

稀疏性和自相似性先验引导的深度学习图像盲超分

葛孙逸^{1,2}, 罗小伟³, 冯世阳^{1,2}, 王 斌^{1,2*}

(1. 复旦大学 电磁波信息科学教育部重点实验室, 上海 200433;

2. 复旦大学 信息学院图像与智能实验室, 上海 200433;

3. 紫光展锐(上海)科技有限公司 音视频技术资源部, 上海 200120)

摘要: 现有的基于深度学习的图像盲超分算法仅利用神经网络端到端地学习低分辨率图像到高分辨率图像的映射, 让网络隐式地学习图像的先验, 导致算法仍产生模糊的超分结果。针对上述问题, 提出一种稀疏性和自相似性先验引导的深度学习图像盲超分算法。首先, 针对不同的低分辨率图像输入, 利用动态线性核估计模块, 有效估计出相应模糊核; 然后, 利用基于快速迭代软阈值收缩算法(FISTA)的深度展开反卷积滤波模块, 显式地对信号的稀疏性先验进行建模, 实现对退化图像的反卷积恢复; 最后, 利用双通道多尺度大感受野恢复模块, 借助图像自相似性先验进行超分恢复。实验结果表明, 相较于现有方法, 所提出算法在公开的 Gaussian8 数据集上达到了 31.66 的峰值信噪比(PSNR)与 0.8725 的结构相似度(SSIM), 在公开的 DIV2K 数据集上实现了 29.08 的 PSNR 与 0.8007 的 SSIM, 其所恢复出的图像不仅具有最高的复原指标, 还具有更佳的视觉效果。

关 键 词: 图像盲超分; 深度学习; 稀疏性先验; 自相似性先验; 深度展开网络

中图分类号: TP751

文献标识码: A

Sparsity and self-similarity priors guided deep learning for blind image super-resolution

GE Sun-Yi^{1,2}, LUO Xiao-Wei³, FENG Shi-Yang^{1,2}, WANG Bin^{1,2*}

(1. Key Laboratory for Information Science of Electromagnetic Waves (MoE), Fudan University, Shanghai 200433, China;

2. Image and Intelligence Laboratory, School of Information Science and Technology, Fudan University, Shanghai 200433, China;

3. Media Technology Resources Department, UNISOC (Shanghai) Technologies Co. Ltd, Shanghai 200120, China)

Abstract: The existing deep learning-based image blind super-resolution algorithms only utilize neural networks to learn the end-to-end mapping from low-resolution (LR) images to high-resolution (HR) images, only allowing the network to implicitly learn image priors, resulting in algorithms that still produce blurry super-resolution results. To address the above issues, a deep learning image blind super-resolution algorithm guided by sparsity and self-similarity priors is proposed. Initially, for various LR image inputs, a dynamic linear kernel estimation module is employed to effectively estimate the corresponding blur kernels; Subsequently, a deep unfolding deconvolution filtering module based on the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) is utilized to explicitly model the sparsity prior of signal, achieving deconvolution restoration of the degraded images; Finally, a dual-path multi-scale large receptive field restoration module leverages the self-similarity prior of images for super-resolution recovery. The experimental results indicate that, compared to existing methods, the proposed algorithm achieves a peak signal-to-noise ratio (PSNR) of 31.66 and a structural similarity index (SSIM) of 0.8725 on the publicly available Gaussian8 dataset, and attains a PSNR of 29.08 and an SSIM of 0.8007 on the DIV2K dataset. The images restored by the proposed algorithm not only exhibit the highest restoration metrics but also superior visual quality.

Key words: blind image super-resolution, deep learning, sparsity prior, self-similarity prior, deep unfolding network

收稿日期: 2024-10-22, 修回日期: 2024-12-03

Received date: 2024-10-22, Revised date: 2024-12-03

基金项目: 国家自然科学基金(62371140)

Foundation item: Supported by National Natural Science Foundation of China (62371140)

作者简介(Biography): 葛孙逸(2000—), 男, 江苏人, 硕士研究生, 主要研究领域为图像盲超分. E-mail: 22210720129@m.fudan.edu.cn

*通讯作者(Corresponding author): E-mail: wangbin@fudan.edu.cn

引言

图像超分旨在从所观测到的低分辨率图像中恢复出相应的高分辨率图像,一直是计算机视觉中的基础任务,在手机图像恢复、老旧相片复原、视频监控和卫星遥感领域中有广泛应用^[1]。近年来,随着深度学习技术的发展,各种先进的图像超分算法(如EDSR^[2],DBPN^[3],CARN^[4]和RCAN^[5]等)相继被提出,大大提升了图像超分性能。现有的图像超分算法,通常假设高分辨率图像按照一种预先定义好的退化过程产生低分辨率图像,即假设图像仅遭受降采样(通常为双三次下采样)退化;但是,现实中的低分辨率图像往往经受了复杂退化过程。由于存在这种潜在域偏差^[6]问题,这类超分算法在真实场景应用中会导致严重的性能退化。为了克服这一重要缺陷,图像盲超分应运而生。

图像盲超分问题^[6]旨在从未知退化的低分辨率图像中恢复出相应的高分辨率图像,其退化过程可建模为如下经典降质模型:

$$y = (x * k_h)_{\downarrow s} + n, \quad (1)$$

其中, y 是降质后的低分辨率图像, x 是原始的高分辨率图像, $*$ 表示二维卷积操作, k_h 表示作用在高分辨率图像 x 上的卷积核, $\downarrow s$ 表示以 s 倍率进行图像下采样, n 则为加性高斯白噪声。图像盲超分是指:在未知 k_h 和 n 的前提下,由观测到的低分辨率图像 y 恢复出原始的高分辨率图像 x 。显然,它考虑了更复杂全面的退化情形,具有更强的泛化能力和实用性。

现有的图像盲超分方法主要分为基于模型和基于深度学习两大类。基于模型的方法^[7-8]通常精心选择合适的先验,将盲超分问题转化为数据保真项加先验项形式的优化问题,进而求解该优化问题以恢复出最终的超分图像。这类方法由于先验知识的引入而具有泛化能力强、物理可解释性好的优点,然而这类方法也存在如下缺点:(1)算法精度较低,在某些图像数据上,由于预设先验知识与数据实际特性存在不匹配或不完全一致情况,算法性能显著下降;(2)推理时间过长,优化问题往往需要几十次甚至上百次迭代运算才能达到较为满意结果,导致其难以满足实际应用的快速求解需求。伴随着人工智能技术的快速发展,基于深度学习的方法能自动学习图像中隐藏的先验信息;同时,由于其直接学习低分辨率图像到高分辨率图像的映射,因而具有较短的推理时间,可克服基于模型方法的不

足,成为图像盲超分的主流。现阶段,基于深度学习的图像盲超分方法可大致分为两类:双阶段方法和端到端方法。

双阶段方法,将图像盲超分问题分解为两个子问题:(1)模糊核估计问题;(2)已知模糊核的图像超分问题,也称为非盲超分问题。近年来,这两个子问题得到了较深入的研究。针对模糊核估计问题,KernelGAN^[9]使用GAN^[10]网络充分挖掘图像内部信息以实现模糊核的有效估计;KernelNet^[11]使用低分辨率图像的梯度图和自卷积策略实现了更加高效的模糊核估计;FKP^[12]则使用归一化流来建模模糊核的分布以实现模糊核的有效估计,然而巨大的计算开销限制了该类方法的应用场景。针对非盲超分问题,ZSSR^[13]针对每张图片单独训练一个网络以进行超分,训练的时间开销较大;SRMD^[14]等则将模糊核展开到和低分辨率图像相同尺寸大小,送入网络进行超分,取得了较好的超分效果;USRNet^[15]将非盲超分问题建模为优化问题,并通过神经网络迭代求解,取得了最好的非盲超分效果。通过组合两类子问题的算法构成双阶段方法,能较好解决图像盲超分问题,具有模块化和可解释性好的优点。然而,双阶段方法却存在算法不适配缺点:模糊核估计算法估计出的模糊核,往往与真实模糊核存在一定的差异,而非盲超分算法只在已知真实模糊核的情况下才能产生较好的超分效果,其模糊核估计算法的微小误差将导致非盲超分算法性能的急剧下降。

为了解决双阶段方法的不适配缺点,端到端的深度学习图像盲超分算法被提出。该类算法将模糊核估计算法和非盲超分算法嵌入到一个网络中,并进行端到端的优化,克服了双阶段方法的不足,大幅提升了图像盲超分性能,因而成为目前图像盲超分的主流。IKC^[16]、AdaTarget^[17]、KOALAnet^[18]和DAN^[19-21]系列算法估计出模糊核的相对表征,然后将该相对表征和低分辨图像送入到后续的非盲超分网络中进行图像超分,取得了较好的图像超分效果,具有精度高的优点,但是,它们仍存在如下缺陷:(1)使用模糊核的相对表征,未能有效利用其结构信息;(2)让网络直接学习低分辨率图像到高分辨率图像的映射,隐式地学习图像先验,导致神经网络往往会优先拟合低频信息^[22]而产生模糊的超分图像;(3)采用小感受野特征,恢复时仅考虑小邻域信息,忽视了图像的自相似性先验。目前最好的

算法 DCLS^[23]重新改写了经典降质模型,在非盲超分网络中先使用最小二乘滤波,利用模糊核的结构信息对模糊的低分辨率图像进行恢复,接着再进行后续的超分工作,由于对模糊核信息的充分利用而取得了目前最好的效果。虽然,它成功解决了缺陷 1,但缺陷 2 和 3 的存在,仍制约了 DCLS 算法性能的进一步提高,成为目前进一步提升算法性能亟需解决的核心问题。

针对上述问题,本文提出一种基于稀疏性和自相似性先验引导的深度学习(Deep Learning guided by Sparsity and Self-Similarity Priors, DLS³P)的图像盲超分算法,该算法主要由三部分构成:(1)动态线性核估计(Dynamic Linear Kernel Estimation, DLKE)模块,该模块采用动态神经网络对模糊核进行估计,针对不同的低分辨率图像产生其相应参数,以实现准确的核估计结果;(2)基于快速迭代软阈值收缩算法^[24](Fast Iterative Shrinkage-Thresholding Algorithm, FISTA)的深度展开反卷积滤波(Deep Unfolding Deconvolution Filtering, DUDF)模块,该模块通过显式地对稀疏性先验进行建模,将反卷积的优化问题转化为广义最小绝对收缩和选择算子(Least Absolute Shrinkage and Selection Operator, LASSO)问题^[25],然后采用 FISTA 算法进行求解,并将其展开为深度网络进行端到端优化,高效完成图像反卷积操作,使得恢复出的图像具有清晰锐利的边缘;(3)双通道多尺度大感受野恢复(Dual-path Multi-scale Large-receptive-field Restoration, DMLR)模块,该模块有效利用图像的自相似性先验,使用不同的膨胀卷积参数来提取不同感受野的图像特征,并进行多尺度特征融合选择出最适于图像超分恢复的特征,以进一步增强算法的超分性能。所提议的算法在公开的 Gaussian8 和 DIV2K²数据集上都取得了最佳的定量结果,同时具有更好的视觉效果,证明了算法的优越性。

本文的主要贡献可简要总结如下:

(1) 提出稀疏性和自相似性先验引导的深度学习图像盲超分算法,该算法结合了基于模型和基于深度学习方法的特点,实现了知识和数据的双驱动,具有精度高、泛化能力强和物理可解释性好的优点。

(2) 提出 DUDF 模块,该模块显式地对信号的稀疏性先验进行建模,并将优化目标展开为深度网

络进行求解,可高效完成反卷积操作,同时恢复出清晰锐利的图像特征。

(3) 提出 DMLR 模块,该模块充分利用图像的自相似性先验,采用多尺度大感受野特征融合策略,使算法在超分恢复时不仅关注小邻域信息,同时还可关注到更大空间范围相似的纹理信息,使最终恢复的图像有更准确的结构信息。

1 相关工作

1.1 降质模型的推导

在式(1)所描述的降质模型下,往往很难从低分辨率图像 y 中恢复出作用于潜在高分辨率图像 x 上的模糊核 k_h ^[23];同时现有算法^[16-20]直接使用神经网络端到端学习模糊核 k_h 和低分辨率图像 y 到高分辨率图像 x 的映射,缺乏明确物理意义的指导。对于式(1),可进行如下改写^[23]:

$$\begin{aligned} y &= \mathcal{F}^{-1}(\mathcal{F}((x * k_h)_{\downarrow s})) + n \\ &= \mathcal{F}^{-1}(\mathcal{F}(x_{\downarrow s}) \frac{\mathcal{F}((x * k_h)_{\downarrow s})}{\mathcal{F}(x_{\downarrow s})}) + n, \\ &= x_{\downarrow s} * \mathcal{F}^{-1}(\frac{\mathcal{F}((x * k_h)_{\downarrow s})}{\mathcal{F}(x_{\downarrow s})}) + n \end{aligned} \quad (2)$$

其中, $\mathcal{F}$ 和 $\mathcal{F}^{-1}$ 分别代表离散傅里叶变换及其逆变换。因此,可得到如下新的降质模型:

$$\begin{cases} y = x_{\downarrow s} * k_l + n \\ k_l = \mathcal{F}^{-1}(\frac{\mathcal{F}((x * k_h)_{\downarrow s})}{\mathcal{F}(x_{\downarrow s})}) \end{cases} \quad (3)$$

其中, $x_{\downarrow s}$ 表示对原始高分辨率图像 x 直接进行 s 倍降采样所得到的图像。

在新降质模型下,核估计任务变成从低分辨率图像 y 估计出作用在低分辨率图像 $x_{\downarrow s}$ 上的等价模糊核 k_l ,后续任务也被分解为了图像反卷积和图像经典超分两个清晰明确的操作,大大降低了图像盲超分任务的难度。

1.2 动态线性核估计(DLKE)模块

不同的低分辨率图像 y 所对应的 $x_{\downarrow s}$ 往往是不一样的,因此对于不同的输入 y ,核估计算法应产生相应的权重,依据具体的输入 y 产生相应的模糊核估计^[23]结果。

因动态神经网络可依据输入产生相应的权重,进而产生相应输出结果,更符合盲超分核估计的场景^[23],所以 DLKE 模块采用动态神经网络实现模糊核的估计^[23]。对每一张低分辨率图像,动态线性核估计 DLKE 模块(如图 1 所示)先使用残差模块提取

初步特征,再经全局池化层和全连接层为每张图像生成了相应的卷积权重,接着通过改变形状(Reshape)操作将生成的卷积权重转变为卷积核,最后使用狄拉克核依次与产生的卷积核进行卷积得到最终的模糊核估计结果。

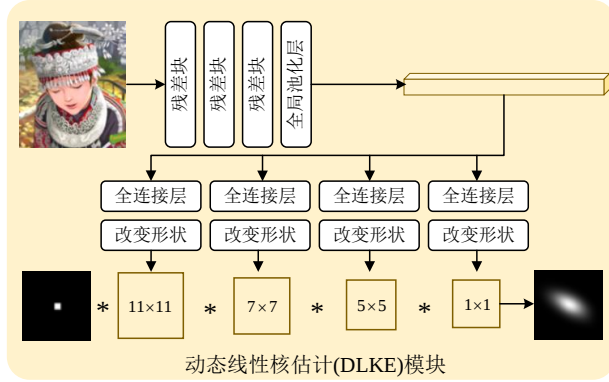


图1 DLKE模块的结构图

Fig. 1 The structure diagram of DLKE Module

2 模型构建

本节将详细阐述所提议的图像盲超分网络。它的整体结构如图2所示,主要由四部分构成:特征提取(Feature Extraction, FE)模块、动态线性核估计(DLKE)模块、基于FISTA算法的深度展开反卷积滤波(DUDF)模块和双通道多尺度大感受野恢复(DMLR)模块。其中,FE模块由卷积层、三个残差块和维度减少操作构成,DLKE模块已在第1节中进行阐述,因此便不再加以赘述。这里,我们先对DUDF模块的设计与实现进行详细阐述;然后,对DMLR模块的结构进行描述;最后,给出网络整体训练的优化目标函数并介绍算法的整体流程。

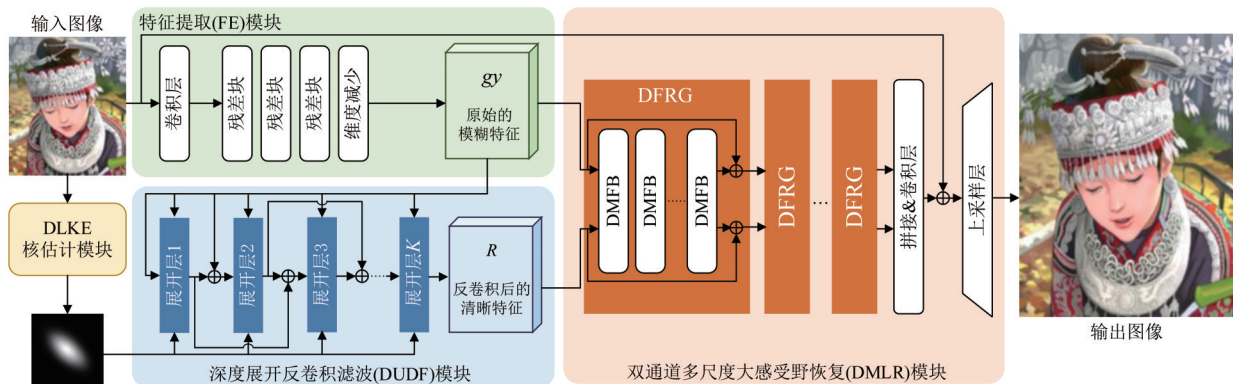


图2 DLS³P方法的整体框架

Fig. 2 The overall framework of DLS³P method

2.1 基于FISTA算法的深度展开反卷积滤波(DUDF)模块

DUDF模块的功能是在估计出模糊核 k_i 后对图像进行反卷积操作,进而得到清晰的特征。

文献^[26,27]已证明:在特征域而不是标准的图像域中进行反卷积操作,会更好地保持图像的结构并产生锐利的边缘。因此,所提议的DUDF模块在特征域上进行反卷积操作,在此情况下,降质模型(3)可表达为矩阵向量形式:

$$g_i y = k_i g_i x_{i_s} + g_i n, \quad (4)$$

其中, g 表示使用卷积神经网络对图像进行特征提取,下标 i 则表示卷积神经网络输出的第 i 张特征图, k_i 表示估计出的模糊核 k_i 所对应的块循环矩阵, y 及 x_{i_s} 是将相应的图像按照字典序展开为列向量形式。DUDF模块是从 $g_i y$ (即每一幅模糊特征图)中恢复出相应的清晰特征图 $g_i x_{i_s}$,接着送入到后续的DMLR模块进行超分恢复。

由于反卷积问题的病态本质,只有选择合适的先验,在优化目标中加入正则项,才能有效恢复出令人满意的清晰特征。但是,现有算法^[23]往往使用平滑性先验,将优化目标表示为

$$\min_{R_i} \frac{1}{2} \|g_i y - k_i R_i\|^2 + \lambda \|L R_i\|^2, \quad (5)$$

其中, λ 为正则项权重系数, L 为拉普拉斯算子所对应的块循环矩阵。该优化问题可由最小二乘法求得相应的解析解,然而却存在如下问题:(1)虽然平滑性先验已被证明在特征域中存在^[23],但仅用平滑性先验将会导致恢复出的特征过于单一,不利于后续超分恢复;(2)使用L2范数进行约束,虽可显式地求得解析解,但是,由于L2范数自身的性质使得最终所求解趋向于平均意义上的最优解,即会产生模

糊的图像边缘特征,丢失高频信息,从而导致最终的超分图像过于模糊。

为解决上述问题,我们使用信号的稀疏性先验来进行清晰特征的恢复。信号的稀疏性先验是指有效的信号在特定变换域上可由少量系数表征,即信号经过某些特定变换算子作用后会产生稀疏性,如离散余弦变换(Discrete Cosine Transform, DCT)、离散傅里叶变换(Discrete Fourier Transform, DFT)和离散小波变换(Discrete Wavelet Transform, DWT)算子等。与原先的拉普拉斯算子类似,在上述域中同样可以对解进行约束,限制解空间,得到令人满意的解。利用稀疏性先验进行信号的复原已在多个领域中被应用^[28-29]。具体地,我们将优化目标改写成:

$$\min_{R_i} \frac{1}{2} \|g_i y - k_i R_i\|^2 + \lambda \|A_i R_i\|_1, \quad (6)$$

其中, A_i 表示每一幅特征图的特定算子(通常为高通滤波算子)所对应的块循环矩阵。由于 L1 范数相较于 L2 范数可以保证解的稀疏性,反映在边缘信息上,意味着只有边缘上具有较大的像素值,靠近边缘的部分强迫其像素值很小,最终边缘就变成一条很细的线(理想情况下),即意味着在反卷积任务中可产生更清晰、锐利的边缘,有利于后续的超分恢复,这里,我们使用 L1 范数来约束。

由于 FISTA 算法^[24]是一种用于解决稀疏优化问题的迭代算法,相较于交替方向乘法(Alternating Direction Method of Multipliers, ADMM)和迭代软阈值收缩算法(Iterative Shrinkage-Thresholding Algorithm, ISTA)具有更快的收敛速率,这里,对该广义 LASSO 形式的优化问题,我们选择使用 FISTA 算法进行求解。对于每一幅模糊特征图 $g_i y$ 使用 FISTA 算法,设定初始值为 ρ_i^1 、 t_i^1 、 R_i^0 和 v_i^1 。具体来说, DUDF 模块使用神经网络可学习参数来初始化 ρ_i^1 和

t_i^1 , 并学习到最优的参数;同时,设置 R_i^0 和 v_i^1 的初始值为模糊特征图 $g_i y$, 接下来迭代求解式(7)-式(11)得到最终的清晰特征图:

$$u_i^k = v_i^k - \rho_i^k k_i^T (k_i v_i^k - g_i y), \quad (7)$$

$$R_i^k = \arg \min_{R_i} \frac{1}{2} \|R_i - u_i^k\|^2 + \lambda \|A_i R_i\|_1, \quad (8)$$

$$t_i^{k+1} = (1 + \sqrt{1 + 4(t_i^k)^2})/2, \quad (9)$$

$$\mu_i^k = (t_i^{k+1} + t_i^k - 1)/t_i^{k+1}, \quad (10)$$

$$v_i^{k+1} = \mu_i^k R_i^k + (1 - \mu_i^k) R_i^{k-1}, \quad (11)$$

DUDF 模块将上述的式(7)-(11)展开为深度神经网络形式, FISTA 算法的每一次迭代都对应着 DUDF 模块的一层神经网络层,其展开的每层具体结构如图3所示。

对于式(7),可将矩阵向量形式的等式转化为标准空间域的等式:

$$u_i^k = v_i^k - \rho_i^k \tilde{k}_i^* (k_i^* v_i^k - g_i y), \quad (12)$$

其中, v_i^k 是第 k 次迭代后 i 通道的清晰特征图, ρ_i^k 是第 k 次迭代中的步长系数(通过网络自动学习得到), k_i 是 DLKE 模块估计出的模糊核, $\tilde{k}_i$ 则是将 k_i 翻转 180° 后的模糊核, $g_i y$ 表示原始的模糊特征图, $*$ 表示二维卷积操作。依式(12),可求出每次迭代过程中的中间变量 u_i^k 。图3中的梯度下降块(Gradient Descent Block, GDB)对应着式(12)的求解。

对于式(8),使用近端映射块(Proximal Mapping Block, PMB)进行求解。具体来说,对于每个算子 A_i 使用卷积神经网络 $\mathcal{H}$ 来拟合算子 A_i , 同时,将原始矩阵向量形式的优化问题(8)转化为标准空间域中的优化问题:

$$\mathcal{R}_i^k = \arg \min_{\mathcal{R}_i} \frac{1}{2} \|\mathcal{R}_i - u_i^k\|^2 + \lambda \|\mathcal{H}(\mathcal{R}_i)\|_1, \quad (13)$$

由文献[30]有如下的近似关系:

$$\|\mathcal{H}(\mathcal{R}_i) - \mathcal{H}(u_i^k)\|^2 \approx \alpha \|\mathcal{R}_i - u_i^k\|, \quad (14)$$

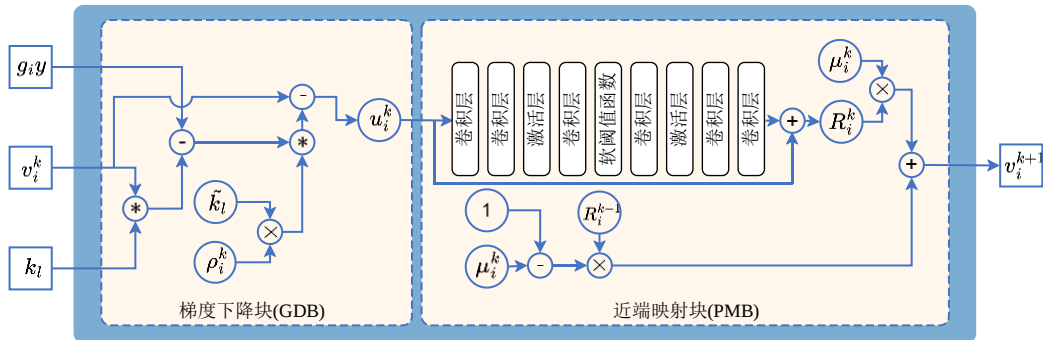


图3 FISTA算法一次迭代所对应的网络层结构

Fig. 3 Structure of a neural network layer that simulates one iteration of FISTA

其中, α 为常数。式(13)可转换为如下形式:

$$\mathcal{R}_i^k = \arg \min_{\mathcal{R}_i} \frac{1}{2} \|\mathcal{H}(\mathcal{R}_i) - \mathcal{H}(u_i^k)\|^2 + \theta \|\mathcal{H}(\mathcal{R}_i)\|_1, \quad (15)$$

其中, $\theta = \lambda\alpha$ 。式(15)有如下显式解:

$$\mathcal{R}_i^k = \tilde{\mathcal{H}}(\text{Soft}(\mathcal{H}(u_i^k), \theta)), \quad (16)$$

其中, Soft 为软阈值操作, $\mathcal{H}$ 和 $\tilde{\mathcal{H}}$ 为神经网络且满足 $\tilde{\mathcal{H}} \circ \mathcal{H} = \mathcal{I}$, 即对任意输入 f , 有 $\tilde{\mathcal{H}}(\mathcal{H}(f)) = f$; θ 为阈值, 由神经网络的可学习参数自动学习到最优的阈值。PMB 整体结构如图3所示。

具体来说, $\mathcal{H}$ 由相继的两层卷积层、激活函数层和卷积层构成; 为了满足条件 $\tilde{\mathcal{H}} \circ \mathcal{H} = \mathcal{I}$, $\tilde{\mathcal{H}}$ 的结构保持与 $\mathcal{H}$ 相对称, 由相继的卷积层、激活函数层和两层卷积层构成。同时, 引入对称损失:

$$\mathcal{L}_{\text{sym}} = \frac{1}{CHW} \sum_{i=1}^C \left\{ \sum_{m=1}^H \sum_{n=1}^W [u_i^1 - \tilde{\mathcal{H}}(\mathcal{H}(u_i^1))]_{mn}^2 \right\}, \quad (17)$$

其中, H 、 W 和 C 分别表示输入特征图的高、宽和通道数, 下标 m 和 n 表示特征图的第 m 行第 n 列像素, u_i^1 表示第 i 个通道特征图第一次迭代后的特征。为减少运算量, 利用网络训练时每批特征的多样性, 对称损失仅使用第一次迭代的特征 u_i^1 计算相应的损失实现网络更新, 可使网络 $\mathcal{H}$ 和 $\tilde{\mathcal{H}}$ 更好满足 $\tilde{\mathcal{H}} \circ \mathcal{H} = \mathcal{I}$ 。

对于式(9)和式(10), DUDF 并未按照原始 FIS-TA 算法更新系数 t_i^k , 而是使用网络可学习参数来学习每次迭代最终的加权系数 μ_i^k , 保证了算法的最优性及简洁性。对于式(11), 算法使用残差连接来完成特征的线性组合, 在此不再赘述。

2.2 双通道多尺度大感受野恢复(DMLR)模块

DMLR 模块主要是依据反卷积后的清晰特征进行图像的超分复原。如图2所示, DMLR 模块由若干双通道特征增强组(Dual-path Feature Refinement Group, DFRG)构成, DFRG 则由若干双通道多尺度融合块(Dual-path Multi-scale Fusion Block, DMFB)堆叠而成, DMFB 的结构图如图4所示。

DMLR 模块的输入包括反卷积前的模糊特征 gy 和反卷积后的清晰特征 $\mathcal{R}$ 。该模块额外引入反卷积前的模糊特征作为输入, 是为了消除 DLKE 模块带来的伪影现象^[23]。使用两种特征进行融合可有效消除伪影, 产生更清晰锐利的图像特征。

对于输入的每种特征, 首先被送入到 DMFB 中进行多尺度特征融合。现有的超分恢复模块往往使用单尺度特征, 且特征的感受野往往固定为 9×9

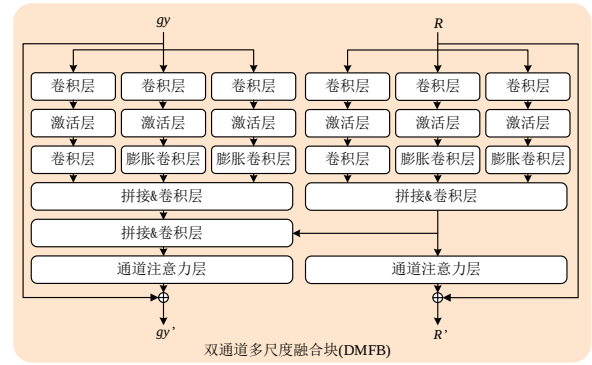


图4 DMFB的结构图

Fig. 4 The structure diagram of DMFB

大小, 网络只能利用邻域信息, 未能有效利用图像的自相似先验, 对于出现在图像各处相似的纹理不能有效地进行捕捉, 使得最终恢复的图像质量较差^[9]。由于相似的图像块通常以不同的大小重复出现于图像上(即图像的自相似先验), 为有效利用该先验, DMFB 首先使用膨胀卷积, 在有限参数量下提取 9×9 、 15×15 和 21×21 感受野的多尺度图像特征, 通过大感受野帮助特征进行长程关系建模, 以涵盖更广阔的空间信息, 捕捉图像中相似的纹理信息。最后, 使用拼接和卷积操作将多尺度特征进行融合, 为各个尺度、各个位置的特征赋予不同权重, 选择出最具表征性的图像特征, 重复迭代多次后可以得到清晰特征。这里, 最终的上采样层选用像素重排(PixelShuffle)层进行上采样。

2.3 优化目标和算法流程

DLS³P 算法的损失主要由三部分损失构成: 对称损失、核估计损失和像素损失。其中, 对称损失已在 2.1 节介绍, 在此主要介绍核估计损失和像素损失。

核估计损失是指 DLKE 模块估计出的模糊核 k_i 与真实模糊核 k_i^{gt} 间的 L1 损失:

$$\mathcal{L}_{\text{ker}} = \frac{1}{N} \|k_i - k_i^{gt}\|_1, \quad (18)$$

其中, N 为模糊核的像素数目。

像素损失则是指最终得到超分辨率图像 X_{sr} 与真实高分辨率图像 X_{hr} 间的 L1 损失:

$$\mathcal{L}_{\text{pix}} = \frac{1}{N'} \|X_{\text{sr}} - X_{\text{hr}}\|_1, \quad (19)$$

其中, N' 为超分辨率图像的像素数目。

综合以上, DLS³P 算法的整体损失函数为

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{ker}} + \mathcal{L}_{\text{pix}} + \mu \mathcal{L}_{\text{sym}}, \quad (20)$$

其中, μ 为对称损失权重的超参数, 在所有试验中固定为 0.01。

另外,表 1 给出了 DLS³P 训练过程的伪代码。首先,固定网络结构:确定 DUDF 模块的展开次数 K 、DFRG 的组数 G 和 DMFB 的块数 B ;接着选择训练过程的相关超参数:训练的总迭代次数 N 、损失函数中的超参数 μ 。然后,循环执行如下 6 个步骤 N 次:(1)将低分辨率图像输入到 DLKE 模块中,得到估计出的模糊核 k ;(2)使用 FE 模块提取低分辨率图像的特征,得到原始的模糊特征 gy ;(3)将估计出的模糊核 k 和模糊特征 gy 送入到 DUDF 模块中进行反卷积操作,得到清晰的特征 $\mathcal{R}$;(4)将模糊特征 gy 和清晰特征 $\mathcal{R}$ 送入到 DMLR 模块中进行超分恢复,得到最终的高分辨率图像;(5)依据估计出的模糊核计算核估计损失 $\mathcal{L}_{\text{ker}}$,依据恢复的高分辨率图像计算像素损失 $\mathcal{L}_{\text{pix}}$,依据 DUDF 模块的中间特征计算对称损失 $\mathcal{L}_{\text{sym}}$,并求出整体损失;(6)反向传播更新网络的参数。最终,我们得到训练好的网络。

算法测试过程的流程则由训练过程的步骤 1~步骤 4 构成,区别在于测试过程无需计算损失并进行反向传播,步骤 4 所得到超分辨率图像 SR 便是测试过程的输出结果,具体的每个步骤便不再赘述。

表 1 DLS³P 算法训练过程的伪代码

Table 1 The pseudocode for the DLS³P algorithm in training process

DLS ³ P 算法训练过程的伪代码
输入:低分辨率图像 LR
DUDF 模块展开层数 K ,DFRG 组数 G ,DMFB 块数 B
训练的总迭代次数 N ,对称损失权重 μ
循环迭代 N 次:
步骤 1: $k = \text{DLKE}(\text{LR})$
步骤 2: $gy = \text{FE}(\text{LR})$
步骤 3: $\mathcal{R} = \text{DUDF}(k, gy)$
步骤 4: $\text{SR} = \text{DMLR}(gy, \mathcal{R})$
步骤 5: $\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{ker}} + \mathcal{L}_{\text{pix}} + \mu \mathcal{L}_{\text{sym}}$
步骤 6:反向传播更新参数
输出:训练好的网络

3 实验结果与分析

3.1 实验数据和设置

数据集:依据先前工作^[16],实验选择公开的 DIV2K^①和 Flickr2K^②数据集作为训练数据集。DIV2K 数据集,源自 NITRE2017 超分辨率挑战赛,

由 800 幅训练集图片、100 幅验证集图片和 100 幅测试集图片组成,每幅图像都是 2K 分辨率图像。Flickr2K 数据集,同样也是图像超分领域中最流行的数据集之一,由 2 650 幅 2K 图像组成。实验使用 DIV2K 训练集和 Flickr2K 数据集共 3 450 幅 2 K 图像作为训练数据集,训练过程中,使用随机裁减策略从 2K 图像中裁剪高分辨率图像,依据经典降质模型来制作低分辨率图像,然后使用制作出的图像对模型进行训练。

依据降质模型中退化模糊核的不同属性,实验可分为退化模糊核分别为各向同性和各向异性高斯模糊核两种情形。下面将针对两种不同设置下的训练及测试细节进行详细介绍。

各向同性高斯模糊核设置:依据文献[16],该实验在各向同性高斯模糊核设置上开展图像盲超分实验。在该设置下,模糊核的大小固定为 21×21 。在训练过程中,针对 2、3 和 4 倍率的图像超分任务,模糊核的宽度分别从 $[0.2, 2.0]$ 、 $[0.2, 3.0]$ 和 $[0.2, 4.0]$ 范围中按均匀分布采样进行设置。对于测试过程,使用 Gaussian8 模糊核在常用的超分测试集 (Set5、Set14、BSD100、Urban100 和 Manga109 数据)上制作图像盲超分测试集。Gaussian8 数据集的模糊核针对 2、3 和 4 倍率的图像超分任务,分别从区间 $[0.8, 1.6]$ 、 $[1.35, 2.40]$ 和 $[1.80, 3.20]$ 上等间隔地选取 8 个模糊核的宽度。高分辨率图像相继经过 Gaussian8 模糊核的模糊和下采样操作,以得到最终的低分辨率图像并进行测试。

各向异性高斯模糊核设置:该实验将退化模糊核设置为各向异性模糊核。在此设置下,对于 4 倍率的图像盲超分任务,模糊核大小固定为 31×31 。在训练过程中,模糊核的宽度及旋转角度按照均匀分布分别从区间 $[0.6, 5.0]$ 和 $[-\pi, \pi]$ 中采样出来。由于 DIV2K 数据集包含了丰富内容和广泛图像来源,非常适合于评估图像超分算法在复杂退化条件下的性能,因此,在测试过程中,实验选用 DIV2K 数据集作为测试数据集。

3.2 评价指标和实验设置

在实验中,采用了两种评价指标来全面评估模型的超分效果:峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 与结构相似度 (Structural Similarity Index, SSIM)。PSNR 作为衡量图像质量的传统标准,

①¹ <https://data.vision.ee.ethz.ch/cvl/DIV2K>

②² <https://cv.snu.ac.kr/research/EDSR/Flickr2K.tar>

专注于量化图像与原始参考图像间在像素级别上的相似程度,是评估超分算法恢复图像细节能力的重要指标。而SSIM不仅考虑了像素值的差异,还融入了对比度和结构信息的考虑,使得评价更为全面且贴近人类视觉感知系统,从而成为评估图像恢复质量不可或缺的工具。评估算法效果时,将图片转化为YCbCr格式,在Y分量上计算相应的PSNR和SSIM值。另外,采用总体参数数量和浮点运算次数(Floating Point Operations, FLOPs)来分别衡量模型的参数量(空间复杂度)和运算量(计算复杂度)。

对于所有设置下的实验,算法DFRG的个数 G 为5,每组DFRG中的DMFB的个数 B 为10。训练中每批次大小设置为64,低分辨率图像被随机裁剪为 64×64 大小送入网络。训练时使用Adam优化器,参数 β_1 和 β_2 分别设置为0.9和0.999。所有模型都使用4×RTX3090 GPUs训练了 6×10^5 次迭代。初始学习率设置为 2×10^{-4} ,并且每 1×10^5 次迭代后,学习率衰减为原先的1/2。

3.3 与其他方法的对比

3.3.1 各向同性高斯模糊核设置

为验证所提出算法在各向同性高斯模糊核设置下的图像超分效果,将所提出算法与当前的其他图像盲超分算法进行对比。对比方法包括:两种经典超分算法(Bicubic和CARN^[4])、三种双阶段算法(Deblur^[31]+CARN^[4]、CARN^[4]+Deblur^[31]和Bicubic+ZSSR^[13])和五种端到端算法(IKC^[16]、DANv1^[19]、DANv2^[20]、AdaTarget^[17]和DCLS^[23])。

表2展示了在Gaussian8数据集上不同方法的定量结果,其最好和次好两个结果分别用粗体和下划线标记。不难看到,在面对Gaussian8数据集中复杂退化场景时,由于预设退化模型过于简化,经典超分算法(Bicubic和CARN)的性能有明显局限性。对于双阶段方法,Deblur+CARN和CARN+Deblur使用去模糊算法Deblur对经典超分的结果进行前处理或后处理,但由于算法的不匹配性而产生了不佳的超分效果;Bicubic+ZSSR算法直接默认设置模糊核为双三次降采样核,却忽略了实际中模糊核的多样性,加之其较长的推理时间,使该算法在性能上仍显不足。对于端到端方法,IKC、AdaTarget及DAN系列算法通过端到端的优化机制,显著提升了超分效果,然而,上述算法在处理模糊核信息时,多采取直接嵌入网络的方式,未能充分挖掘和利用模糊核的深层信息;DCLS通过改写降质模型,更有效

地利用了估计出的模糊核信息,然而,由于其采用最小二乘滤波策略,以及未能有效利用图像的自相似性先验,在一定程度上带来了更模糊的超分结果。本文算法则巧妙地融入了稀疏性与自相似性先验。通过显式建模稀疏先验,不仅更有效地利用了模糊核信息,还激发了图像中更为丰富的特征表达;同时通过多尺度大感受野特征融合策略充分利用了自相似先验。在稀疏和自相似双重先验的联合指导下进行端到端的优化,本文方法最终生成了更为清晰、锐利的超分辨率图像,在Gaussian8数据集上取得了最佳的超分效果,体现了所提出算法设计的先进性与有效性。值得注意的是,我们的算法在Set5数据集,×2和×3超分倍率上的定量效果不如DCLS方法,其主要原因是Set5数据集只有五张图像,该数据集的图片数量较少;同时,五张图像的高频细节较少,超分所需要恢复的纹理较少,因此现有的DCLS算法也足以产生令人满意的效果。对于Gaussian8中的其他子数据集,这些数据集的尺寸更大,图片种类更加多样,因而更具挑战性,更能衡量算法的综合性能。所提议的算法在这些数据集上都取得了最高的定量指标,证明了我们算法的优越性和可靠性。

3.3.2 各向异性高斯模糊核设置

各向异性模糊核设置的图像盲超分任务往往更加具有挑战性,更能衡量算法性能。将所提出DLS³P算法与四种经典超分算法(Bicubic、DBPN^[3]、EDSR^[2]和RCAN^[5])、三种双阶段算法(Bicubic+ZSSR^[13]、KernelGAN^[9]+ZSSR^[13]、KernelGAN^[9]+SRMD^[14]和KernelGAN^[9]+USRNet^[15])和七种端到端算法(IKC^[16]、DANv1^[19]、DANv2^[20]、AdaTarget^[17]、KOALANet^[18]、DCLS^[23]和KernelDAN^[21])进行性能比较。

表3展示了各种算法在DIV2K数据集上的定量结果,其最好和次好两个结果也分别用粗体和下划线标记。不难看出,所提议DLS³P算法在PSNR和SSIM指标上相对于次优算法取得了大幅提升。在面对更为复杂的退化场景时,经典超分算法(Bicubic、DBPN、EDSR和RCAN)性能出现了明显的下降。对于双阶段算法,Bicubic+ZSSR直接假设模糊核为双三次降采样核,本质上与经典超分算法的设置相同,因而限制了其性能上限;通过引入KernelGAN进行模糊核的预估计,KernelGAN+ZSSR性能超越了经典超分算法,但效果仍显不足;进一步,

表2 不同算法在 Gaussian8数据集上的定量结果

Table 2 Quantitative results of different algorithms on Gaussian8 dataset

Method	Scale	Set5		Set14		BSD100		Urban100		Manga109		Average	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bicubic	2	28.82	0.8577	26.02	0.7634	25.92	0.7310	23.14	0.7258	25.60	0.8498	25.90	0.7855
CARN ^[4]		30.99	0.8779	28.10	0.7879	26.78	0.7286	25.27	0.7630	26.86	0.8606	27.60	0.8036
Deblur ^[31] +CARN ^[4]		24.20	0.7496	21.12	0.6170	22.69	0.6471	18.89	0.5895	21.54	0.7946	21.69	0.6796
CARN ^[4] +Deblur ^[31]		31.27	0.8974	29.03	0.8267	28.72	0.8033	25.62	0.7981	29.58	0.9134	28.84	0.8478
Bicubic+ZSSR ^[13]		31.08	0.8786	28.35	0.7933	27.92	0.7632	25.25	0.7618	28.05	0.8769	28.13	0.8148
IKC ^[16]		37.19	0.9526	32.94	0.9024	31.51	0.8790	29.85	0.8928	36.93	0.9667	33.68	0.9187
DANv1 ^[19]		37.34	0.9526	33.08	0.9041	31.76	0.8858	30.60	0.9060	37.23	0.9710	34.00	0.9239
DANv2 ^[20]		37.60	0.9544	33.44	0.9094	32.00	0.8904	31.43	0.9174	38.07	0.9734	34.51	0.9290
DCLS ^[23]		<u>37.63</u>	<u>0.9554</u>	<u>33.46</u>	<u>0.9103</u>	<u>32.04</u>	<u>0.8907</u>	<u>31.69</u>	<u>0.9202</u>	<u>38.31</u>	<u>0.9740</u>	<u>34.63</u>	<u>0.9301</u>
DLS ³ P (Ours)		37.64	<u>0.9546</u>	33.65	0.9105	32.11	0.8916	31.90	0.9220	38.51	0.9744	34.76	0.9306
Bicubic	3	26.21	0.7766	24.01	0.6662	24.25	0.6356	21.39	0.6203	22.98	0.7576	23.77	0.6913
CARN ^[4]		27.26	0.7855	25.06	0.6676	25.85	0.6566	22.67	0.6323	23.85	0.7620	24.94	0.7008
Deblur ^[31] +CARN ^[4]		19.05	0.5226	17.61	0.4558	20.51	0.5331	16.72	0.5895	18.38	0.6118	18.45	0.5426
CARN ^[4] +Deblur ^[31]		30.31	0.8562	27.57	0.7531	27.14	0.7152	24.45	0.7241	27.67	0.8592	27.43	0.7816
Bicubic+ZSSR ^[13]		28.25	0.7989	26.15	0.6942	26.06	0.6633	23.26	0.6534	25.19	0.7914	25.78	0.7202
IKC ^[16]		33.06	0.9146	29.38	0.8233	28.53	0.7899	24.43	0.8302	32.43	0.9316	29.57	0.8579
DANv1 ^[19]		34.04	0.9199	30.09	0.8287	28.94	0.7919	27.65	0.8352	33.16	0.9382	30.78	0.8628
DANv2 ^[20]		34.12	0.9209	30.20	0.8309	29.03	0.7948	27.83	0.8395	33.28	0.9400	30.89	0.8652
DCLS ^[23]		34.21	<u>0.9218</u>	<u>30.29</u>	<u>0.8329</u>	<u>29.07</u>	<u>0.7956</u>	<u>28.03</u>	<u>0.8444</u>	<u>33.54</u>	<u>0.9414</u>	<u>31.03</u>	<u>0.8672</u>
DLS ³ P (Ours)		<u>34.15</u>	0.9219	30.31	0.8341	29.12	0.7979	28.21	0.8486	33.81	0.9429	31.12	0.8690
Bicubic	4	24.57	0.7108	22.79	0.6032	23.29	0.5786	20.35	0.5532	21.50	0.6933	22.50	0.6278
CARN ^[4]		26.57	0.7420	24.62	0.6226	24.79	0.5963	22.17	0.5865	21.85	0.6834	24.00	0.6462
Deblur ^[31] +CARN ^[4]		18.10	0.4843	16.59	0.3994	18.46	0.4481	15.47	0.3872	16.78	0.5371	17.08	0.4512
CARN ^[4] +Deblur ^[31]		28.69	0.8092	26.40	0.6926	26.10	0.6528	23.46	0.6597	25.84	0.8035	26.10	0.7236
Bicubic+ZSSR ^[13]		26.45	0.7279	24.78	0.6268	24.97	0.5989	22.11	0.5805	23.53	0.7240	24.37	0.6516
IKC ^[16]		31.67	0.8829	28.31	0.7643	27.37	0.7192	25.33	0.7504	28.91	0.8782	28.32	0.7990
DANv1 ^[19]		31.89	0.8864	28.42	0.7687	27.51	0.7248	25.86	0.7721	30.50	<u>0.9037</u>	28.84	0.8111
DANv2 ^[20]		32.00	0.8885	28.50	0.7715	27.56	0.7277	25.94	0.7748	30.45	<u>0.9037</u>	28.89	0.8132
AdaTarget ^[17]		31.58	0.8814	28.14	0.7626	27.43	0.7216	25.72	0.7683	29.97	0.8955	28.57	0.8059
DCLS ^[23]		<u>32.12</u>	<u>0.8890</u>	<u>28.54</u>	<u>0.7728</u>	<u>27.60</u>	<u>0.7285</u>	<u>26.15</u>	<u>0.7809</u>	<u>30.86</u>	0.9086	<u>29.05</u>	<u>0.8160</u>
DLS ³ P (Ours)		32.13	0.8906	28.62	0.7752	27.63	0.7305	26.24	0.7849	30.94	0.9086	29.11	0.8180

KernelGAN+SRMD 结合了更为先进的非盲超分技术 SRMD,性能有了一定的提升;KernelGAN+USRNet 结合了当前最为先进的非盲超分算法 USRNet,但 USRNet 对模糊核过于敏感,微小的模糊核估计误差带来了算法巨大的性能下降,体现了双阶段算法的不适配性的缺点。对于端到端的盲超分算法,IKC、AdaTarget 和 KOALAnet 使超分性能进一步增强,但这些算法未能充分利用模糊核的有效信息,因而无法达到最优性能;DAN 系列中最新算法 KernelDAN 相对先前版本的算法有了较大性能提升,但只是加入了特征提取器,在特征域中进行图像复原,仍未能有效利用模糊核的信息,因而其性

能没有本质提升;DCLS 通过改写降质模型,有效地利用了模糊核的信息,但是最小二乘滤波带来了更模糊的超分结果,小感受野特征也使得算法只专注于局部信息。相比之下,我们提出的 DLS³P 算法,创新性地 将稀疏性和自相似性先验融入到深度学习方法中并进行端到端的优化,同时利用了知识驱动与数据驱动优势,因而取得了最高的定量指标。

3.4 消融实验分析

3.4.1 DUDF 模块迭代次数对超分结果的影响

我们将 FISTA 算法深度展开为 DUDF 模块,而 FISTA 算法作为一种迭代算法,其迭代次数往往会影响算法能否收敛到一个较优解,因此,我们考察

表3 不同算法在DIV2KRR数据集上的定量结果

Table 3 Quantitative results of different algorithms on DIV2KRR dataset

Method	DIV2KRR	
	PSNR	SSIM
Bicubic	25.33	0.679 5
EDSR ^[2] (2017)	25.64	0.692 8
DBPN ^[3] (2018)	25.58	0.691 0
RCAN ^[5] (2018)	25.66	0.693 6
Bicubic+ZSSR ^[13] (2019)	25.61	0.691 1
KernelGAN ^[9] +ZSSR ^[13] (2019)	26.81	0.731 6
KernelGAN ^[9] +SRMD ^[14] (2019)	27.51	0.726 5
KernelGAN ^[9] +USRNet ^[15] (2020)	20.06	0.535 9
IKC ^[16] (2019)	27.70	0.766 8
DANv1 ^[19] (2020)	27.55	0.758 2
DANv2 ^[20] (2021)	28.74	0.789 3
AdaTarget ^[17] (2021)	28.42	0.785 4
KOALAnet ^[18] (2021)	27.77	0.763 7
DCLS ^[23] (2022)	28.99	0.794 6
KernelDAN ^[21] (2023)	28.90	0.796 1
DLS ³ P (Ours)	29.08	0.800 7

不同迭代次数对超分结果的影响。为了节省实验耗时,选择在各向同性高斯模糊核设置下进行实验,采用原始的双通道单尺度特征恢复模块,迭代次数从2开始逐渐递增。图5展示了Urban100和Manga109数据集上的PSNR值随迭代次数的变化曲线。

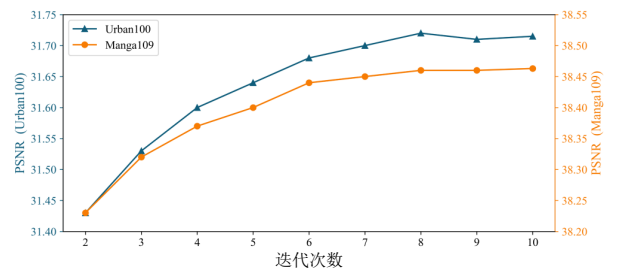


图5 在Urban100和Manga109数据集上PSNR值随迭代次数变化曲线

Fig. 5 PSNR values variation curve with iterations on Urban100 and Manga109 datasets

由图5不难发现:在迭代次数较为有限的情况下,FISTA算法未能较好的收敛,导致反卷积处理后的特征依旧保留着明显模糊痕迹,最终的PSNR值较低;当迭代次数提升至大约8次时,算法展现出了优异的收敛性能,收敛到一个较优解,使得反卷积后的特征细节清晰可辨,超分辨率图像的PSNR值

也随之显著提升。值得注意的是,当进一步增加迭代次数时,PSNR指标趋于饱和,几乎不再有明显提升,这揭示了算法在迭代次数为8次时已达到性能与效率的较佳平衡点。因此,综合考虑算法效果与效率,本实验最终固定迭代次数为8次,以此作为平衡两者需求的最佳选择。

更进一步,与传统基于模型的迭代算法相比,DUDF模块需要迭代的次数明显较小,这一显著优势主要归功于深度展开网络方法的独特之处。由于深度学习是在图像特征空间而不是在标准空间域内对问题进行求解,有效降低了问题复杂度,还凭借其数据驱动的学习模式,在有限迭代次数内便能迅速捕捉到图像的内在深层特征,从而实现高效且高质量的收敛效果。

3.4.2 DUDF模块反卷积通道数对超分结果的影响

DUDF模块对每一通道的模糊特征图进行单独反卷积操作,对每一通道得到相应的清晰特征图,因此,反卷积特征的通道数代表着用于后续超分的清晰特征的数量,对最终超分结果有重要影响。

在实验中,固定DUDF模块迭代次数为8,采用原始的双通道单尺度特征恢复模块,反卷积特征通道数分别设置为16、32和64。表4展示了在Urban100数据集上PSNR值随反卷积特征通道数的变化情况,其中基准算法选择为DCLS算法,我们的算法相对于DCLS算法来说,是将DCLS中的最小二乘滤波模块替换为DUDF模块,与DCLS算法的比较能更客观地反映出所提议DUDF模块的有效性。

表4 PSNR值随反卷积特征通道数的变化

Table 4 PSNR values variation with the number of deconvolution feature channels

反卷积特征通道数	16 (DCLS)	16	32	64
Urban100	31.69	31.72	31.78	31.80

通过表4不难发现,当反卷积特征通道数仅为16时,本文算法相较于基准算法提升相对较小,这主要是因为当反卷积通道数固定为16时,反卷积后仅得到16幅清晰特征,其表征能力仍然有限,因此最终恢复图像PSNR值也相对较低;随着反卷积特征通道数增加到32,特征表征能力和多样性大大增加,进而最终的PSNR值也得到了较大提升;虽然通道数持续增大仍会带来收益,但收益较小,相反会

使得网络参数量急剧增大,因此,最终我们固定反卷积通道数为32。值得注意的是,DCLS算法固定其反卷积特征通道数为16,主要是因为DCLS算法对每一通道特征图采用离散傅里叶变换进行最小二乘滤波,会产生极大的计算开销,然而我们的算法使用全卷积架构大大减少了计算开销,因此,可在相近时间开销下对更多特征通道数进行反卷积操作,以获取更丰富的特征表征。

3.4.3 DMLR 模块多尺度特征对超分结果的影响

本节主要探究DMLR模块中多尺度特征对超分结果的影响。更大感受野和更多尺度的特征往往可以涵盖图像上更广阔的空间信息,其充分利用了图像的自相似先验,对最终超分结果有重要影响。

在实验中,固定DUDF模块的迭代次数为8,反卷积特征通道数设置为32,我们在原始9×9单感受野特征上依次加入15×15和21×21感受野的特征图。表5展示了随着不同感受野特征的加入Urban100数据集上PSNR值的变化。

表5 PSNR 值随不同感受野特征的变化

Table 5 PSNR values variation with different receptive field features

9×9 特征	15×15 特征	21×21 特征	Urban100
√	×	×	31.78
√	√	×	31.85
√	√	√	31.90

由表5不难看出,随着多尺度特征的引入,所提算法在Urban100数据上PSNR值取得了显著增加,这主要是由于多尺度大感受野的特征能充分利用图像的自相似先验,即相似的图像块以不同的大小重复出现于图像上,因而不同感受野的特征可以有效地捕获块重复信息。

3.4.4 参数量和运算量分析

本节对所提出模型的参数量和运算量进行分析。表6分别对比了目前常见模型及本文模型的参数量和FLOPs,其中FLOPs是在输入图像的尺寸为270×180条件下进行统计的。

由表6可看到,尽管本文所提出算法在超分辨率重建方面展现出了最优异的效果,但其模型参数量也相对较大,这一特性主要归因于DMLR模块中的多尺度特征融合机制,该机制虽显著提升了重建质量,却不可避免地增加了参数规模。

表6 不同模型的参数量和计算量对比

Table 6 Comparison of the parameter and computational complexity among the different models

方法	参数量 (M)	FLOPs (G)
IKC (2019)	5.29	2178.72
DANv1 (2020)	4.33	926.72
DANv2 (2021)	4.71	918.12
DCLS (2022)	19.05	—
DLS ³ P (Ours)	34.40	1224.00

在运算效率方面,本文算法与经典算法如DANv1^[19]和DANv2^[20]保持了相当的水平。值得注意的是,DCLS^[23]算法因使用离散傅里叶变换,其浮点运算次数(FLOPs)难以直接量化评估,这在一定程度上限制了对其实算效率的直观比较。仅从实际推理时间进行考察,本文算法与当前最先进的盲超分辨率算法(如DANv1^[19]、DANv2^[20]以及DCLS^[23])均保持了相近的推理速度,充分验证了本文算法在实际应用中的有效性。

在服务器上部署时,鉴于硬件资源的充足性,本文算法能充分发挥其性能优势,有效应对真实世界中低分辨率图像的挑战,恢复出令人满意的高分辨率图像质量。另外,我们的后续研究将聚焦于模型的优化与轻量化设计,旨在通过减少模型参数与计算复杂度,使得该算法能够在计算资源和存储空间受限的端侧设备上实现成功部署,从而进一步拓宽其应用场景,满足更多元化的需求。

3.4.5 可视化结果

本节将对算法恢复出的超分图像进行可视化的展示并进行分析,与实验设置一样,我们可视化的结果也分为各向同性和各向异性高斯模糊核两种设置。

图6展示了各向同性设置下的可视化结果。可视化图片来自Urban100建筑物数据集,由于建筑物图像细节纹理较难恢复,Urbn100数据集盲超分任务往往极具挑战。实验是在放大倍率为2,模糊核宽度设置为1.6的情况下进行的。我们选取了2倍率下的最大模糊核宽度,意味着低分辨率图像受到最严重的退化,在严重退化图像上的超分结果更能衡量算法的性能。

如图6所示,本文算法取得了最佳视觉观感效果。在绿色方框的复原图像中,本文算法恢复出了最为清晰锐利的边缘(清晰的镜子轮廓线),这主要是因为算法充分利用稀疏性先验反卷积恢复了清

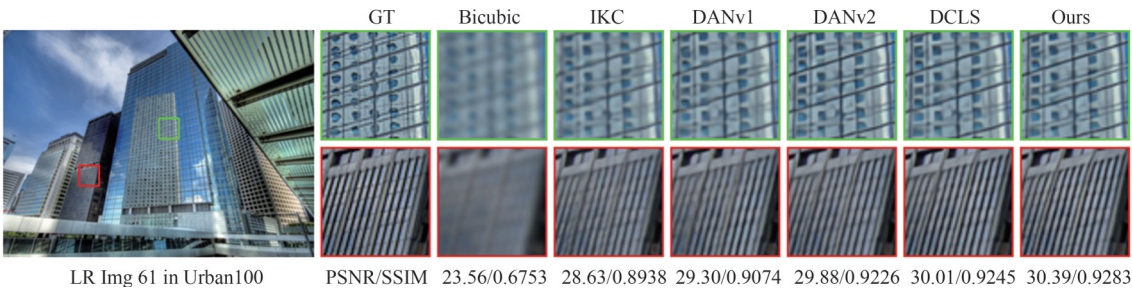


图6 Urban100数据集图像61的可视化结果
Fig. 6 Visual results of Img 61 in Urban100 dataset

晰的图像特征,进而恢复出了最为锐利的图像;DCLS取得了次优的结果,这主要是因为最小二乘滤波恢复出的图像往往是比较模糊的;IKC及DAN系列算法恢复出的结果则更为模糊,因为它们并没有充分利用模糊核的空间信息。在红色方框中,本文算法恢复出了最为真实准确的建筑物外表面形状(平行的建筑物外表轮廓线);而DCLS等算法恢复出的轮廓线则出现了相交和重合现象,显然是不符合我们预期的。这主要是因为本文算法采用了多尺度大感受野特征融合策略,能关注到更大空间范围内的参照信息,有效利用了图像的自相似先验,因而取得了最为真实准确的恢复效果。

图7展示了各向异性设置下的可视化结果。如图7所示,本文算法取得了最佳视觉效果。在绿色方框的复原图像中,本文算法恢复出了最为准确的图像结构,即白色集装箱上的网格结构(绿色箭头所示);相较之下,其余算法恢复出的纹理过于模糊,结构遭到损坏。这主要是由于本文算法的多尺度大感受野特征融合策略捕获到最上侧集装箱相似的纹理结构,提供了超分恢复的参照信息。在红色方框中,本文算法取得最为清晰锐利的图像边缘(红色箭头所示),这主要是因为本文算法充分利用了稀疏性先验,而稀疏性先验在超分恢复任务中会带来清晰锐利的图像纹理,因而取得了最佳的视觉

感观效果。
3.4.6 小结

通过与目前SOTA方法的性能对比以及消融实验结果分析,所提出DLS³P算法具有如下主要优点:

(1) 精度高:本文算法取得了最高的定量指标,超过了目前SOTA图像盲超分方法。这一优点主要归功于本文算法结合了基于模型和基于深度学习两种方法的优点,在深度学习中充分融入图像的稀疏性和自相似性先验知识,并结合端到端的优化策略,实现了知识与数据双驱动图像盲超分算法。

(2) 泛化能力强:本文算法在包括自然景致、城市建筑以及精致动漫场景的Gaussian8和DIV2KRR数据集上都取得了最高的定量指标,展现出了强大的泛化能力。这主要是因为本文算法将稀疏性与自相似性这两种先验知识融入网络设计之中,赋予了算法更好的物理可解释性,使其能在不同领域的图像数据上充分运用这些先验知识,从而实现广泛而精准的超分辨率重建效果。

(3) 超参数易调节:在DLS³P中,DUDF模块的迭代次数是需要调节的超参数。实验中,本文算法在各个不同的设置和不同的数据集上均采用了相同参数设置,均取得了良好的超分结果,不用针对不同的设置和数据集重新进行超参数调节,因而本文算法具有超参数易调节的特点。

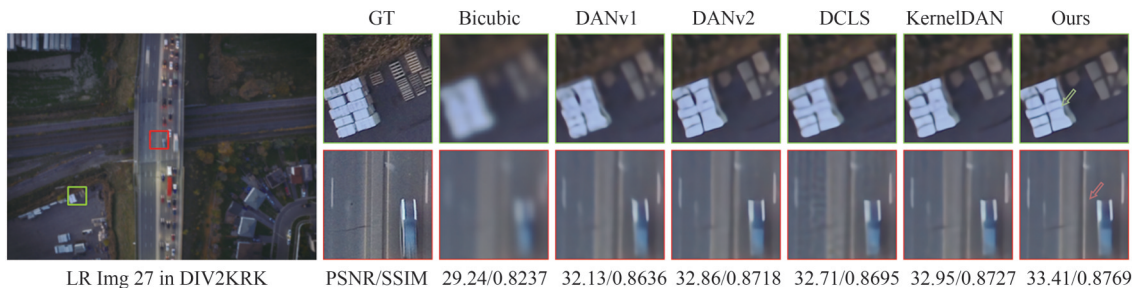


图7 DIV2KRR数据集的图像27可视化结果
Fig. 7 Visual results of Img 27 in DIV2KRR dataset

另外,DLS³P算法的参数量相对仍较大,这主要源于DMLR模块所使用的多尺度特征融合机制,该机制通过引入大量卷积层参数,实现了对图像深层次、多尺度特征的精细捕捉与融合,从而极大地提升了图像重建质量。然而,这一操作也伴随着参数规模的显著增长,对存储资源提出了更高要求。未来,我们将聚焦于DMLR模块的轻量化设计,以进一步大大降低算法参量的复杂度。

4 结论

本文提出了一种稀疏性和自相似性先验引导的深度学习图像盲超分算法(即DLS³P)。在算法中,我们精心设计了一个多模块协作的网络框架,其中DLKE模块从低分辨率图像中精确估计出模糊核,DUDF和DMLR模块则分别引入稀疏性和自相似性先验,有效地实现了反卷积与超分恢复任务,产生了令人满意的图像超分结果。所提议方法巧妙地将稀疏性和自相似性先验知识融入深度学习框架之中,实现了知识驱动与数据驱动策略的优势互补,具有精度高、泛化能力强和超参数易调节等优点,在实际应用中有更为重要的应用价值和意义。

在未来的工作中,我们将聚焦于DMLR模块的轻量化设计,通过采用先进的模型压缩技术(如知识蒸馏、剪枝策略等),大大减少DMLR模块参数规模,以降低算法的参量复杂度。

References

- [1] Wang Z, Chen J, Hoi S C H. Deep learning for image super-resolution: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(10): 3365–3387.
- [2] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super-resolution[C]. IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017.
- [3] Haris M, Shakhnarovich G, Ukita N. Deep back-projection networks for super-resolution[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [4] Ahn N, Kang B, Sohn K A. Fast, accurate, and lightweight super-resolution with cascading residual network[C]. European Conference on Computer Vision, 2018.
- [5] Zhang Y, Li K, Li K, et al. Image super-resolution using very deep residual channel attention networks[C]. European Conference on Computer Vision, 2018.
- [6] Liu A, Liu Y, Gu J, et al. Blind image super-resolution: a survey and beyond[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(5): 5461–5480.
- [7] Wang Q, Tang X, Shum H. Patch based blind image super resolution[C]. IEEE International Conference on Computer Vision, 2005.
- [8] He Y, Yap K H, Chen L, et al. A soft MAP framework for blind super-resolution image reconstruction[J]. Image and Vision Computing, 2009, 27(4): 364–373.
- [9] Bell-Kligler S, Shocher A, Irani M. Blind super-resolution kernel estimation using an internal-GAN[C]. Advances in Neural Information Processing Systems, 2019.
- [10] Goodfellow I, Pouget-Abadie J, Mirza M, et al. GAN (Generative Adversarial Nets)[J]. arXiv preprint arXiv: 1406.2661, 2014.
- [11] Yamac M, Ataman B, Nawaz A. KernelNet: a blind super-resolution kernel estimation network[C]. IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2021.
- [12] Liang J, Zhang K, Gu S, et al. Flow-based kernel prior with application to blind super-resolution[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2021.
- [13] Shocher A, Cohen N, Irani M. Zero-shot super-resolution using deep internal learning[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [14] Zhang K, Zuo W, Zhang L. Learning a single convolutional super-resolution network for multiple degradations[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [15] Zhang K, Van Gool L, Timofte. Deep unfolding network for image super-resolution[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2020.
- [16] Gu J, Lu H, Zuo W, et al. Blind super-resolution with iterative kernel correction[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2019.
- [17] Jo Y, Oh S W, Vajda P, et al. Tackling the ill-posedness of super-resolution through adaptive target generation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2021.
- [18] Kim S Y, Sim H, Kim M. KOALAnet: Blind super-resolution using kernel-oriented adaptive local adjustment[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2021.
- [19] Luo Z, Huang Y, Li S, et al. Unfolding the alternating optimization for blind super resolution[C]. Advances in Neural Information Processing Systems, 2020.
- [20] Luo Z, Huang Y, Li S, et al. End-to-end alternating optimization for blind super resolution[J]. arXiv preprint arXiv:2105.06878, 2021.
- [21] Luo Z, Huang Y, Li S, et al. End-to-end alternating optimization for real-world blind super resolution[J]. International Journal of Computer Vision, 2023, 131(12): 3152–3169.
- [22] Xu Z, Zhang Y, Luo T, et al. Frequency principle: fourier analysis sheds light on deep neural networks[J]. arXiv preprint arXiv:1901.06523, 2019.
- [23] Luo Z, Huang H, Yu L, et al. Deep constrained least squares for blind image super-resolution[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2022.
- [24] Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems[J]. SIAM Journal on Imaging Sciences, 2009, 2(1): 183–202.

-
- [25] Ali A, Tibshirani R J. The generalized lasso problem and uniqueness[J]. *Electronic Journal of Statistics*, 2019, 13 (2): 2307–2347.
- [26] Son H, Lee S. Fast non-blind deconvolution via regularized residual networks with long/short skip-connections [C]. *IEEE International Conference on Computational Photography*, 2017.
- [27] Dong J, Roth S, Schiele B. DWDN: Deep wiener deconvolution network for non-blind image deblurring[J]. *IEEE Transactions on Pattern Analysis and Machine intelligence*, 2022, 44(12): 9960–9976.
- [28] Guerquin-Kern M, Haberland M, Pruessmann K P, et al. A fast wavelet-based reconstruction method for magnetic resonance imaging[J]. *IEEE Transactions on Medical Imaging*, 2011, 30(9): 1649–1660.
- [29] Wang, Shuihua, Dong, et al. Exponential wavelet iterative shrinkage thresholding algorithm for compressed sensing magnetic resonance imaging [J]. *Information Sciences*, 2015, 322: 115–132.
- [30] Zhang J, Ghanem B. ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2018.
- [31] Pan J, Sun D, Pfister H, et al. Deblurring images via dark channel prior[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(10): 2315–2328.