) 分析图像去模糊面临的挑战,并讨论大模型和多模态学习在解决复杂模糊场景中的潜力与未来的研究方向。

本文第 2 章探讨了模糊现象的成因、去模糊技术的分类以及图像质量评估的指标体系;第 3 章详细介绍了去模糊方法,包括非盲去模糊方法和盲去模糊方法;第 4 章详细介绍了基于深度学习的去模糊方法;第 5 章介绍了用于评估去模糊性能的基准数据集并进行了去模糊性能评估;第 6 章讨论了当前研究领域面临的挑战,并展望了未来研究的发展方向;最后总结全文。

## 2 图像模糊成因及评价方法

### 2.1 问题表述

导致图像模糊的因素众多,如运动模糊、离焦模糊、高斯模糊和混合模糊,如图 1 所示。图像模糊的数学表达式如下<sup>[3]</sup>:

$$B = \Phi(I_s, k) + N \quad (1)$$

其中,  $\Phi$  为图像模糊函数,  $B$  为模糊图像,  $I_s$  为清晰图像,  $k$  为

模糊核,  $N$  为噪声。根据是否知道图像模糊函数  $\Phi$ , 去模糊方法可以分为非盲去模糊方法和盲去模糊方法。图像去模糊的目标是恢复出一张清晰的图像, 即寻找模糊函数的逆过程, 其表达式如下:

$$I_{db} = \Phi^{-1}(B, k, N) \quad (2)$$

其中,  $\Phi^{-1}$  为去模糊模型,  $I_{db}$  为去模糊图像。



图 1 常见的模糊类型

Fig. 1 Common blurry types

#### 2.1.1 运动模糊

运动模糊一般是由相机的曝光时间过长引起的,许多方法直接将运动模糊过程建模为卷积过程,并假设整个图像的模糊是均匀的。

$$B = K * I_s + N \quad (3)$$

其中,  $K$  为模糊核,  $N$  为高斯噪声。运动模糊的早期研究中常常利用边缘、纹理和图像梯度等信息来对图像进行恢复<sup>[56-60]</sup>。典型做法是先利用成像物理模型估计空间不变的模糊核,再通过卷积核的平移不变性、梯度先验和总变差最小化来恢复图像细节<sup>[61-65]</sup>。然而,在近几年的研究中,解决模糊核的不一致问题成为了热点研究方向<sup>[66]</sup>。

#### 2.1.2 离焦模糊

对于离焦模糊现象,若被拍摄场景未处于相机的焦平面,则图像的清晰度将受到显著影响。在焦点前后,光线逐渐汇聚和发散,使得点光源的影像呈现模糊状态,形成一个扩大的光斑,这种光斑被称为弥散圆。在实际场景中,人眼感知影像的清晰度与放大倍率、投影距离及观察距离密切相关。当弥散圆的直径小于人眼的分辨能力时,在一定范围内的模糊可能无法被察觉,这种难以分辨的弥散圆被定义为容许弥散圆。在焦点的前后各存在一个容许弥散圆,这两个弥散圆之间的空间距离被称为景深。当拍摄物体超出这一景深范围时,便会产生离焦模糊。这种模糊常使用点扩散函数(Point Spread Function)来建模<sup>[67]</sup>。

$$K(x, y) = \begin{cases} \frac{1}{\pi r^2}, & \text{if } (x-k)^2 + (y-l)^2 \leq r^2 \\ 0, & \text{else} \end{cases} \quad (4)$$

其中,  $(k, l)$  是点扩散函数的中心,  $r$  是模糊的半径。解决离焦模糊的相关研究有使用模糊检测或者编码孔径的方式,这些

先验知识可以通过正则化项或先验模型的形式加入到去模糊算法中,以提供额外的约束和指导图像去模糊<sup>[15,34,68-69]</sup>。除此之外,最近有许多基于深度学习的方法被用于估计景深,从而用于处理离焦模糊<sup>[70]</sup>。

### 2.1.3 高斯模糊

高斯模糊一般指使用高斯模糊核进行卷积后得到的模糊图像。高斯模糊核的定义如下<sup>[14]</sup>:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (5)$$

其中,  $\sigma$  表示高斯模糊的标准差,也称为模糊半径,模糊半径越长,图像模糊程度越高; $G(x, y)$ 表示模糊核在位置 $(x, y)$ 处的像素值。现有研究普遍采用基于先验约束的策略对高斯模糊进行复原<sup>[67-71]</sup>,鉴于高斯模糊核通常由人工设定且其参数精确已知,仅面向高斯退化的去模糊技术已趋于成熟且表现出良好的鲁棒性,例如以小波域反卷积为代表的传统算法亦可在该场景下取得优异复原效果<sup>[72]</sup>。然而,此类方法的性能高度依赖于对退化过程的解析建模,故主要适用于模糊核已知的场景<sup>[73]</sup>。

### 2.1.4 混合模糊

真实世界中的模糊一般都是多种模糊因素的混合,如相机抖动、物体运动和深度变化等,模糊核较难估计,因此近年来的研究都使用卷积神经网络(CNN)等深度学习模型,学习到图像的特征和模糊核的表示,从而实现图像去模糊<sup>[19,66,70,74]</sup>。图像去模糊是一个复杂的问题,仍然面临许多挑战和难题。例如,处理运动模糊、模糊核估计和处理复杂场景中的模糊等问题仍然是研究的重点。此外,如何平衡去模糊效果和图像细节保留也是一个关键问题。将传统的模糊核估计模型的方法与深度学习模型相结合,以充分利用两种方法的优势。

## 2.2 图像质量评价方法

图像质量评价方法(IQA)一般分为主观评价和客观评价。主观评价常常是基于人类的判断,通常不需要参考图像。一个代表性的主观评价指标是MOS(Mean Opinion Score)<sup>[75]</sup>,它将图像质量分为1级-5级。MOS值取决于总体样本,且通常考虑到意见得分。由于这类方法需要耗费大量人力,效率较低,因此大多数现有方法都是通过客观评价方法来评价的,而根据需要的真实参考图像的情况客观评价方法又可以进一步分为3类:全参考、半参考和无参考。

### 2.2.1 全参考图像质量评价(FR-IQA)

全参考图像质量评估(IQA)的核心方法论涉及将恢复图像与参考图像进行对比,以量化地评估图像质量。该方法依赖于一系列精心构建的特征,用于衡量恢复图像与参考图像之间的差异。在此领域内,众多评估指标被广泛采纳,包括但不限于峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)<sup>[76]</sup>、结构相似度指数(Structural Similarity Index, SSIM)<sup>[77]</sup>、加权信噪比(Weighted Signal-to-Noise Ratio, WSNR)<sup>[78]</sup>、多尺度结构相似度指数(Multi-Scale Structural Similarity Index, MS-SSIM)<sup>[79]</sup>、信息保真度准则(Information Fidelity Criterion, IFC)<sup>[80]</sup>、噪声质量评价(Noise Quality Metric, NQM)<sup>[81]</sup>、无失真质量指数(Universal Image Quality Index, UIQI)<sup>[82]</sup>、

平均绝对误差(Mean Absolute Error, MAE)<sup>[83]</sup>、视觉信息保真度(Visual Information Fidelity, VIF)<sup>[84]</sup>,以及学习感知图像块相似度<sup>[85]</sup>(Learned Perceptual Image Patch Similarity, LPIPS)。在这些指标中,PSNR和SSIM因具有简洁性和直观性而成为最常用的评估工具。特别是LPIPS指标,因其评估结果与人类视觉感知的高一致性,被认为是最为接近人类评判标准的指标。

### 2.2.2 半参考图像质量评价(RR-IQA)

半参考IQA指标的核心理念是利用参考图像的统计特性进行比较,而无需依赖于像素级别的精确匹配信息。这种方法允许在没有完整参考图像的情况下,对图像恢复质量进行评估。目前,FID(Fréchet Inception Distance)<sup>[86]</sup>作为一种流行的半参考指标,通过对比恢复图像集合与参考图像集合在Inception模型特征空间中的统计分布差异来评估图像质量。具体而言,FID计算这两组图像特征的均值和协方差矩阵之间的Fréchet距离,也被称作Wasserstein-2距离<sup>[87]</sup>。FID值降低表明恢复图像与参考图像在特征统计层面更加接近,从而反映出更高的恢复质量。半参考指标在全参考指标和无参考指标之间架起了桥梁,融合了两者的优势。一方面,它能够从参考图像中提取并学习到丰富的信息;另一方面,它对参考图像的要求较为宽松,不强制要求恢复图像与参考图像在像素级别上完全对应。尽管如此,半参考指标在评估过程中可能忽略了恢复图像的真实性,因此它们更多地被应用于图像生成模型的评估中。这种评估方法在图像生成领域尤为重要,因为它能够提供一种在没有精确像素级参考信息的情况下,对生成图像质量进行量化的手段。

### 2.2.3 无参考图像质量评价(NR-IQA)

相比于FR-IQA和RR-IQA指标需要高清真实图像作为参考,无参考IQA指标可以直接对模糊图像进行质量评定,目前的方法有BIQI<sup>[17]</sup>,BLINDS2<sup>[26]</sup>,BRISQUE<sup>[88]</sup>,CORNIA<sup>[89]</sup>,DIIVINE<sup>[31]</sup>,NIQE<sup>[90]</sup>和SSEQ<sup>[91]</sup>等。此外,还有一些指标通过测量对目标检测<sup>[92]</sup>和图像识别<sup>[93]</sup>等视觉任务精度的影响,来评估图像去模糊算法的性能。

## 3 图像去模糊方法

自20世纪60年代非盲去模糊算法提出以来,图像去模糊方法经历了半个多世纪的发展,逐步形成了一个复杂而系统的研究领域。早期的非盲去模糊方法依赖于对模糊核的准确估计,采用经典的傅里叶变换、维纳滤波等技术来恢复图像。然而,这些方法对模糊模型的假设较为严格,因此在实际应用中的效果往往受到限制。2006年,随着盲图像去模糊算法的提出<sup>[94]</sup>,研究重点逐渐转向在未知模糊核条件下的图像复原。盲去模糊方法通过结合图像的先验信息和模糊核的联合估计,在图像处理的多个领域取得了突破性进展。这一时期的研究方法包括基于稀疏表示<sup>[95]</sup>、全变分正则化<sup>[96]</sup>和自然图像统计模型<sup>[36]</sup>的方法,显著提升了复原效果。2015年以后,深度学习的兴起为图像去模糊研究注入了新的活力<sup>[97]</sup>。基于卷积神经网络(CNN)的去模糊方法通过端到端的训练模型,能够直接学习从模糊图像到清晰图像的映射关系,大大简化了传统方法中复杂的模糊核估计过程。此后,各种改进

的深度学习架构被相继提出,如生成对抗网络(GAN)、注意力机制和Transformer模型等,使去模糊算法在处理复杂模糊场景和动态模糊时展现出更高的鲁棒性和性能。图2给出了图像去模糊方法中具有代表性的算法。

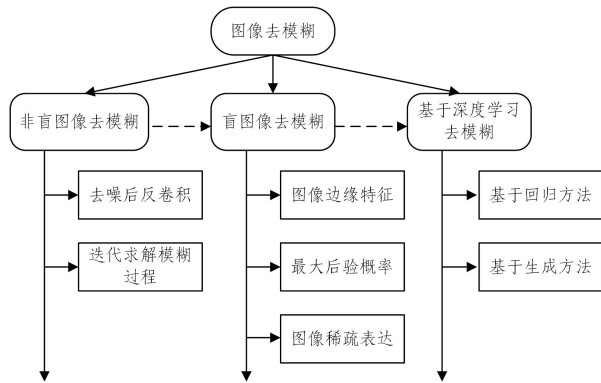


图2 常见去模糊方法

Fig.2 Common deblurring methods

### 3.1 非盲图像去模糊方法

非盲图像去模糊算法的核心在于已知模糊核(点扩散函数,PSF)的前提下,通过逆向计算恢复清晰图像。早期研究中,Wiener提出的维纳滤波成为经典方法之一,该方法利用信号与噪声的二阶统计特性,以最小均方误差为准则进行最优估计,适用于高斯噪声和线性模糊场景。随后,Kalifa等<sup>[98]</sup>结合小波技术,引入镜像小波,通过频域匹配与阈值收缩抑制噪声,进一步提升去模糊性能。Chen等<sup>[99]</sup>引入非局部均值滤波,有效保留了模糊图像中的细腻纹理信息,而Danielyan等<sup>[100]</sup>将三维块匹配滤波(Block Matching 3D Filtering, BM3D)应用于去模糊过程,显著提高了噪声抑制效果。尽管上述方法对去噪和模糊复原具有一定成效,但在缺乏先验信息时易出现振铃效应或噪声残留问题。因此,许多研究集中于迭代优化方法,以提高去模糊效果。Lucy基于贝叶斯理论提出了Richardson-Lucy迭代算法,通过最大似然估计不断优化图像复原结果。Yuan等<sup>[101]</sup>设计了渐进式双边正则化方法,通过多尺度迭代有效恢复图像细节,同时避免振铃效应。Krishnan等<sup>[102]</sup>利用稀疏表达优化图像梯度分布,从而提升去模糊质量。Zoran等<sup>[103]</sup>采用高斯混合模型建模图像块特征,以此为正则项提升复原质量。Kheradmand等<sup>[104]</sup>基于核相似度估计权重,调整拉普拉斯正则化和数据保真项的平衡。Lanza等<sup>[105]</sup>提出变分正则化模型,通过参数化稀疏表达处理高斯噪声干扰。Khetkeeree等<sup>[106]</sup>引入维纳滤波正则项,减轻了对先验信息的依赖,进一步提高了复原效果。此外,一些方法在特定噪声场景下表现出显著优势。例如, Lucy-Richardson方法在处理泊松噪声时表现优异,而全变分去模糊方法在高频细节的保留方面具有独特的优势。然而,这类算法的性能通常高度依赖于模糊核的精确估计,而在实际应用中,噪声往往会被算法放大,从而影响去模糊效果。因此,非盲去模糊方法更适用于模糊核已知且噪声水平较低的场景,它能够在较理想的条件下实现更优的去模糊效果。未来的研究方向可能集中于增强算法的鲁棒性,降低对先验信息的依赖,同时平衡去模糊与去噪性能。

### 3.2 盲图像去模糊方法

盲图像去模糊在模糊核未知的条件下,通过同时估计模糊核和清晰图像来实现图像复原,其研究方法主要包括基于自然图像先验的正则化方法<sup>[36]</sup>、变分贝叶斯模型<sup>[18]</sup>和最大后验概率估计<sup>[66]</sup>等。基于自然图像先验的正则化方法利用图像的稀疏性和边缘特性,通过正则化约束优化模糊核推断。Fergus等<sup>[94]</sup>首次通过最大后验概率方法联合估计模糊核和清晰图像,为盲去模糊提供了开创性的框架。随后,Shan等<sup>[107]</sup>基于局部区域约束优化梯度分布,Xu等<sup>[108]</sup>结合稀疏表达改进去模糊效果,Levin等<sup>[109]</sup>进一步引入变分正则化显著提升了模型性能,Pan等<sup>[27]</sup>基于暗通道稀疏性进行模糊核估算,而Yan等<sup>[110]</sup>结合暗通道与明通道信息进一步提升了去模糊效果,Yang等<sup>[111]</sup>采用变分贝叶斯模型开发了自适应边缘选择算法,有效提高了图像复原质量。稀疏表达在盲去模糊中也发挥了重要作用。例如,Krishnan等<sup>[112]</sup>利用高频稀疏性进行模糊核估计,Xu等<sup>[95]</sup>提出统一框架处理均匀和非均匀模糊问题,Chen等<sup>[113]</sup>基于局部梯度先验改善了细节复原能力,Zhao等<sup>[114]</sup>利用边缘先验进行去模糊。尽管盲去模糊算法已取得重要进展,但在应对复杂模糊核和噪声场景时仍面临诸多挑战。未来研究需要进一步结合多模态先验、深度学习和高效优化算法,以提升算法的鲁棒性和适应性,推动其在计算机视觉领域的深层次应用。

## 4 深度学习去模糊方法

深度学习去模糊方法是近年来图像恢复领域的一个重要研究方向,主要通过深度学习模型自动学习图像中的模糊特征并进行恢复。与传统的去模糊方法不同,深度学习方法不依赖于手工设计的特征和显式的模糊模型,而是通过数据驱动的方式自动学习图像的结构和模糊过程,从而实现去模糊。

### 4.1 基本原理

#### 4.1.1 基于回归的方法

基于回归的方法的核心思想是利用深度学习模型(如Resnet<sup>[115]</sup>,Transformer<sup>[116]</sup>和U-Net<sup>[117]</sup>等),通过大规模训练数据学习模糊图像与清晰图像之间的映射关系。具体来说,给定一张模糊图像 $B$ ,模型通过网络学习到模糊核 $k$ 或者直接从模糊图像中恢复清晰图像 $I_{db}$ 。训练过程中,深度学习网络会通过最小化损失函数来优化网络参数,使恢复的图像尽可能接近真实的清晰图像,常用的有基于像素的损失函数,如均方误差(Mean Square Error, MSE)损失。

$$\mathcal{L}_{\text{pixel}}(I_{db}, I_s) = \|I_{db} - I_s\|_2^2 \quad (6)$$

其中,  $\|\cdot\|_2^2$  是欧几里得范数,用于衡量差异。此外,为了进一步提高去模糊效果,通常会引入感知损失。感知损失是一种用于衡量图像重建或生成任务中感知质量的损失函数。与传统的像素级损失不同,感知损失通过比较高层特征来评估生成图像与目标图像之间的差异,这些高层特征通常来自预训练的卷积神经网络(如VGG<sup>[118]</sup>网络)的某些层。假设预训练网络为 $\phi_l$ ,  $\phi_l(x)$ 表示在第 $l$ 层提取的特征映射, $\lambda_l$ 是平衡每一层损失的权重, $I_s$ 为清晰图像, $I_{db}$ 是去模糊图像,感知损失定义为两者在特征空间中的差异。

$$\mathcal{L}_{\text{perceptual}}(I_{db}, I_s) = \sum_l \lambda_l \|\phi_l(I_{db}) - \phi_l(I_s)\|_2^2 \quad (7)$$

感知损失更关注图像的结构和感知特性,能更好地保留视觉上的相似性。因此,基于回归方法的损失函数一般为:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{pixel}} + \mathcal{L}_{\text{perceptual}} \quad (8)$$

进一步地,为平衡两个损失函数的权重,通常采用基于不确定性的自适应权重分配<sup>[119]</sup>:

$$\mathcal{L}_{\text{total}} = \frac{1}{2\sigma_{\text{pixel}}^2} \mathcal{L}_{\text{pixel}} + \frac{1}{2\sigma_{\text{perceptual}}^2} \mathcal{L}_{\text{perceptual}} + \log \sigma_{\text{pixel}} + \log \sigma_{\text{perceptual}} \quad (9)$$

其中, $\sigma^2$ 表示不确定性方差。

#### 4.1.2 基于生成的方法

基于生成的方法进一步分为生成对抗模型<sup>[49-55]</sup> (Generative Adversarial Network, GAN) 和扩散模型<sup>[48]</sup> (Diffusion Model, DM)。GAN 是由 Goodfellow 等在 2014 年提出的深度学习模型,该模型通过生成器 (Generator) 与判别器 (Discriminator) 之间的对抗训练,能够生成逼真的数据 (如图像、音频等)。在 GAN 中,判别器的作用是评估生成图像是否真实,从而提高图像恢复质量。生成器  $G$  负责生成恢复图像,而判别器  $D$  则判断生成图像与真实图像之间的差异。GAN 的训练目标是优化生成器和判别器之间的对抗损失。

$$\mathcal{L}_{\text{GAN}} = E_{x \sim p_{\text{data}}} [\log D(x)] + E_{z \sim p_z} [\log(1 - D(G(z)))] \quad (10)$$

其中, $p_{\text{data}}$ 是真实图像的分布; $p_z$ 是生成图像的分布; $G(z)$ 是生成器生成的图像; $D(x)$ 是判别器的输出,表示图像  $x$  为真实图像的概率。通过联合训练生成器和判别器,生成器能够恢复出更为真实和自然的清晰图像。

扩散模型去模糊的原理是基于逐步添加噪声的过程,并在反向过程中恢复图像。具体而言,扩散模型包含前向过程和反向过程。在前向过程中,模型从清晰图像  $I_t$  开始,逐步加入噪声,生成模糊图像。这个过程是一个马尔可夫过程<sup>[120]</sup>,每一步的图像状态  $I_t$  都由前一状态  $I_{t-1}$  加上一个小的噪声扰动生成,直到得到完全的随机噪声。其表达式为:

$$q(I_t | I_{t-1}) = \mathcal{N}(I_t; I_{t-1}, \beta_t I_{t-1}) \quad (11)$$

其中, $\mathcal{N}(\mu, \sigma^2)$ 表示一个均值为  $\mu$ 、协方差矩阵为  $\sigma^2$  的高斯分布; $\beta_t$ 是每个时间步的噪声添加强度。这个过程中获得的图像序列将被用于训练一个深度学习网络模型,模型接收当前时间步的噪声图像  $I_t$ ,并预测加入噪声前的图像  $I_{t-1}$ 。反向过程则通过训练好的模型逐步去除噪声,从噪声图像开始,逐步恢复出清晰图像。这个过程通常使用反向条件分布来减少迭代次数,加速图像恢复过程,其计算式为:

$$p_\theta(I_{t-1} | I_t) = \mathcal{N}(I_{t-1}; \mu_\theta(I_t, t), \Sigma_\theta(I_t, t)) \quad (12)$$

基于生成的方法通常能够生成更加细致和丰富的图像纹理,使得生成的图像在视觉上更加真实。然而,这种方法也面临一定的挑战,特别是在去模糊任务中。尽管生成的图像可能具有更多的细节和纹理,但这些细节和纹理有时并非真实的图像内容,而是由模型生成的虚假纹理。这种虚假的纹理可能与真实图像不一致,从而影响到去模糊效果的准确性和质量。特别是在图像恢复过程中,生成器可能会过度拟合训练数据,导致过度修复某些图像细节,这样反而破坏了图像的真实感,最终使得恢复后的图像出现不自然的伪影或失真。因此,如何平衡生成细节和保持真实图像内容之间的关系,成为生成模型在去模糊任务中需要解决的关键问题。

## 4.2 深度学习模型架构

常见的去模糊架构可以分为以下两大类:基于神经网络架构和基于 Transformer 架构。

### 4.2.1 基于神经网络架构

神经网络架构通常由各种类型的网络层通过连接和堆叠的方式构建而成。常见的神经网络层有卷积层、密集层、注意力层和循环层。其计算复杂度如表 1 所列。

表 1 神经网络常用层的计算复杂度

Table 1 Computational complexity of common layers in neural networks

网络层类型	每层计算复杂度	序列操作	最大路径长度
卷积层	$O(k \cdot n^2 \cdot D^2)$	$O(1)$	$O(\log_k(n))$
密集层	$O(n^2 \cdot D^2)$	$O(1)$	$O(1)$
注意力层	$O(n^2 \cdot D)$	$O(1)$	$O(1)$
循环层	$O(n \cdot D^2)$	$O(n)$	$O(n)$

注:其中  $n$  是序列长度,  $D$  表示维度,  $k$  是卷积核大小。

卷积层在从大规模数据中学习具有良好泛化能力的模式方面表现出色,因此它们在去模糊任务中相较于传统方法具有显著优势。卷积操作通过提供局部连接性和位移不变性,为去模糊问题提供了坚实的理论基础。以 ResNet 模型<sup>[115]</sup> 为代表的神经网络架构通过残差连接,显著提升了网络训练速度和去模糊效果。残差连接通过跳跃连接 (Skip Connections) 将输入直接传递给后续层,使网络能够更容易地学习输入图像与目标图像之间的差异,而非直接预测清晰图像。这种设计不仅提升了去模糊性能,还有效缓解了梯度消失问题。在此基础上,U-Net 模型<sup>[117]</sup> 进一步通过精心设计的跳跃连接,将编码器提取的低级特征直接传递给解码器。这种方法可以更好地保留图像的细节信息,特别适用于处理局部模糊场景。Kupyn 等<sup>[38,121]</sup> 则结合生成对抗网络 (GAN) 的思想,提出了单尺度去模糊网络 Deblur-GAN。该方法中,生成器网络采用残差连接和卷积层来实现去模糊,而判别器网络则采用马尔可夫判别器来评估生成图像的质量。尽管卷积层提升了模型的效率和泛化能力,但它也面临两大主要挑战:1) 卷积算子的感受野有限,难以有效建模长距离像素之间的依赖关系;2) 卷积滤波器在推理过程中权重保持静态,缺乏根据输入内容动态调整的灵活性。密集层 (Dense Layers) 提供了有效的解决方案,它有助于改善特征传播、减少参数数量并缓解梯度消失问题。Purohit 等<sup>[122]</sup> 提出了一种由区域自适应模块组成的区域自适应密集网络,用于学习模糊图像中的空间变化偏移。这些区域自适应模块被整合到一个密集层的自编码器架构中。此外,Zhang 等<sup>[123]</sup> 也应用密集层构建了去模糊网络,用密集层代替卷积层,用于更精细的特征提取。在多个尺度上进行图像特征提取的循环层方法也得到广泛研究。SRN<sup>[124]</sup> 通过共享权值深度自编码器,采用由粗到精的结构来消除运动模糊,而 PSS-SRN<sup>[125]</sup> 则通过引入可选的参数共享方案进一步提高了 SRN 的性能。注意力机制的引入帮助深度网络集中于图像的关键区域并进行去模糊处理,Zhao 等<sup>[126]</sup> 提出了一种基于残差注意力和多特征融合的去模糊网络,通过双分支结构和多特征融合模块,有效地捕捉和整合局部不变属性和非局部信息,以提高动态场景下非均匀盲去模糊的效果。Shen 等<sup>[127]</sup> 提出的基于注意力的深度去模糊方法通过 3 个独立的分支,从前景、背景和全局进行

模糊去除,有效提升了去模糊效果。

#### 4.2.2 基于Transformer架构

近年来基于Transformer架构的去模糊方法获得了广泛的关注。类似于神经网络架构,Transformer架构由多个Transformer模块通过堆叠和连接构成。最初,Transformer模型由Vaswani等提出,旨在处理自然语言处理中的序列数据<sup>[116]</sup>。其出色的表达能力和灵活性使其迅速在多个领域得到应用,尤其是在视觉任务中,Vision Transformer(ViT)在图像处理领域得到广泛推广<sup>[137]</sup>。尽管从理论和可解释性角度来看,其他模型如卷积神经网络(CNN)可能在某些任务中具有优势<sup>[138]</sup>,但在实际性能上,尚无其他模型超越Transformer架构。因此,研究者们集中精力深入分析Transformer架构在去模糊等任务中取得优异表现的原因,并探索其潜力。大量实验表明,Transformer的高效性主要得益于其中Transformer块的Token Mixer<sup>[136]</sup>和前馈网络<sup>[138]</sup>(Feed-Forward Network,FFN)的设计。其中,Token Mixer中的自注意力机制被公认为其卓越性能的核心驱动力。然而,随着数据的增加,自注意力机制的复杂度呈二次增长,使其在高分辨率图像处理中的应用变得不可行。为平衡速度与精度,常将图像分割成中等大小的不重叠图像块并编码成序列进行

处理,以此学习这些图像块之间的相互关系。同时,许多方法在不同的图像处理应用中引入多种Token Mixer以降低复杂度。一种常见的策略是仅在局部图像区域内使用注意力机制,例如Swin Transformer<sup>[139]</sup>通过层级窗口注意力(Hierarchical Window Attention)逐步聚合特征,而Neighborhood Attention Transformer<sup>[140]</sup>则聚焦于图像的局部区域,从而减少全局计算。此外,还有方法引入稀疏注意力机制(如Sparse Transformer<sup>[128]</sup>)或结合局部与全局注意力(如Longformer<sup>[129]</sup>,BigBird<sup>[130]</sup>),以进一步优化处理效率。与此同时,低秩分解(如Linformer<sup>[133]</sup>)或核函数近似(如Linear Transformer<sup>[135]</sup>,Performer<sup>[136]</sup>)等技术也得到了广泛应用,以更大程度地降低复杂度并提升模型对高分辨率图像的适应性。此外,研究者们还提出了多种前馈网络设计方案,如Mlp<sup>[141]</sup>,GluMlp<sup>[142]</sup>,GatedMlp<sup>[143]</sup>,ConvMlp<sup>[144]</sup>,SimpleGate<sup>[145]</sup>等。表2列出了不同Transformer架构在Long-Range Arena(LRA)基准测试中的性能比较,数据来源于原始文献<sup>[146]</sup>,旨在评估各架构在处理长序列时的计算性能。这些设计在不同的低层次视觉任务中各具优势。然而,当模型参数量接近时,这些设计的性能差异较小,表明在某些情况下,模型的性能更多依赖于参数优化和网络架构的整体设计,而非单一设计的优劣。

表2 Long-Range Arena基准上的各Transformer变体的长距离依赖性能对比

Table 2 Long-range dependency performance comparison of Long-Range Arena benchmark across different Transformer architectures

Transformer 架构	性能得分	每秒计算次数 (1000)	每秒计算次数 (4000)	峰值内存 (1000)/GB	峰值内存 (4000)/GB
Transformer <sup>[116]</sup>	54.39	8.1	1.4	0.85	9.48
Sparse Transformer <sup>[128]</sup>	51.24	—	—	—	—
Longformer <sup>[129]</sup>	53.46	—	—	—	—
BigBird <sup>[130]</sup>	55.01	7.4	1.5	0.77	2.88
Sinkhorn Transformer <sup>[131]</sup>	51.39	9.1	5.3	0.47	1.48
Reformer <sup>[132]</sup>	50.67	4.4	1.1	0.48	2.28
Linformer <sup>[133]</sup>	51.36	9.3	7.7	0.37	0.99
Synthesizer <sup>[134]</sup>	51.39	8.7	1.9	0.65	6.99
Linear Transformer <sup>[135]</sup>	50.55	9.1	7.8	0.37	1.03
Performer <sup>[136]</sup>	51.41	9.5	8	0.37	1.06

注:“—”表示对应文献无相应数据;每秒计算的步骤数越多,模型性能越优。

## 5 基准数据集与去模糊性能评估

### 5.1 基准数据集

由于模糊图像对应的参考清晰图像较难获取,因此数据

集一般分为真实数据集和合成数据集。真实数据集包含真实世界的模糊图像以及对应的参考清晰图像。合成数据集是对清晰图像进行人工模糊以获取对应的模糊图像。表3列出了图像去模糊数据集的信息。

表3 图像去模糊数据集

Table 3 Image deblurring datasets

数据集	合成模糊	真实模糊	参考图像/幅	模糊图像/幅	应用	训练对/测试对
Levin等 <sup>[33]</sup>	×	√	4	32	运动模糊	—
Sun等 <sup>[147]</sup>	√	×	80	640	运动模糊	—
Köhler等 <sup>[40]</sup>	√	×	4	48	运动模糊	—
Lai等 <sup>[37]</sup>	√	√	108	300	运动模糊	—
GoPro <sup>[148]</sup>	√	√	3214	3214	运动模糊	2103/1111
HIDE <sup>[127]</sup>	√	×	8422	8422	运动模糊	6397/2025
Blur-DVS <sup>[149]</sup>	√	√	2178	2918	运动模糊	1782/396
Middlebury <sup>[150]</sup>	√	√	—	—	高焦模糊	—
Deblur-NeRF <sup>[151]</sup>	√	√	—	—	均有	—
CUHK <sup>[152]</sup>	×	√	—	—	均有	854/146
DPD <sup>[9]</sup>	×	√	500	500	离焦模糊	425/75
Real Blur <sup>[153]</sup>	×	√	4738	4738	运动模糊	3758/980
RealDOF <sup>[154]</sup>	×	√	50	50	离焦模糊	0/50
ReLoBlur <sup>[155]</sup>	×	√	2405	2405	运动模糊	2010/395
RSBlur <sup>[156]</sup>	×	√	13358	13358	运动模糊	9998/3360

注:“—”表示对应文献无相应数据。

### 5.1.1 合成数据集

早期提出的合成数据集一般使用模糊核对清晰图像进行卷积以获取对应的模糊图像。Levin 等<sup>[35]</sup>将相机安装在三脚架上,通过锁定 Z 轴旋转手柄并允许 X 和 Y 方向的运动,来捕捉由真实相机抖动引起的模糊。采集到的图像使用 8 个均匀模糊核进行卷积,构建了一个包含 32 幅模糊图像的数据集。Sun 等<sup>[147]</sup>扩展了该数据集,通过应用 8 个模糊核,共生成了 640 张模糊图像。与 Levin 等的数据集类似,该数据集也仅包含均匀模糊图像,因此不足以用来训练鲁棒的卷积神经网络(CNN)模型。为了模拟非均匀模糊,Köhler 等<sup>[40]</sup>使用 Stewart 平台(即一种机械臂)记录 6D 相机运动,并捕捉打印图片。该数据集包含 4 张潜在的清晰图像和 12 条相机运动轨迹,总共生成了 48 张非均匀模糊图像。Lai 等<sup>[37]</sup>提供了一个数据集,包括 100 张真实模糊图像和 200 张通过均匀模糊核和 6D 相机轨迹生成的合成模糊图像。该数据集的图像覆盖了多种场景,如户外、人物面部、文本和低光照条件,可用于评估在多种环境下的去模糊方法。Nah 等<sup>[148]</sup>使用 GoPro Hero4 Black 相机拍摄并用帧平均法发布了 GoPro 数据集,其包含 3 214 对图像,训练集和测试集分别包含 2 103 对和 1 111 对图像。Shen 等<sup>[127]</sup>主要关注行人和街景,创建了一个包含相机抖动与物体运动的模糊数据集 HIDE。该数据集包括 6 397 对训练图像和 2 025 对测试图像。与 GoPro 数据集类似,HIDE 数据集的模糊图像通过平均连续的 11 帧生成,中心帧作为清晰图像。Scharstein 等<sup>[150]</sup>维护了一个专门用于立体匹配和深度估计研究的数据集 Middlebury,该数据集包含了多个 2001—2021 年间收集的数据集,并且至今仍在持续更新。不同于传统的使用卷积模糊方法和帧平均法,Ma 等<sup>[151]</sup>提出一个基于深度学习三维重建 NeRF(Neural Radiance Fields)技术构建的数据集 Deblur-NeRF,以一个新的三维视角处理去模糊任务。其使用 Canon EOS RP 捕获清晰图像,生成了 20 个真实场景,还捕获模糊图像,生成了 10 个模糊场景,以进行定量分析。

### 5.1.2 真实数据集

人工合成模糊图像在实际场景中的应用存在一定局限性。近年来,随着相机硬件技术的进步,获取真实模糊图像及其对应的清晰图像成为可能。Shi 等<sup>[152]</sup>提出的 CUHK 数据集为模糊检测与去模糊任务提供了重要的基准数据,但由于该数据集缺乏对应的参考清晰图像,其主要用于模糊区域检测与质量评价,而无法直接支持端到端的图像复原算法训练。为解决这一不足,Abuolaim 等<sup>[13]</sup>基于双像素相机构建了一个离焦模糊数据集 DPD(Dual-Pixel Defocus Deblurring)。该数据集包含 500 个精心设计的场景,共 2 000 张图像,包括 500 张离焦模糊图像、1 000 张双像素(DP)子孔径视图,以及 500 张对应的全清晰图像。由于该方式构建的数据集是通过调整曝光时间来拍摄离焦场景,收集训练数据需要一个耗时且繁琐的设置过程,需要调整相机的光圈。因此,Abuolaim 进一步给出了一种合成双像素图像的方法,用于缓解真实数据采集的瓶颈<sup>[157]</sup>。Jiang 等<sup>[149]</sup>通过 DAVIS240C 相机采集了一个大型户外模糊数据集 Blur-DVS,其中包含 1 782 对训练图像、396 对测试图像,并提供 740 张真实模糊图像。Rim

等<sup>[153]</sup>提出了真实世界模糊图像数据集 Real Blur,该数据集分为 Real Blur-R(原始图像)和 Real Blur-J(JPEG 图像)两个子集,共包含 9 476 对图像。Lee 等<sup>[154]</sup>提供了一个针对离焦模糊的测试数据集 RealDOF,该数据集由双像素相机拍摄,包含 50 对测试图像。Li 等<sup>[155]</sup>提出了一个局部运动模糊数据集 ReLoBlur,该数据集以真实场景为基础,展示了局部运动物体与背景之间的混叠现象,包含 2 010 对训练图像和 395 对测试图像。为进一步提升数据集的多样性,Rim 等<sup>[156]</sup>还提出了 RSBlur 数据集,其训练集、验证集和测试集分别包含 8 878 张、1 120 张和 3 360 张模糊图像。

## 5.2 去模糊性能评估

去模糊性能通常以模糊类型进行划分,近年来的研究主要集中于运动去模糊和离焦去模糊任务。在去模糊性能评估中,常用的指标包括峰值信噪比(PSNR)、结构相似度指数(SSIM)、平均绝对误差(MAE)和学习感知图像块相似度(LPIPS)。其中,PSNR 用于衡量图像的细节和清晰度,其值越大,表示恢复图像与原始图像的差异越小,质量越好。类似地,SSIM 更关注图像的结构相似度,值越接近 1,表示恢复图像与原始图像的结构相似度越高,恢复效果越好。相反,MAE 通过计算图像像素级差异来评估质量,值越小,表示恢复图像与原始图像的像素差异越小,质量越好。LPIPS 则是一种基于深度学习的感知质量评估方法,其值越小,表示恢复图像在视觉感知上与原始图像的差异越小,质量越好。去模糊算法的评估通常需要在至少 3 个不同的数据集上进行对比实验,这不仅是为了验证其去模糊能力,还旨在展示算法在多种场景下的泛化性能和适应性。

### 5.2.1 运动去模糊

目前,运动去模糊任务的研究主要基于合成数据集和真实数据集展开,代表性数据集包括 GoPro 数据集、HIDE 数据集以及真实场景数据集 RealBlur 数据集。运动去模糊的性能比较如表 4 所列。其中,传统基于卷积神经网络(CNN)的方法,例如 DeepDeblur<sup>[140]</sup>,通过端到端学习模糊图像与清晰图像之间的映射,取得了显著效果。然而,后续研究进一步优化了特征提取与建模机制,例如 SRN<sup>[124]</sup>通过递归网络逐步恢复图像清晰度,而 DMPHN<sup>[158]</sup>利用多路径递归卷积网络针对不同模糊区域进行独立处理,从而提升了性能。生成对抗网络(GAN)的引入为去模糊任务提供了新的视角,例如 DBGAN<sup>[123]</sup>通过对抗训练生成更具真实感的清晰图像。此外,多任务学习方法在去模糊任务中的应用也表现出较高的潜力,如 MT-RNN<sup>[159]</sup>同时结合了去模糊与高频细节增强任务,有效提升了整体性能。近年来,多阶段网络架构(如 MPRNet<sup>[160]</sup>)以及基于注意力机制的模型(如 HINet<sup>[161]</sup>, NAFNet64<sup>[135]</sup>)展现出了在局部细节恢复与全局依赖建模上的显著优势。值得注意的是,Transformer 架构的引入为图像去模糊研究带来了新的突破。例如,Uformer<sup>[162]</sup>通过结合 U-Net 的分层结构与 Transformer 的全局建模能力实现了性能提升。对频域特征的利用也成为近年来的研究热点,例如 FFTformer<sup>[163]</sup>和 UFPNet<sup>[164]</sup>通过结合频域特征建模与空间域优化,显著提升了图像重建的精度。与此同时,结合多分辨

率分层建模和多阶段优化的模型(如 MRLPFNet<sup>[165]</sup> 和 Restormer<sup>[166]</sup>)也表现出卓越的性能,在去模糊任务中受到广泛

关注。图3展示了常见的运动去模糊算法在 GoPro 数据集上的可视化结果。

表4 运动去模糊性能比较

Table 4 Comparison of motion deblurring performance

方法	GoPro 数据集		HIDE 数据集		RealBlur-R 数据集		RealBlur-J 数据集	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DeepDeblur <sup>[148]</sup>	29.08	0.914	25.73	0.874	32.51	0.841	27.87	0.827
SRN <sup>[124]</sup>	30.26	0.934	28.36	0.915	35.66	0.947	28.56	0.867
DMPHN <sup>[158]</sup>	31.2	0.94	29.09	0.924	35.7	0.948	28.42	0.86
DBGAN <sup>[123]</sup>	31.1	0.942	28.94	0.915	33.78	0.909	24.93	0.745
MT-RNN <sup>[159]</sup>	31.15	0.945	29.15	0.918	35.79	0.951	28.44	0.862
MPRNet <sup>[160]</sup>	32.66	0.959	30.96	0.939	35.99	0.952	28.7	0.873
HINet <sup>[161]</sup>	32.71	0.959	30.32	0.932	—	—	—	—
MIMO-UNet+ <sup>[167]</sup>	32.45	0.957	29.99	0.93	35.54	0.947	27.63	0.837
Whang <sup>[168]</sup>	33.23	0.963	—	—	—	—	—	—
Uformer <sup>[162]</sup>	33.06	0.967	30.9	0.953	36.19	0.956	29.09	0.886
NAFNet64 <sup>[145]</sup>	33.69	0.967	31.32	0.943	35.84	0.952	27.94	0.854
Stripformer <sup>[169]</sup>	33.08	0.962	31.03	0.94	—	—	—	—
Restormer <sup>[166]</sup>	32.92	0.961	31.22	0.942	36.19	0.957	28.96	0.879
DeepRFT+ <sup>[170]</sup>	33.52	0.965	31.66	0.946	36.11	0.955	28.9	0.881
FFTformer <sup>[163]</sup>	34.21	0.968	31.62	0.946	—	—	—	—
UFPNet <sup>[164]</sup>	34.06	0.968	31.74	0.947	36.25	0.953	29.87	0.884
MRLPFNet <sup>[165]</sup>	34.01	0.968	31.63	0.947	—	—	—	—
RevD-S <sup>[171]</sup>	<b>34.35</b>	<b>0.97</b>	<b>32.08</b>	<b>0.95</b>	<b>36.56</b>	<b>0.957</b>	<b>30.09</b>	<b>0.892</b>

注:算法模型均仅在 GoPro 数据集上进行训练,并在多个数据集上分别进行性能评估。其中加粗字体表示各列最优结果,“—”表示对应文献无相应数据。

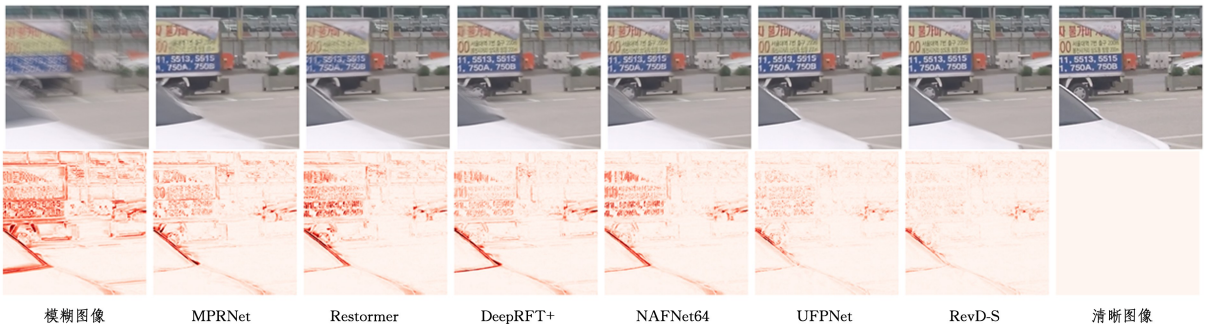


图3 运动去模糊可视化结果

Fig. 3 Visualization results of motion deblurring

### 5.2.2 离焦去模糊

离焦去模糊任务的基准数据集主要有 DPD 数据集、RealDOF 数据集和 Middlebury 数据集。表5列出了近年来离焦去模糊算法在 DPD 数据集上的性能结果。其中, Karaali 等<sup>[172]</sup>提出的 EBDB 方法利用边缘信息来估计散焦模糊,并采用自适应尺度选择策略,提高了模糊估计的精度,并且在图像边缘的细节恢复方面取得了良好的效果。Lee 等<sup>[173]</sup>提出的 DMENet 模型通过结合域适应技术,优化了散焦图像的深度估计,从而提高了图像去模糊的整体效果。此外, Shi 等<sup>[174]</sup>提出的 JNB 方法聚焦于可察觉的散焦模糊检测与估计,在去模糊过程中以感知质量为先,确保复原图像更加精准。Abuolaim 等<sup>[13]</sup>提出的 DPDNet 则利用相机的双像素数据来进行散焦去模糊,这种方法能直接利用相机硬件的双像素信息,有效恢复图像清晰度。Son 等<sup>[175]</sup>提出的 KPAC 方法使用核共享并行空洞卷积来处理单图像的散焦去模糊,能够应对模糊核的多样性,从而提升去模糊的效果。Lee 等<sup>[154]</sup>

设计的 IFAN 方法通过迭代优化滤波器来适应单图像的散焦去模糊任务,从而在处理复杂模糊时取得较好的效果。IRNeXt 方法<sup>[176]</sup>重新审视了卷积网络的设计,通过改进的网络结构提升了去模糊性能。SFNet<sup>[177]</sup>和 FSNet<sup>[178]</sup>方法则通过选择性频率网络和频率选择策略来优化图像恢复,尤其是在去模糊效果的提升方面取得了重要突破。MambaIR<sup>[179]</sup>是一种基于状态空间模型的简单基线方法,旨在通过状态空间模型处理图像的去模糊任务,提供了一种创新的解决思路。Li 等<sup>[180]</sup>提出的 GRL 模型通过结合锚定条纹自注意力、窗口自注意力和通道注意力增强的卷积来有效地显式建模全局、区域和局部范围内的图像层次结构,以实现图像恢复。Gao 等<sup>[181]</sup>提出一种高效的图像去模糊网络 ALGNet,它利用选择性状态空间模型来聚合丰富准确的特征,并通过聚合局部和全局信息块来有效捕捉和整合局部不变属性和非局部信息,从而在保持计算效率的同时消除长距离退化扰动。图4展示了在 DPD 数据集上的离焦去模糊可视化结果。

表 5 离焦去模糊性能比较

Table 5 Comparison of defocus deblurring performance

方法	室内场景				室外场景				综合表现			
	PSNR	SSIM	MAE	LPPIPS	PSNR	SSIM	MAE	LPPIPS	PSNR	SSIM	MAE	LPPIPS
EBDB <sup>[172]</sup>	25.77	0.772	0.040	0.297	21.25	0.599	0.058	0.373	23.45	0.683	0.049	0.336
DMENet <sup>[173]</sup>	25.50	0.788	0.038	0.298	21.43	0.644	0.063	0.397	23.41	0.714	0.051	0.349
JNB <sup>[174]</sup>	26.73	0.828	0.031	0.273	21.10	0.608	0.064	0.355	23.84	0.715	0.048	0.315
DPDNet <sup>[13]</sup>	26.54	0.816	0.031	0.239	22.25	0.682	0.056	0.313	24.34	0.747	0.044	0.277
KPAC <sup>[175]</sup>	27.97	0.852	0.026	0.182	22.62	0.701	0.053	0.269	25.22	0.774	0.040	0.227
IFAN <sup>[154]</sup>	28.11	0.861	0.026	0.179	22.76	0.720	0.052	0.254	25.37	0.789	0.039	0.217
Restormer <sup>[166]</sup>	28.87	0.882	0.025	0.145	23.24	0.743	0.050	0.209	25.98	0.811	0.038	0.178
IRNeXt <sup>[176]</sup>	29.22	0.879	0.024	0.167	23.53	0.752	0.049	0.244	26.30	0.814	0.037	0.206
SFNet <sup>[177]</sup>	29.16	0.878	<b>0.023</b>	0.168	23.45	0.747	0.049	0.244	26.23	0.811	0.037	0.207
FSNet <sup>[178]</sup>	29.14	0.878	0.024	0.166	23.45	0.747	0.050	0.246	26.22	0.811	0.037	0.207
MambaIR <sup>[179]</sup>	28.89	0.879	0.026	0.171	23.36	0.738	0.051	0.243	26.11	0.809	0.039	0.202
GRL <sup>[180]</sup>	29.06	0.886	0.024	<b>0.139</b>	23.45	<b>0.761</b>	0.049	<b>0.196</b>	26.18	<b>0.822</b>	0.037	<b>0.168</b>
ALGNet <sup>[181]</sup>	<b>29.37</b>	<b>0.898</b>	<b>0.023</b>	0.147	<b>23.68</b>	0.755	<b>0.048</b>	0.223	<b>26.45</b>	0.821	<b>0.036</b>	0.186

注:其中加粗字体表示各列最优结果。

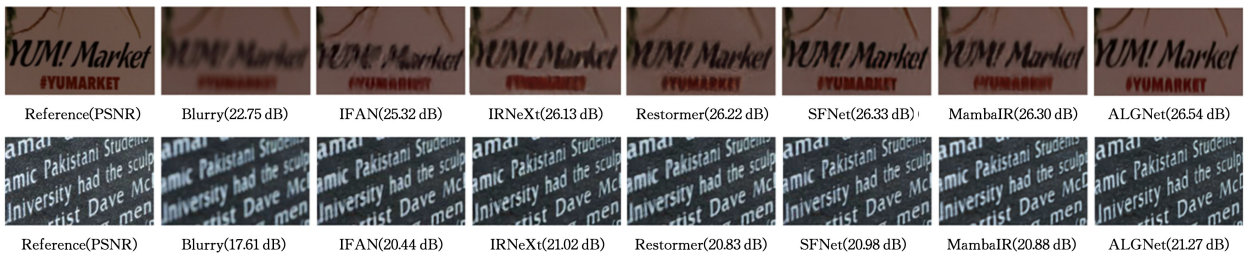


图 4 离焦去模糊可视化结果

Fig. 4 Visualization results of defocus deblurring

## 6 面临的挑战以及未来的发展方向

回顾本文所述的去模糊工作可知,传统的图像去模糊方法在去噪、迭代优化、图像边缘信息和稀疏表达等方面取得了显著进展。这些方法通常依赖于人工选择特征,并针对特定类型的模糊进行优化,因此它们在面对多样化模糊类型和复杂场景时,往往泛化能力有限。尽管如此,这些传统方法为图像处理领域提供了重要的理论基础,为其他相关方法提供了借鉴思路。目前,随着深度学习技术的飞速发展,基于深度学习的图像去模糊方法已经成为该领域的主流。这些方法通过自动化特征学习,能够在海量数据的训练基础上提取更加精准的图像特征,不仅在性能上超越了传统方法,还能更好地适应各种模糊环境,显著提高了去模糊效果的质量和速度。基于卷积神经网络(CNN)、生成对抗网络(GAN)等技术的图像去模糊算法在实际应用中展现了出色的性能,能够有效应对运动模糊、散焦模糊以及其他复杂的图像模糊类型。然而,目前的深度学习方法仍然面临一些局限和挑战。

### 6.1 评价指标的全面性与适用性

现有用于衡量去模糊效果的评价体系仍存在不完善之处,全参考评价指标(Full-Reference Metrics)(如峰值信噪比(PSNR)和结构相似度(SSIM))在实际应用中通常优于无参考评价指标(No-Reference Metrics),但其评价结果更多依赖于像素级差异或全局结构相似性,主要用于衡量去模糊图像与参考清晰图像之间的整体一致性。然而,参考图像的“清晰”并不意味着其所有区域均优于原始模糊图像,某些局部区域甚至可能比原始图像更加模糊,从而影响评价的可靠性。

参考图像本身的质量对于去模糊效果的评估至关重要,但传统全参考指标往往忽视了这一关键因素。相比之下,无参考评价指标无需依赖参考图像,更贴近真实应用场景,能够在仅有待评估图像的情况下对其质量进行估计。然而,现有无参考指标在去模糊任务中往往侧重于图像的自然性、纹理复杂度或统计特征,对运动模糊、离焦模糊等特定退化类型的感知效果的捕捉能力有限。此外,无参考指标的稳定性和与人类主观感知的一致性仍有待提高,这限制了其在去模糊方法评价中的广泛应用。因此,构建能够综合反映局部与全局质量变化并兼顾不同类型模糊感知特性的评价体系,是未来去模糊算法研究的重要方向。

展望未来,随着深度学习技术的持续发展,构建能够融合全参考与无参考评价优势的专用图像质量评估神经网络,将成为该领域的重要研究方向。此类模型应有效避免当前评价体系中常见的矛盾,例如同一幅图像在无参考评价中被判定为质量提升,而在全参考评价中却显示质量下降。新的去模糊效果评价方法应在设计上全面覆盖图像细节恢复能力、自然感知度、边缘保持效果以及局部与全局一致性等多维度特性。

### 6.2 模型泛化能力与数据集多样性

无论是传统去模糊算法还是基于深度学习的方法,现有模型通常依赖大规模图像数据集进行训练。然而,这些数据集在模糊类型与场景覆盖范围上仍存在局限,难以全面反映真实应用中的多样化模糊特征。此外,当前主流的训练方式依然基于传统的梯度下降优化策略,使得模型性能在很大程度上受制于训练数据的质量。当数据集覆盖不足或分布单一

时,模型往往难以具备超出训练集范围的泛化能力,难以在未见过的模糊类型或复杂成像条件下实现稳健表现。

未来应探索更加多样化和高覆盖度的数据采集与构建方式,以提升模型在不同模糊类型与复杂场景下的训练泛化能力。此外,应重点发展更接近人脑思维与学习机制的模型训练范式,以突破传统梯度下降优化策略的局限,从根本上提升模型在未知模糊类型下的鲁棒性。

### 6.3 虚假细节的生成与修复质量

随着去模糊算法的不断优化,生成的去模糊图像的质量不断提升。然而,仍存在一个问题是生成图像可能会出现虚假细节,尤其是在处理极度模糊或低质量图像时。这些虚假细节不仅可能影响图像的视觉效果,还可能在某些应用中造成误导。

未来的研究需要更加关注模型在细节恢复中的可控性和真实性,通过设计更先进的生成模型和正则化方法,减少虚假细节的生成,提高模型对图像真实细节的重建能力。此外,可以通过对生成的图像进行后处理、增强图像质量检查等手段,来进一步提高去模糊图像的视觉真实感。

### 6.4 伦理与隐私问题

图像去模糊技术被广泛应用于各类视觉任务中,包括视频监控、医学影像处理和社交媒体图像增强等。尽管这些技术为人们提供了便利,但也带来了一些伦理和隐私问题。例如,去模糊技术可能被恶意用于修复不真实的图像,进而在社交网络中传播虚假信息;或者在监控视频中恢复原本应被模糊的面部特征,从而侵犯个人隐私。

未来的研究需要制定相应的技术标准和法律法规,确保图像去模糊技术的使用符合社会伦理规范。开发技术透明度高且可追溯的去模糊方法,并在模型设计中加入隐私保护机制,如差分隐私和数据加密等,将有助于避免技术滥用。此外,社会各界应加强对AI技术伦理的讨论,推动技术创新与道德规范的同步发展,确保技术进步惠及社会的各个方面。

### 6.5 基于大模型与多模态学习的发展方向

随着大模型(如GPT<sup>[182]</sup>,BERT<sup>[183]</sup>,DALL·E<sup>[184]</sup>等)的不断发展,图像去模糊领域也受到大规模预训练模型的启发。去模糊方法同样可以通过海量数据的预训练,学习到更加丰富的特征表示,进而在处理不同模糊类型时展现出更强的泛化能力。

展望未来,去模糊算法的发展有望充分受益于大模型与多模态学习技术的进步。未来的去模糊输入将不再局限于单一模态(如单幅图像),而是通过融合多种模态的信息,实现更优的复原效果。例如,将图像的语义信息、上下文数据或多视角观测结果引入去模糊过程,可有效增强模型的预测能力,并提升其对多种模糊类型及复杂成像条件的适应性。在大模型持续优化与计算资源不断提升的背景下,如何设计适用于大规模训练的高效去模糊算法,并在多模态信息融合中保持平衡与有效性,将成为未来的重要研究课题。通过多模态的协同学习,图像去模糊技术有望在复杂真实环境中实现更精确、更具真实感的复原效果,同时显著拓展其应用场景与实用价值。

**结束语** 本综述围绕“图像去模糊技术的研究进展”这一核心主题,从传统方法与深度学习方法两个方向对相关研究

进行了多维度、多层次的系统回顾与总结分析。首先,从经典去模糊算法、深度学习模型、数据集构建及评价指标等方面,对现有图像去模糊技术进行了全面梳理与归纳,系统性地总结了不同方法在理论依据、实际效果及适用场景上的优势与局限性。其次,从评价指标的全面性、模型的泛化性、虚假细节生成、伦理隐私问题及大规模多模态模型的潜力等方面,深入探讨了当前研究所面临的主要挑战与发展瓶颈。最后,基于对现有局限的分析与对新兴技术的展望,提出了未来研究的发展方向,包括实现基于多模态学习的去模糊能力、构建更具普适性的评价体系,以及在保障伦理与隐私的前提下推动真实世界场景中的应用落地。本文旨在通过对图像去模糊技术的回顾与展望,为相关领域的研究人员提供系统化的参考与指导,为进一步提升图像去模糊技术的理论与应用水平奠定基础,同时为推进其在实际场景中的广泛应用提供新思路与方向。

## 参考文献

- [1] ZUO H, FU C, LI S, et al. Adversarial Blur-Deblur Network for Robust UAV Tracking[J]. IEEE Robotics and Automation Letters, 2023, 8(2): 1101-1108.
- [2] SHORTEN C, KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning[J/OL]. <https://doi.org/10.1186/s40537-019-0197-0>.
- [3] SINGH D, SAHU M R. A survey on various image deblurring techniques[J]. International Journal of Advanced Research in Computer and Communication Engineering, 2013, 116: 15-18.
- [4] CHEN X, HE X, YANG J, et al. An effective document image deblurring algorithm[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2011.
- [5] USHA RANI B, ARUNA R, VELRAJKUMAR P, et al. Image Denoising Using Deblur Generative Adversarial Network Denoising U-Net[J/OL]. Journal of Circuits, Systems and Computers, 2024: 2550069. <https://doi.org/10.1142/5s0218126625500690>.
- [6] REN S, HE K, GIRSHICK R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[C]// Advances in Neural Information Processing Systems, 2015.
- [7] HE K, GKIOXARI G, DOLLÁR P, et al. Mask r-cnn[C]// IEEE International Conference on Computer Vision, 2017.
- [8] WANG J, XIANG L Y, HUANG Y B, et al. Prior-guided blind restoration algorithm for iris image[J]. Computer Science, 2023, 50(12): 148-155.
- [9] HUANG Y Q, ZHENG P Y, JIANG M R, et al. Reconstruction of solar spot diagrams using gated fusion network and residual Fourier transform[J]. Computer Science, 2023, 50(S2): 278-284.
- [10] WANG J M, LI X F, YE L, et al. Medical image deblurring based on channel attention structure generative adversarial network[J]. Computer Science, 2021, 48(S1): 101-106.
- [11] PAN J S. Research progress on image deblurring methods based on deep learning[J]. Computer Science, 2021, 48(3): 9-13.
- [12] CHO S, LEE S. Fast motion deblurring[C]// ACM SIGGRAPH Asia, 2009.

- [13] ABUOLAIM A, BROWN M S. Defocus deblurring using dual-pixel data [C] // European Conference on Computer Vision. 2020.
- [14] HUMMEL R A, KIMIA B, ZUCKER S W. Deblurring gaussian blur[J]. Computer Vision, Graphics, and Image Processing, 1987, 38(1):66-80.
- [15] BIGDELI S A, ZWICKER M, FAVARO P, et al. Deep mean-shift priors for image restoration[C] // Advances in Neural Information Processing Systems. 2017:763-772.
- [16] MATHIEU M, COUPRIE C, LECUN Y. Deep multi-scale video prediction beyond mean square error[J]. arXiv:1511.05440, 2015.
- [17] MOORTHY A K, BOVIK A C. A two-step framework for constructing blind image quality indices[J]. IEEE Signal Processing Letters, 2010, 17(5):513-516.
- [18] BABACAN S D, MOLINA R, DO M N, et al. Bayesian blind deconvolution with general sparse image priors[C] // European Conference on Computer Vision. 2012.
- [19] CHO S, WANG J, LEE S. Handling outliers in non-blind image deconvolution[C] // IEEE International Conference on Computer Vision. 2011.
- [20] REN W, ZHANG J, MA L, et al. Deep non-blind deconvolution via generalized low-rank approximation[C] // Proceedings of the 32<sup>nd</sup> International Conference on Neural Information Processing Systems. 2018.
- [21] SCHMIDT U, ROTHER C, NOWOZIN S, et al. Discriminative non-blind deblurring[C] // IEEE Conference on Computer Vision and Pattern Recognition. 2013.
- [22] SCHULER C J, CHRISTOPHER BURGER H, HARMELING S, et al. A machine learning approach for non-blind image deconvolution[C] // IEEE Conference on Computer Vision and Pattern Recognition. 2013:1067-1074.
- [23] VASU S, REDDY MALIGIREDDY V, RAJAGOPALAN A. Non-blind Deblurring: Handling Kernel Uncertainty with CNNs [C] // IEEE Conference on Computer Vision and Pattern Recognition. 2018.
- [24] ZHANG J, PAN J, LAI W S, et al. Learning fully convolutional networks for iterative non-blind deconvolution[C] // IEEE Conference on Computer Vision and Pattern Recognition. 2017.
- [25] WIESCHOLLEK P, HIRSCH M, SCHOLKOPF B, et al. Learning blind motion deblurring[C] // IEEE International Conference on Computer Vision. 2017.
- [26] SAAD M A, BOVIK A C, CHARRIER C. Blind image quality assessment: A natural scene statistics approach in the DCT domain[J]. IEEE Transactions on Image Processing, 2012, 21(8):3339-3352.
- [27] PAN J, SUN D, PFISTER H, et al. Blind image deblurring using dark channel prior[C] // IEEE Conference on Computer Vision and Pattern Recognition. 2016.
- [28] PAN J, DONG J, TAI Y W, et al. Learning discriminative data fitting functions for blind image deblurring[C] // IEEE International Conference on Computer Vision. 2017.
- [29] NIMISHA T M, KUMAR SINGH A, RAJAGOPALAN A N. Blur-invariant deep learning for blind-deblurring[C] // IEEE International Conference on Computer Vision. 2017.
- [30] NAN Y, QUAN Y, JI H. Variational-EM-based Deep Learning for Noise-blind Image Deblurring [C] // IEEE Conference on Computer Vision and Pattern Recognition. 2020.
- [31] MOORTHY A K, BOVIK A C. Blind image quality assessment: From natural scene statistics to perceptual quality [J]. IEEE Transactions on Image Processing, 2011, 20(12):3350-3364.
- [32] MICHAELI T, IRANI M. Blind deblurring using internal patch recurrence [C] // European Conference on Computer Vision. 2014.
- [33] LI Y, TOFIGHI M, GENG J, et al. Deep algorithm unrolling for blind image deblurring[J]. arXiv:1902.03493, 2019.
- [34] LI L, PAN J, LAI W S, et al. Learning a discriminative prior for blind image deblurring[C] // IEEE Conference on Computer Vision and Pattern Recognition. 2018.
- [35] LEVIN A, WEISS Y, DURAND F, et al. Understanding and evaluating blind deconvolution algorithms[C] // IEEE Conference on Computer Vision and Pattern Recognition. 2009.
- [36] LEVIN A. Blind motion deblurring using image statistics[C] // Advances in Neural Information Processing Systems. 2007:841-848.
- [37] LAI W S, HUANG J B, HU Z, et al. A comparative study for single image blind deblurring[C] // IEEE Conference on Computer Vision and Pattern Recognition. 2016.
- [38] KUPYN O, BUDZAN V, MYKHAILYCH M, et al. Deblurgan: Blind motion deblurring using conditional adversarial networks [C] // IEEE Conference on Computer Vision and Pattern Recognition. 2018.
- [39] KRISHNAN D, TAY T, FERGUS R. Blind deconvolution using a normalized sparsity measure[C] // IEEE Conference on Computer Vision and Pattern Recognition. 2011.
- [40] KÖHLER R, HIRSCH M, MOHLER B, et al. Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database[C] // European Conference on Computer Vision. 2012.
- [41] FIORI S, UNCINI A, PIAZZA F. Blind deconvolution by modified Bussgang algorithm[C] // The IEEE International Symposium on Circuits and Systems. 1999:1-4.
- [42] CHAN T F, WONG C K. Total variation blind deconvolution [J]. IEEE Transactions on Image Processing, 1998, 7(3):370-375.
- [43] CHAKRABARTI A. A neural approach to blind motion deblurring[C] // European Conference on Computer Vision. 2016.
- [44] AL-AMEEN Z, SULONG G. Deblurring computed tomography medical images using a novel amended landweber algorithm[J]. Interdisciplinary Sciences: Computational Life Sciences, 2015, 7(3):319-325.
- [45] ESLAMI S A, HEES N, WEBER T, et al. Attend, infer, repeat: Fast scene understanding with generative models[C] // Advances in Neural Information Processing Systems. 2016.
- [46] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural computation, 1989, 1(4):541-551.
- [47] MARTIN D, FOWLKES C, TAL D, et al. A database of human

- segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C]// IEEE International Conference on Computer Vision, 2001.
- [48] ZHAO W, ZHENG B, LIN Q, et al. Enhancing diversity of defocus blur detectors via cross-ensemble network[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2019.
- [49] DU H, LI Z, NIYATO D, et al. Enabling AI-Generated Content Services in Wireless Edge Networks[J]. IEEE Wireless Communications, 2024, 31(3): 226-234.
- [50] SHAO L, CHEN B, ZHANG Z, et al. Artificial intelligence generated content (AIGC) in medicine: A narrative review [J]. Mathematical Biosciences and Engineering, 2024, 21(1): 1672-1711.
- [51] LEI Y, LI J, LI Z, et al. Prompt learning in computer vision: a survey[J]. Frontiers of Information Technology & Electronic Engineering, 2024, 25(1): 42-63.
- [52] WU T, CHEN Z, HE D, et al. CDDM: Channel Denoising Diffusion Models for Wireless Semantic Communications[J]. IEEE Transactions on Wireless Communications, 2024, 23(9): 11168-11183.
- [53] DENTON E L, CHINTALA S, FERGUS R, et al. Deep generative image models using a Laplacian pyramid of adversarial networks[C]// Advances in Neural Information Processing Systems, 2015.
- [54] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]// Advances in Neural Information Processing Systems, 2014.
- [55] HO J, ERMON S. Generative adversarial imitation learning [C]// Advances in Neural Information Processing Systems, 2016.
- [56] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]// IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [57] PAN J, LIU Y, DONG J, et al. Physics-based generative adversarial models for image restoration and beyond[J]. arXiv:1808.00605, 2018.
- [58] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv:1511.06434, 2015.
- [59] REED S, AKATA Z, YAN X, et al. Generative adversarial text to image synthesis[J]. arXiv:1605.05396, 2016.
- [60] BORACCHI G, FOI A. Modeling the performance of image restoration from motion blur[J]. IEEE Transactions on Image Processing, 2012, 21(8): 3502-3517.
- [61] SUN J, CAO W, XU Z, et al. Learning a convolutional neural network for non-uniform motion blur removal[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2015.
- [62] XU L, JIA J. Two-phase kernel estimation for robust motion deblurring[C]// European Conference on Computer Vision, 2010.
- [63] XU X, PAN J, ZHANG Y J, et al. Motion blur kernel estimation via deep learning[J]. IEEE Transactions on Image Processing, 2017, 27(1): 194-205.
- [64] ZHAO H X, CHAO M L, NI F C. Overdrive LUT optimization for LCD by box motion blur measurement and gamma-based thresholding method[C]// SID Symposium Digest of Technical Papers, 2008: 117-120.
- [65] GONG D, ZHANG Z, SHI Q, et al. Learning Deep Gradient Descent Optimization for Image Deconvolution[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(12): 5468-5482.
- [66] GOLDSTEIN A, FATTAL R. Blur-kernel estimation from spectral irregularities[C]// European Conference on Computer Vision, 2012.
- [67] SON C H, PARK H M. A pair of noisy/blurry patches-based PSF estimation and channel-dependent deblurring [J]. IEEE Transactions on Image Processing, 2011, 57(4): 1791-1799.
- [68] SUN L, CHO S, WANG J, et al. Edge-based blur kernel estimation using patch priors[C]// IEEE International Conference on Computational Photography, 2013.
- [69] ALJADAANY R, PAL D K, SAVVIDES M. Douglas-rachford networks: Learning both the image prior and data fidelity terms for blind image deconvolution[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2019.
- [70] ZHANG K, ZUO W, GU S, et al. Learning deep CNN denoiser prior for image restoration[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [71] CHEN F, MA J. An empirical identification method of Gaussian blur parameter for image deblurring[J]. IEEE Transactions on Signal Processing, 2009, 57(7): 2467-2478.
- [72] GETREUER P. A survey of Gaussian convolution algorithms [J]. Image Processing on Line, 2013, 2013: 286-310.
- [73] VAIRY M, VENKATESH Y V. Deblurring Gaussian blur using a wavelet array transform [J]. Pattern Recognition, 1995, 28(7): 965-976.
- [74] SUN D, YANG X, LIU M Y, et al. PWC-Net: Cnns for optical flow using pyramid, warping, and cost volume[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [75] HOSSFELD T, HEEGAARD P E, VARELA M, et al. QoE beyond the MOS: an in-depth look at QoE via better metrics and their relation to MOS[J]. Quality and User Experience, 2016, 1(1): 2.
- [76] HORE A, ZIOU D. Image quality metrics: PSNR vs. SSIM[C]// IEEE International Conference on Pattern Recognition, 2010.
- [77] 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.
- [78] MITSU T, VARKUR K L. Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms[C]// IEEE International Conference on Acoustics, Speech, and Signal Processing, 1993.
- [79] WANG Z, SIMONCELLI E P, BOVIK A C. Multiscale structural similarity for image quality assessment[C]// The Asilomar Conference on Signals, Systems, and Computers, 2003.
- [80] SHEIKH H R, BOVIK A C, DE VECIANA G. An information fidelity criterion for image quality assessment using natural scene statistics[J]. IEEE Transactions on Image Processing, 2005, 14(12): 2117-2128.

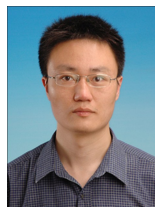
- [81] DAMERA-VENKATA N, KITE T D, GEISLER W S, et al. Image quality assessment based on a degradation model[J]. *IEEE Transactions on Image Processing*, 2000, 9(4): 636-650.
- [82] WANG Z, BOVIK A C. A universal image quality index[J]. *IEEE signal processing letters*, 2002, 9(3): 81-84.
- [83] WUN L M, PEARN W L. Assessing the statistical characteristics of the mean absolute error or forecasting[J]. *International Journal of Forecasting*, 1991, 7(3): 335-337.
- [84] SHEIKH H R, BOVIK A C. Image information and visual quality[J]. *IEEE Transactions on Image Processing*, 2006, 15(2): 430-444.
- [85] ZHANG R, ISOLA P, EFROS A A, et al. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. 2018.
- [86] SHARVIT Y. The puzzle of free indirect discourse[J]. *Linguistics and Philosophy*, 2008, 31(3): 353-395.
- [87] BEIER F, BEINERT R, STEIDL G. On a Linear Gromov-Wasserstein Distance[J]. *IEEE Transactions on Image Processing*, 2022, 31: 7292-7305.
- [88] MITTAL A, MOORTHY A K, BOVIK A C. No-reference image quality assessment in the spatial domain[J]. *IEEE Transactions on Image Processing*, 2012, 21(12): 4695-4708.
- [89] YE P, KUMAR J, KANG L, et al. Unsupervised feature learning framework for no-reference image quality assessment[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. 2012.
- [90] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a "completely blind" image quality analyzer[J]. *IEEE Signal Processing Letters*, 2012, 20(3): 209-212.
- [91] LIU L, LIU B, HUANG H, et al. No-reference image quality assessment based on spatial and spectral entropies[J]. *Signal Processing: Image Communication*, 2014, 29(8): 856-863.
- [92] TANG C, WANG P, ZHANG C, et al. Salient object detection via weighted low rank matrix recovery[J]. *IEEE Signal Processing Letters*, 2016, 24(4): 490-494.
- [93] JIANG P, LING H, YU J, et al. Salient region detection by ufo: Uniqueness, focusness and objectness[C]// *IEEE International Conference on Computer Vision*. 2013.
- [94] FERGUS R, SINGH B, HERTZMANN A, et al. Removing camera shake from a single photograph[C]// *ACM SIGGRAPH*. 2006.
- [95] XU L, ZHENG S, JIA J. Unnatural l0 sparse representation for natural image deblurring[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. 2013.
- [96] PAN J, HU Z, SU Z, et al. Deblurring text images via L0-regularized intensity and gradient prior[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
- [97] HRADIŠ M, KOTERA J, ZEMCIK P, et al. Convolutional neural networks for direct text deblurring[C]// *British Machine Vision Conference*. 2015.
- [98] KALIFA J, MALLAT S, ROUGE B. Deconvolution by thresholding in mirror wavelet bases[J]. *IEEE Transactions on Image Processing*, 2003, 12(4): 446-457.
- [99] CHEN F, HUANG X, CHEN W. Texture-Preserving Image Deblurring[J]. *IEEE Signal Processing Letters*, 2010, 17(12): 1018-1021.
- [100] DANIELYAN A, KATKOVNIK V, EGAZARIAN K. BM3D frames and variational image deblurring[J]. *IEEE Transactions on Image Processing*, 2011, 21(4): 1715-1728.
- [101] YUAN L, SUN J, QUAN L, et al. Progressive inter-scale and intra-scale non-blind image deconvolution[J]. *ACM Transactions on Graphics*, 2008, 27(3): 1-10.
- [102] KRISHNAN D, FERGUS R. Dark flash photography[J]. *ACM Transactions on Graphics*, 2009, 28(3): 1-11.
- [103] ZORAN D, WEISS Y. From learning models of natural image patches to whole image restoration[C]// *IEEE International Conference on Computer Vision*. 2011.
- [104] KHERADMAND A, MILANFAR P. A general framework for regularized, similarity-based image restoration[J]. *IEEE Transactions on Image Processing*, 2014, 23(12): 5136-5151.
- [105] LANZA A, MORIGI S, SGALLARI F. Convex Image Denoising via Non-convex Regularization with Parameter Selection[J]. *Journal of Mathematical Imaging and Vision*, 2016, 56(2): 195-220.
- [106] KHETKEEREE S, LIANGROCAPART S. Image Restoration Using Optimized Weiner Filtering Based on Modified Tikhonov Regularization[C]// *2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP)*. Wuxi, China; IEEE, 2019: 1015-1020.
- [107] SHAN Q, XIONG W, JIA J. Rotational motion deblurring of a rigid object from a single image[C]// *IEEE International Conference on Computer Vision*. 2007.
- [108] XU Y, LI Q Y. Blind deblurring of binary images based on sparse characteristics[J]. *Computer Science*, 2018, 45(3): 255-259.
- [109] LEVIN A, WEISS Y, DURAND F, et al. Efficient marginal likelihood optimization in blind deconvolution[C]// *CVPR 2011*. 2011: 2657-2664.
- [110] YAN Y, REN W, GUO Y, et al. Image Deblurring via Extreme Channels Prior[C]// *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI; IEEE, 2017: 6978-6986.
- [111] YANG L, JI H. A Variational EM Framework With Adaptive Edge Selection for Blind Motion Deblurring[C]// *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA; IEEE, 2019: 10159-10168.
- [112] KRISHNAN D, FERGUS R. Fast image deconvolution using hyper-Laplacian priors[C]// *Advances in Neural Information Processing Systems*. 2009.
- [113] CHEN L, FANG F, WANG T, et al. Blind Image Deblurring With Local Maximum Gradient Prior[C]// *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA; IEEE, 2019: 1742-1750.
- [114] ZHAO Z G, CHEN Y Y, ZHAO Y, et al. Motion deblurring based on edge prior model[J]. *Computer Science*, 2015, 42(5): 305-308.
- [115] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. 2016.
- [116] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is

- all you need[J]. arXiv:1706.03762,2017.
- [117]RONNEBERGER O,FISCHER P,BROX T. U-net:Convolutional networks for biomedical image segmentation[C]//The International Conference on Medical Image Computing and Computer-assisted Intervention. 2015.
- [118]SIMONYAN K,ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv:1409.1556,2014.
- [119]CIPOLLA R,GAL Y,KENDALL A. Multi-task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA; IEEE, 2018: 7482-7491.
- [120]GIRARDIN V. Entropy Maximization for Markov and Semi-Markov Processes[J]. Methodology and Computing in Applied Probability,2004,6(1):109-127.
- [121]KUPYN O,MARTYNIUK T,WU J,et al. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better[C]//IEEE International Conference on Computer Vision. 2019.
- [122]PUROHIT K,RAJAGOPALAN A. Region-Adaptive Dense Network for Efficient Motion Deblurring [J]. arXiv: 1903.11394,2019.
- [123]ZHANG K,LUO W,ZHONG Y,et al. Deblurring by realistic blurring[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2020.
- [124]TAO X,GAO H,SHEN X,et al. Scale-recurrent network for deep image deblurring[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2018.
- [125]GAO H,TAO X,SHEN X,et al. Dynamic scene deblurring with parameter selective sharing and nested skip connections[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2019.
- [126]ZHAO Q,ZHOU D M,YANG H,et al. Image deblurring via residual attention and multi-feature fusion[J]. Computer Science, 2023,50(1):147-155.
- [127]SHEN Z,WANG W,LU X,et al. Human-Aware Motion Deblurring[C]//IEEE International Conference on Computer Vision. 2019.
- [128]CHILD R,GRAY S,RADFORD A,et al. Generating Long Sequences with Sparse Transformers [J]. arXiv: 1904. 10509, 2019.
- [129]BELTAGY I,PETERS M E,COHAN A. Longformer: The Long-Document Transformer[J]. arXiv:2004.05150,2020.
- [130]ZAHEER M,GURUGANESH G,DUBEY K A,et al. Big Bird: Transformers for Longer Sequences[C]//NeurIPS. 2020;17283-17297.
- [131]TAY Y,BAHRI D,YANG L,et al. Sparse Sinkhorn Attention [C]//ICML. 2020;9438-9447.
- [132]KITAIEV N,KAISER L,LEVSKAYA A. Reformer: The Efficient Transformer[C]//ICLR. 2020.
- [133]WANG S,LI B Z,KHABSA M,et al. Linformer: Self-Attention with Linear Complexity[J]. arXiv:2004.04768,2020.
- [134]TAY Y,BAHRI D,METZLER D,et al. Synthesizer: Rethinking Self-Attention in Transformer Models[J]/[OL]. <https://proceedings.mlr.press/v139/tay21a/tay21a.pdf>.
- [135]KATHAROPOULOS A,VYAS A,PAPPAS N,et al. Transformers are RNNs: Fast Autoregressive Transformers with Linear Attention[C]//ICML. 2020;5156-5165.
- [136]CHOROMANSKI K,LIKHOSHERSTOV V,DOHAN D,et al. Rethinking Attention with Performers[J]. arXiv:2009.14794,2020.
- [137]DOSOVITSKIY A,BEYER L,KOLESNIKOV A,et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale[J]. arXiv:2010.11929,2020.
- [138]LECUN Y,BOTTOU L,BENGIO Y,et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE,1998,86(11):2278-2324.
- [139]LIU Z,LIN Y,CAO Y,et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows[J]. arXiv: 2103.14030,2021.
- [140]HASSANI A,WALTON S,LI J,et al. Neighborhood Attention Transformer[J]. arXiv:2204.07143,2022.
- [141]BURGER H C,SCHULER C J,HARMEILING S. Image denoising: Can plain neural networks compete with BM3D? [C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence,RI:IEEE,2012:2392-2399.
- [142]DAUPHIN Y N,FAN A,AULI M,et al. Language Modeling with Gated Convolutional Networks [J]. arXiv: 1612. 08083, 2016.
- [143]LIU H,DAI Z,SO D R,et al. Pay Attention to MLPs[J]. arXiv: 2105.08050,2021.
- [144]LI J,HASSANI A,WALTON S,et al. ConvMLP: Hierarchical Convolutional MLPs for Vision[J]. arXiv:2109.04454,2021.
- [145]CHEN L,CHU X,ZHANG X,et al. Simple Baselines for Image Restoration[J]. arXiv:2204.04676,2022.
- [146]TAY Y,DEGHANI M,ABNAR S,et al. Long Range Arena: A Benchmark for Efficient Transformers[C]//ICLR. 2021.
- [147]SUN L,HAYS J. Super-resolution from internet-scale scene matching[C]//IEEE International Conference on Computational Photography. 2012.
- [148]NAH S,HYUN KIM T,MU LEE K. Deep multi-scale convolutional neural network for dynamic scene deblurring[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2017.
- [149]JIANG Z,ZHANG Y,ZOU D,et al. Learning Event-Based Motion Deblurring[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA; IEEE,2020:3317-3326.
- [150]SCHARSTEIN D,BRIGGS A J. Real-time recognition of self-similar landmarks [J]. Image and Vision Computing, 2001, 19(11):763-772.
- [151]MA L,LI X,LIAO J,et al. Deblur-NeRF: Neural Radiance Fields from Blurry Images[J]. arXiv:2111.14292,2021.
- [152]SHI J P,XU L,JIA J Y. Discriminative Blur Detection Features [C]//IEEE Conference on Computer Vision and Pattern Recognition. 2014.
- [153]RIM J,LEE H,WON J,et al. Real-world blur dataset for learning and benchmarking deblurring algorithms [C] // European Conference on Computer Vision. 2020.
- [154]LEE J,SON H,RIM J,et al. Iterative filter adaptive network for

- single image defocus deblurring[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021:2034-2042.
- [155]LI H,ZHANG Z,JIANG T, et al. Real-World Deep Local Motion Deblurring[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(1):1314-1322.
- [156]RIM J,KIM G,KIM J, et al. Realistic Blur Synthesis for Learning Image Deblurring[J]. arXiv:2202.08771, 2022.
- [157]ABUOLAIM A,DELBRACIO M,KELLY D, et al. Learning to Reduce Defocus Blur by Realistically Modeling Dual-Pixel Data [C]//ICCV. 2021.
- [158]ZHANG H,DAI Y,LI H, et al. Deep Stacked Hierarchical Multi-patch Network for Image Deblurring[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2019.
- [159]PARK P 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.
- [160]ZAMIR S W,ARORA A,KHAN S H, et al. Multi-Stage Progressive Image Restoration[C]//CVPR. 2021.
- [161]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.
- [162]WANG Z,CUN X,BAO J, et al. Uformer: A General U-Shaped Transformer for Image Restoration[C]//CVPR. 2022.
- [163]KONG L,DONG J,GE J, et al. Efficient Frequency Domain-based Transformers for High-Quality Image Deblurring [C] // CVPR. 2023.
- [164]FANG Z,WU F,DONG W, et al. Self-supervised Non-uniform Kernel Estimation with Flow-based Motion Prior for Blind Image Deblurring[C]//CVPR. 2023.
- [165]DONG J,PAN J,YANG Z, et al. Multi-Scale Residual Low-Pass Filter Network for Image Deblurring[C]//ICCV. 2023.
- [166]ZAMIR S W,ARORA A,KHAN S, et al. Restormer: Efficient Transformer for High-Resolution Image Restoration [C] // CVPR. 2022.
- [167]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.
- [168]WHANG J,DELBRACIO M,TALEBI H, et al. Deblurring via Stochastic Refinement[C]//CVPR. 2022.
- [169] TSAI F J,PENG Y T,LIN Y Y, et al. Stripformer: Strip Transformer for Fast Image Deblurring[C]//ECCV. 2022.
- [170]MAO X,LIU Y,LIU F, et al. Intriguing Findings of Frequency Selection for Image Deblurring[C]//AAAI. 2023.
- [171]MAO X,LI Q,WANG Y. AdaRevD: Adaptive Patch Exiting Reversible Decoder Pushes the Limit of Image Deblurring[J]. arXiv:2406.09135, 2024.
- [172]KARAALI A,JUNG C R. Edge-based defocus blur estimation with adaptive scale selection[J]. IEEE Transactions on Image Processing, 2018, 27(3):1126-1137.
- [173]LEE J,LEE S,CHO S, et al. Deep Defocus Map Estimation Using Domain Adaptation[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2019.
- [174]SHI J,XU L,JIA J. Just noticeable defocus blur detection and estimation[C]//CVPR. 2015.
- [175]SON H,LEE J,CHO S, et al. Single Image Defocus Deblurring Using Kernel-Sharing Parallel Atrous Convolutions[C]//ICCV. 2021.
- [176]CUI Y,REN W,CAO X, et al. Revitalizing Convolutional Network for Image Restoration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(12):9423-9438.
- [177]CUI Y,TAO Y,BING Z, et al. Selective Frequency Network for Image Restoration[C]//The Eleventh International Conference on Learning Representations. 2023.
- [178]CUI Y,REN W,CAO X, et al. Image Restoration via Frequency Selection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(2):1093-1108.
- [179]GUO H,LI J,DAI T, et al. MambaIR: A Simple Baseline for Image Restoration with State-Space Model [J]. arXiv: 2402.15648, 2024.
- [180]LI Y W,FAN Y C, XIANG X Y, et al. Efficient and Explicit Modelling of Image Hierarchies for Image Restoration [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2023.
- [181]GAO H,DANG D. Learning Enriched Features via Selective State Spaces Model for Efficient Image Deblurring[J]. arXiv: 2403.20106, 2024.
- [182]OPENAI, ACHIAM J, ADLER S, et al. GPT-4 Technical Report[J]. arXiv:2303.08774, 2023.
- [183]DEVLIN J,CHANG M W,LEE K, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [J]. arXiv:1810.04805, 2018.
- [184]RAMESH A,PAVLOV M,GOH G, et al. Zero-Shot Text-to-Image Generation[J]. arXiv:2102.12092, 2021.



**CHEN Kang**, born in 1997, postgraduate. His main research interests include image deblurring, image/video analysis and related visual problems.



**LIU Yuanjie**, born in 1985, Ph.D, associate professor, is a senior member of CCF (No. I0004S). His main research interests include pattern recognition and image processing.