

# 基于近端梯度下降算法展开的盲图像去模糊方法

杨 涛<sup>1</sup> 朱文球<sup>1</sup>  
YANG Tao ZHU Wenqiu

## 摘 要

目前在盲图像去模糊任务中,深度学习技术已展现出卓越的性能。然而,当前多数方法主要聚焦于启发式网络架构的构建,而对模糊核与清晰图像之间物理生成机制的显式嵌入关注不足。这在一定程度上限制了方法对不同模糊核的通用性和模型的可解释性。为克服这一局限性,提出一种创新的模型驱动深度神经网络。通过迭代算法精确求解盲图像去模糊的优化模型,并将迭代步骤巧妙地嵌入相应的网络模块中,从而实现网络学习过程与盲图像去模糊物理机制的深度融合。实验结果表明,所提出的方法在准确性和通用性方面均显著优于现有的代表性盲图像去模糊技术,能够有效应对各种模糊类型,为盲图像去模糊研究提供了新的思路和方法。

## 关键词

图像去模糊;近端梯度下降;算法展开;深度学习

doi: 10.3969/j.issn.1672-9528.2024.10.023

## 0 引言

图像去模糊的主要目标是从捕获的模糊图像中重构出高质量的清晰图像。在实际应用中,非盲图像去模糊方法需要依赖已知的模糊核,然而这在多数场景下难以实现。为解决此难题,研究焦点逐渐转向模糊核未知的盲图像去模糊领域。

目前,盲图像去模糊方法主要可分为两类:基于优化模型的方法和基于深度学习的方法<sup>[1]</sup>。基于优化模型的盲图像去模糊工作旨在将模糊核和清晰图像的先验知识制定为一种优化算法,以约束图像去模糊问题的求解空间。由于涉及迭代计算,基于优化模型的方法通常很耗时。此外,手动设计的先验通常基于自然图像的统计特征,而这些特征无法全面反映真实场景中图像的复杂性。

随着深度卷积神经网络(CNNs)在成像任务中的广泛应用,有研究提出使用端到端网络直接从单一图像中去模糊,无需预先估计模糊核,简化了传统的两步去模糊流程。然而,使用单一网络对任意模糊核进行去卷积仍面临巨大挑战<sup>[2]</sup>。由于未充分考虑图像去模糊任务的物理生成机制,深度卷积神经网络在处理特定模糊核时的表现仍不及优化模型方法。

算法展开策略提出将迭代算法展开为网络结构训练,利用数据驱动神经网络求解优化模型。文献[3]将基于迭代的L0正则化盲图像去模糊方法展开到深度网络中,兼顾了优化模型和深度卷积神经网络的优点,在适应任意模糊核的同时

大幅提高了求解效率。然而,现有的算法展开盲图像去模糊方法多使用深度神经网络隐式估计模糊核,而未显示求解数据项。

为克服上述问题,本文提出一种新的算法展开盲图像去模糊网络。在模糊图像退化机制引导下,同时更新模糊核和清晰图像。总而言之,本文的主要贡献包括以下几点。

(1)提出了一种新的模型驱动的算法展开盲去模糊网络,该网络采用迭代算法求解经典退化模型,并将迭代步骤映射到相应的网络模块中,以联合估计模糊核和清晰图像。迭代过程严格遵循模糊图像的物理生成机制来学习模糊核和清晰图像,因此每个网络模块都具有明确的物理可解释性。

(2)通过显式求解模糊核数据项,学习得到的模糊核呈现出具有特定物理意义的显式结构。由于嵌入了模糊图像物理生成机制,所估计的模糊核能够适应不同类型的模糊。

(3)在包含多种模糊类型的合成数据集上进行了大量实验,结果全面证明了所提方法在盲图像去模糊性能上优于当前主流方法,从而验证了所提方法的合理性和有效性。

## 1 理论与方法

### 1.1 优化模型

图像去模糊是典型的病态逆问题,图像模糊过程可以描述为:

$$y = A * x + n \quad (1)$$

式中:  $y$  表示模糊图像,  $x$  表示清晰图像,  $A$  表示模糊核,  $n$  表示噪声,  $*$  表示卷积操作。

1. 湖南工业大学计算机学院 湖南株洲 412007

[基金项目] 湖南省教育厅青年项目(22B0586); 湖南省教育厅教改项目(2022JGYB186)

根据贝叶斯公式推理, 可将盲图像去模糊问题建模为优化模型:

$$\begin{aligned} \min_{k,x} & \|y - k * x\|_2^2 + \lambda J(x) + \beta \Phi(k) \\ \text{s.t. } & k_i \geq 0, \sum_i k_i = 1 \end{aligned} \quad (2)$$

式中:  $\Phi(k)$  和  $J(x)$  分别表示提供模糊核  $k$  和清晰图像  $x$  先验知识的正则项;  $\lambda$  和  $\beta$  表示正则项的权重系数。数据保真度项 (即方程 (2) 的目标函数中的第一项) 表示模糊图像物理生成机制, 用于指导模糊核  $k$  和清晰图像  $x$  的迭代更新, 而先验项  $\Phi(k)$  和  $J(x)$  用于约束解空间, 提供高质量解。本文摒弃了传统基于优化模型的图像去模糊方法中常用的手工制作先验函数, 转而采用数据驱动的策略。借助深度神经网络 (DNN), 以端到端的方式直接从数据中提取  $x$  和  $k$  背后的隐式先验知识。这一创新方法已在众多视觉任务中被充分验证<sup>[4]</sup>, 其有效性得到了广泛认可。

## 1.2 近端梯度下降算法

为了构建一种与优化算法中迭代步骤紧密对应的网络模块, 以增强算法的可解释性和通用性, 本文通过近端梯度算法交替迭代更新模糊核  $k$  和清晰图像  $x$ , 以实现问题的优化求解。

近端梯度下降算法是一种优化算法, 主要用于解决具有约束条件或非光滑目标函数的优化问题。它的核心思想是在每次迭代中, 首先计算目标函数在当前点的梯度, 然后根据梯度信息和一个近端算子来更新参数。近端操作通常是一个非线性的映射, 用于处理正则化项带来的非光滑性。

简单来说, 近端梯度下降算法就是在梯度下降的基础上增加一个近端项来修正参数的更新方向, 以确保算法能够收敛到满足约束条件的最优解。值得一提的是, 该优化算法仅涉及基础的近端梯度运算, 这一特性使得其能够轻松嵌入到网络模块中, 实现算法展开为网络。以下是该优化算法的求解步骤。

通过公式 (3) 建立优化模型更新模糊核  $k$ :

$$\begin{aligned} \min_k & \|k^{(t-1)} - (k^{(t-1)} - \delta_1 \nabla f(k^{(t-1)}))\|^2 + \beta \Phi(k) \\ \text{s.t. } & k_i \geq 0, \sum_i k_i = 1 \end{aligned} \quad (3)$$

式中:  $\delta_1$  表示步长,  $t$  表示迭代次数,  $f(k^{(t-1)})$  的具体形式为  $\|y - k^{(t-1)} * x^{(t-1)}\|^2$ 。使用近端算子  $\text{prox}(\bullet)$ , 可将  $k$  子问题通过公式 (4) 求解:

$$k^{(t)} = \text{prox}(k^{(t-1)} - (k^{(t-1)} - \delta_1 \nabla f(k^{(t-1)}))) \quad (4)$$

通过公式 (5) 建立优化模型更新模糊核  $x$ :

$$\min_x \|x - (x^{(t-1)} - \delta_2 \nabla h(x^{(t-1)}))\|^2 + \lambda J(x) \quad (5)$$

式中:  $\delta_2$  表示步长,  $t$  表示迭代次数,  $h(k^{(t-1)})$  具体形式为  $\|y - k^{(t-1)} * x^{(t-1)}\|^2$ 。  $k$  子问题通过公式 (6) 求解:

$$x^{(t)} = \text{prox}(x^{(t-1)} - (x^{(t-1)} - \delta_2 \nabla f(x^{(t-1)}))) \quad (6)$$

近端算子  $\text{prox}(\bullet)$  一般采用软阈值收缩<sup>[5]</sup>, 计算成本和伪影可能会随着迭代次数的增加而增加。本文采用卷积神经网络代替近端算子, 通过数据驱动学习更加丰富的先验信息和近端映射方式具体细节见 1.3 节。

## 1.3 网络结构

近端梯度下降算法本质仍然是传统基于模型的去模糊方法, 需要不断迭代计算直到收敛。为减少计算复杂度和时间成本, 在本节中, 利用算法展开策略将带核偏差图像去模糊方法展为深度神经网络。通过展开方程 (4) 和方程 (6) 中每个迭代步骤作为相应的网络模块, 构建一个端到端的盲图像去模糊网络。网络包含 3 个模块: 梯度下降模块、 $x$  子问题近端网络 (Prox-xNet)、 $k$  子问题近端网络 (Prox-kNet)。

梯度下降模块分别将  $k$  子问题与  $x$  子问题梯度下降过程展开, 为简化模型训练, 使用可训练的线性层替代超参数  $\delta_1$  和  $\delta_2$ , 通过数据驱动对参数进行学习。这一设计避免了训练前繁琐的超参数调整, 同时大幅提升了计算时间。

由于近端映射与恒等映射的近似性, 本研究采用配备残差连接的 U-Net 网络架构, 作为解决  $x$  子问题的近端算子。该模块由特征提取与特征重建两大部分组成, 整体结构形似字母 U。在特征提取阶段包含三个卷积块, 每一卷积块均包含两个配备  $3*3$  卷积核的卷积层, 并辅以 ReLU 激活函数, 以增强网络对图像低级特征的学习能力。每个卷积块后均设有一个最大池化层, 旨在缩减特征图的尺寸, 同时确保关键信息的留存, 从而帮助网络更好地捕捉多尺度的特征。进入特征重建阶段, 同样采用三个反卷积块, 每个反卷积块由两个配备  $3*3$  卷积核的卷积层及 ReLU 激活函数构成, 旨在协助网络深入学习图像的高级语义信息。每个反卷积块后都设置有上采样层, 旨在逐步增加特征图的尺寸, 与特征提取阶段的池化层形成对应关系, 从而确保特征图能够恢复至与原始图像相近的尺寸。此外, 特征重建中的各层均通过跳跃连接与特征提取阶段的对应层相连, 此举有助于实现底层与高层语义信息的有效融合。最终, 使用长残差连接将其与输入相加, 从而有效规避了梯度消失的问题。

鉴于模糊核  $k$  的大小远小于输入图像  $x$  的大小, 采用体积较小的残差块 (ResBlock) 作为  $k$  子问题的近端算子。该模块由 2 个卷积块和残差连接组成, 卷积块包含两个配备  $3*3$  卷积核的卷积层和 ReLU 激活函数, 最终与输入相加保留整体特征信息。整体网络模型根据公式 (4) 与公式 (6)

输入输出关系进行连接，为缩小网络参数数量，迭代次数设置为  $T=5$ ，同时不同迭代次数的近端网络设置为共享参数，线性层初始化  $\delta_1=0.1$ ， $\delta_2=0.1$ ，卷积层初始化方式设置为均值为 0、方差为 1 的 Kaiming 初始化。整体网络输入为模糊图像，输出为清晰图像。总体网络结构如图 1 所示。

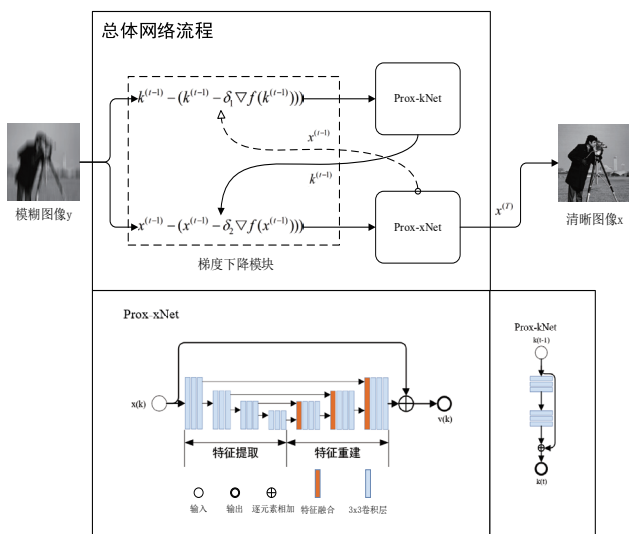


图 1 网络结构图

## 2 实验与结果分析

### 2.1 实验环境

网络模型训练环境 CPU 为 R5 4600H，内存为 16 GB，显卡为显存容量 4 GB 的 GTX1650 显卡；软件环境为 Windows 10 操作系统，深度学习环境为 PyTorch。训练数据集使用 BSD300 训练数据集，将图像随机裁剪为 128 像素  $\times$  128 像素的图像，最终得到 10 000 个清晰图像，经过与模糊核卷积后加上高斯噪声得到模糊图像。测试数据集使用 Set10。网络训练采用 Adam 优化器，参数设置为  $\beta_1=0.9$ ， $\beta_2=0.999$ ， $\varepsilon=10^{-8}$ ，网络初始学习率设置为  $4 \times 10^{-4}$ ，每 40 轮下降 0.5 倍，总计进行 200 轮训练，每轮训练对数据进行批归一化处理，批处理大小设置为 4。

### 2.2 图像去模糊结果

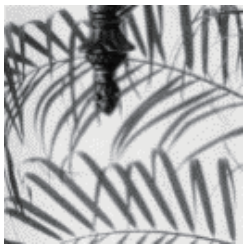
为了更直观地展示所提出的网络在图像去模糊方面的效果，本文与几种主流的去模糊方法进行了详细比较。这些方法包括三种基于模型的去模糊方法：EPLL<sup>[6]</sup>（图像块对数似然算法）、IDDBM3D<sup>[7]</sup>（基于迭代解耦去模糊的 BM3D 方法）以及 NCSR<sup>[8]</sup>（非局部稀疏表示）。此外，还对比了两种基于深度学习的方法：MLP<sup>[9]</sup> 和 IRCNN<sup>[10]</sup>。在表 1 中，本文详细列出了这些方法在处理 4 幅图像（如图 2 所示），以及面对三种不同类型的模糊核时，所得到的峰值信噪比数据。这一全面的比较旨在提供一个清晰、客观的性能评估。

表 1 图像去模糊结果

| 方法                             | 蝴蝶    | 树叶    | 鹦鹉    | 海星    |
|--------------------------------|-------|-------|-------|-------|
| 标准差为 1.6 的高斯模糊，噪声等级为 2         |       |       |       |       |
| EPLL                           | 24.60 | 23.55 | 29.62 | 27.41 |
| IDDBM3D                        | 28.50 | 28.69 | 30.90 | 29.52 |
| NCSR                           | 28.31 | 27.51 | 30.39 | 29.39 |
| MLP                            | 28.84 | 28.86 | 31.21 | 29.71 |
| IRCNN                          | 29.90 | 29.74 | 32.07 | 30.69 |
| 本文方法                           | 30.12 | 30.34 | 31.70 | 31.01 |
| 19 $\times$ 19 运动模糊，噪声等级为 7.65 |       |       |       |       |
| EPLL                           | 21.57 | 20.59 | 26.33 | 24.36 |
| IRCNN                          | 27.12 | 27.07 | 29.57 | 26.98 |
| 本文方法                           | 27.35 | 27.35 | 29.92 | 27.12 |
| 17 $\times$ 17 运动模糊，噪声等级为 2.55 |       |       |       |       |
| EPLL                           | 25.05 | 23.57 | 30.71 | 28.10 |
| IRCNN                          | 30.88 | 30.37 | 33.15 | 30.10 |
| 本文方法                           | 30.65 | 31.05 | 33.16 | 30.97 |



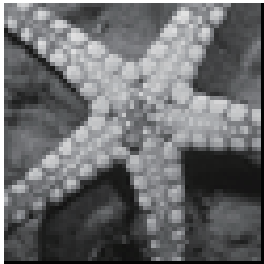
(a) 蝴蝶



(b) 树叶



(c) 鹦鹉



(d) 海星

图 2 清晰图像示例

根据表 1 的数据分析，可以清晰地观察到，在高斯模糊的情形下，本文所提出的方法相较于其他对比方法展现出明显的优势。对于两个运动模糊核的处理，IRCNN 和本文的方法均表现出比传统基于模型的方法更出色的性能。特别值得一提的是，由于本文方法中融入了残差连接，对噪声具有进一步的抑制作用，因此在高噪声水平的环境下，本文方法的性能略胜一筹，甚至略优于 IRCNN。此外，IRCNN 需要通过多次迭代才能达到理想的去模糊效果，而本文的方法则仅需执行  $T=5$  个阶段，效率显著提高。



图3直观地展示了清晰图像、模糊图像,以及经过EPLL、IDDBM3D、NCSR、MLP、IRCNN和本文方法处理后的高斯模糊去模糊效果图。从图3中可以清晰地看到,本文的方法在保留图像细节方面表现出色,例如蝴蝶翅膀上的锯齿纹理得到了更好的还原。与此同时,EPLL和IDDBM3D等方法由于利用了非局部特征,在处理后的图像边界部分仍然存在一定的模糊。

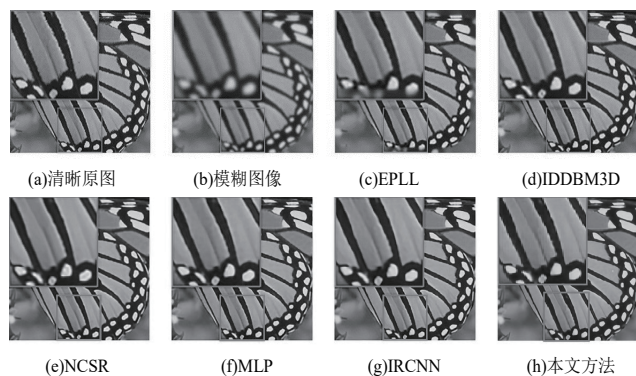


图3 高斯去模糊对比图

图4运动模糊的去模糊视觉效果,依次展示了4幅图像清晰原图、模糊图像和去模糊后的图像,本文方法能在去模糊后仍然保留良好人眼视觉效果。

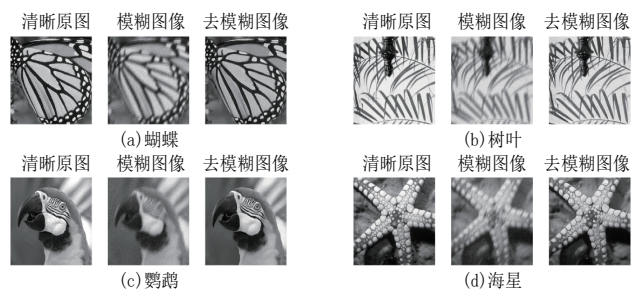


图4 运动模糊去模糊视觉效果图

综上所述,与基于模型的传统方法相比,基于深度学习的方法在去模糊方面展现出显著的优势。其中,本文所提出的方法更能适应高水平的噪声环境与高斯模糊,不仅无需进行迭代计算,而且能够更好地保留图像的精细纹理,为图像处理领域提供了一种新的高效解决方案。

### 3 结论

本文提出了一种创新的盲图像去模糊网络,该网络通过深度展开近端梯度算法,将迭代步骤融入网络设计中。此举不仅在网络学习过程中嵌入了模糊核与清晰图像之间的物理关系,显著提升了网络的可解释性,还通过显式求解模糊核数据项,使网络能够灵活适应不同类型的模糊,从而大大增强了其通用性。实验结果表明,所提网络有优秀的图像去模糊性能。

### 参考文献:

- [1] 胡张颖,周全,陈明举,等.图像去模糊研究综述[J].中国图象图形学报,2024,29(4):841-861.
- [2] 王旭晴,魏伟波,杨光宇,等.基于算法展开的图像盲去模糊深度学习网络[J].山东大学学报(工学版),2023,53(6):35-46.
- [3] 侯瑞峰,张鹏程,张丽媛,等.基于全变分正则项展开的迭代去噪网络[J].计算机应用,2024,44(3):916-921.
- [4] 张弛,贾丽媛,冯福双.基于广义迭代收缩的图像复原新算法[J].控制工程,2016,23(6):844-847.
- [5] ZORAN D, WEISS Y. From learning models of natural image patches to whole image restoration[C]//2011 International Conference on Computer Vision. Piscataway: IEEE, 2011: 479-486.
- [6] DANIELYAN A, KATKOVNIK V, EGIAZARIAN K. BM3D frames and variational image deblurring[J]. IEEE transactions on image processing, 2011, 21(4): 1715-1728.
- [7] WEI S D, ZHANG L, SHI G, et al. Nonlocally centralized sparse representation for image restoration[J]. IEEE transactions on image processing, 2012, 22(4): 1620-1630.
- [8] BURGER H C, SCHULER C J, HARMELING S. Image denoising: can plain neural networks compete with BM3D?[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2012: 2392-2399.
- [9] WANG M Z, ZUO W, GU S, et al. Learning deep CNN denoiser prior for image restoration[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 3929-3938.
- [10] LEVIN A, WEISS Y, DURAND F, et al. Understanding and evaluating blind deconvolution algorithms[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2009: 1964-1971.

### 【作者简介】

杨涛(2001—),男,湖南衡阳人,硕士研究生,研究方向:图像处理。

朱文球(1969—),男,湖南攸县人,硕士,教授,硕士生导师,研究方向:人工智能。

(收稿日期:2024-07-17)