

文章编号: 1671-0576(2023)01-0045-05

超分辨光学图像深度展开网络

史梦阳¹, 张 弦², 王 丽³, 施君南², 高叶盛¹

(1. 上海交通大学电子工程系, 上海 200240; 2. 上海无线电设备研究所, 上海 201109;
3. 中国人民解放军 96962 部队, 北京 102206)

摘 要: 提出了一种用于光学图像超分辨的深度展开网络。该网络将传统优化算法和深度学习算法相结合, 使得最终的网络既具有可解释性, 又拥有神经网络的强大拟合性。为了增强网络的目标图像多尺度特征提取能力, 提出了多尺度空洞卷积注意力模块, 该模块可以有效地提取目标图像的多尺度特征, 并对重要特征赋予较大的权重, 提高网络性能。实验表明, 该网络能够有效恢复图像的细节信息, 实现降质图像较好的超分辨重建效果。

关键词: 超分辨光学图像; 深度展开网络; 可解释; 多尺度; 深度学习

中图分类号: TP391

文献标志码: A

DOI: 10.3969/j.issn.1671-0576.2023.01.008

Deep Unfolding Network for Super-resolution Optical Image

SHI Mengyang¹, ZHANG Xian², WANG Li³, SHI Junnan², GAO Yesheng¹

(1. Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; 2. Shanghai Radio Equipment Research Institute, Shanghai 201109, China;
3. Unit 96962 of People's Liberation Army of China, Beijing 102206, China)

Abstract: A deep unfolding network for super-resolution optical image was presented. The network combines traditional optimization algorithm with deep learning algorithm to make the final network interpretable and fit well with the neural network. To enhance the ability of extracting multi-scale features from target images, a multi-scale void convolution attention module was proposed. The module can effectively extract multi-scale features from target images, and give larger weights to important features to improve network performance. The experimental results show that the network can effectively restore the image details and achieve a better super-resolution reconstruction of degraded images.

Key words: super-resolution optical image; deep unfolding network; interpretable; multi-scale; deep learning

收稿日期: 2022-08-16

基金项目: 上海航天先进技术联合研究基金(USCAST2020-5)

作者简介: 史梦阳(1996—), 男, 博士研究生, 主要从事雷达信号处理技术研究。

通信作者: 高叶盛(1983—), 男, 博士, 副教授, 主要从事雷达信号处理技术研究。

0 引言

光学遥感图像处理中,目标检测一直是一个具有挑战性的课题。由于遥感的目标图像通常分辨率较低,并且伴随不同程度的降质,导致获得的目标特征不清晰,给目标检测带来难度。提升目标识别算法或处理硬件的性能都可以得到高质量的图像。但是硬件性能的提升通常需要很高的成本,因此针对图像恢复的超分辨算法有着更广阔的应用前景。图像超分辨算法通常分为多张图像超分辨算法和单张图像超分辨算法,由于单张图像超分辨算法更为方便和直接,因此应用更为广泛。

单张图像超分辨算法主要包括 3 种,分别为基于插值的超分辨算法、基于优化的超分辨算法和基于学习的超分辨算法。

基于插值的超分辨算法^[1-2]包括最近邻插值、双线性插值、三次插值以及各种基于插值的改进超分辨算法。基于插值的超分辨算法具有重建速度快的优点,但是重建效果有限,不能很好地恢复图像的细节信息。

基于优化的超分辨算法重点研究从高分辨率图像到低分辨率图像的退化过程。LI 等^[3]提出了一种基于通用隐马尔可夫树(hiding Markov tree, HMT)模型的遥感图像超分辨的最大后验概率方法。HMT 理论用于建立一个先验模型,以便从一系列模糊、下采样和噪声污染的低分辨率图像中重建图像。SCHULTZ 等^[4]提出了一种利用最大后验概率估计高分辨率图像的贝叶斯方法。

基于学习的超分辨算法主要包括基于稀疏表示的方法和基于神经网络的方法。YANG 等^[5]首先提出了一种基于稀疏表示的图像超分辨方法,该方法通过学习一个冗余字典,建立了高分辨率图像和低分辨率图像的映射。ZHENG 等^[6]首先将稀疏表示法应用于遥感图像超分辨任务。基于优化和稀疏表示的超分辨算法可以利用先验信息,但也存在计算量大的缺点。

近年来,神经网络被广泛地应用于图像的超分辨算法。DONG 等^[7]在 2015 年首次将卷积神经网络(convolutional neural network, CNN)引

入图像超分辨算法中,取得了良好的效果。图像多蒸馏网络^[8](information multi-distillation network, IMDN)是一个轻量级的图像超分辨网络,有较高的实用价值。LIEBEL 等^[9]提出了用于遥感图像的卷积神经网络。LEI 等^[10]提出了局部-全局组合网络(local-global combined network, LGCnet), LGCnet 充分提取了目标图像的局部信息与全局信息,因此可以得到很好的重建效果。HAUT 等^[11]通过引入注意力机制,以很少的计算代价提高了遥感图像的重建质量。神经网络具有良好的性能,然而由于其独特的结构,导致其缺乏可解释性,不能灵活地利用传统的先验知识。

最近,一种将传统优化算法与神经网络相结合的图像重建方法被提出。该方法将传统的优化过程展开为一个网络,称为深度展开网络,算法中的参数可以通过训练得到。深度先验图像去噪网络^[12](learning deep CNN denoiser prior for image restoration, IRCNN)和深度即插即用超分辨率网络^[13](deep plug-and-play super-resolution, DPSR)都是深度展开网络,都可以很好地利用图像的先验信息,并且取得了较好的图像重建效果。深度残差去噪网络^[14](beyond a Gaussian denoiser: residual learning of deep CNN for image denoising, DNCNN)也是一个有效的去噪神经网络。ZHANG 等^[15]提出了超分辨率图像深度展开网络(deep unfolding network for image super-resolution, USRNet)用于处理多重退化的图像超分辨问题,并获得了良好的结果。USRNet 可以充分利用先验知识,但是并没有对网络进行良好的优化,导致网络的整体性能提升有限。

对于遥感图像处理算法,可解释性至关重要。并且遥感图像的数量有限,因此如何充分利用先验信息减小网络拟合的压力也是需要研究的。为了提高算法的可解释性和图像重建性能,本文提出一种基于多特征提取的深度展开图像超分辨网络。首先利用半二次分裂算法对目标函数进行迭代求解,接着将其展开为一个神经网络。为使网络可以充分提取目标图像的信息,提出一种多尺度空洞卷积注意力模块。该模块通过设计不同空洞率的空洞卷积组,扩大网络的感受野,使得网络

可以充分挖掘图像的信息。在此基础上,引入通道注意力机制和空间注意力机制,使得网络可以更加注意重要的特征图和重要的空间位置信息。最终,在网络具有良好的可解释性的同时,得到较好的图像超分辨重建效果。

1 方法与理论

1.1 优化方法

在遥感图像采集过程中,由于设备和环境的限制,图像通常受多种退化因素影响。本文考虑了图像模糊、随机噪声和图像降采样等几种典型的退化因素,使超分辨模型更加实用。退化模型可以表示为

$$y = (k \otimes x) \downarrow_s + n \quad (1)$$

式中: y 表示低分辨率图像; k 表示模糊核; \otimes 为卷积运算符; x 表示高分辨率图像; \downarrow 表示降采样; S 表示降采样因子; n 表示噪声。基于此退化模型,目标函数可以表示为

$$\operatorname{argmin}_x \| y - (k \otimes x) \downarrow_s \|^2 + \lambda \phi(x) \quad (2)$$

式中: $\operatorname{argmin}(\cdot)$ 表示使目标函数为最小值时的自变量取值函数; $\| \cdot \|$ 表示矩阵的 2-范数运算; λ 表示权重系数; $\phi(x)$ 表示先验约束项。

在传统优化算法中,先验约束通常可以选择稀疏先验或者低秩先验。但是很多情况下,算法性能与先验约束的选择息息相关,如果先验约束选择不合理,会影响算法的性能。本文将先验约束设置为一个隐式先验,此隐式先验表达可以通过端到端的网络学习得到。这种方式可以降低先验约束选择的难度,并能大大提高算法性能。为了求解式(2),引入辅助变量 z ,式(2)可以被重新表示为

$$\begin{aligned} \operatorname{argmin} \| y - (k \otimes x) \downarrow_s \|^2 + \lambda \phi(z), \\ \text{s. t. } x = z \end{aligned} \quad (3)$$

拉格朗日函数 $L_{\lambda,\beta}$ 可以用于处理有约束条件的问题。其表达式为

$$\begin{aligned} L_{\lambda,\beta} = \operatorname{argmin}_x \| y - (k \otimes x) \|^2 + \\ \lambda \phi(z) + \beta \| z - x \|^2 \end{aligned} \quad (4)$$

式中: β 表示惩罚系数。可以使用半二次分裂法^[16]来对式(4)进行迭代求解。式(4)可以被分解为两个子问题,迭代公式分别表示为

$$\begin{aligned} x_{i+1} = \operatorname{argmin}_x \| y - (k \otimes x) \|^2 + \\ \beta \| z_i - x_i \|^2 \end{aligned} \quad (5)$$

$$z_{i+1} = \operatorname{argmin}_z \beta \| z_i - x_{i+1} \|^2 + \lambda \phi(z_i) \quad (6)$$

式中: x_i 表示第 i 次迭代的高分辨率图像; z_i 表示第 i 次迭代的辅助变量。

对式(5)进行闭式解求解,闭式解通过快速傅里叶变换(FFT)在频域求得,可以表示为

$$x_{i+1} = F^{-1} \left(\frac{\overline{F(k)} F(y) + \beta F(z_{i-1})}{F(k) F(k) + \beta} \right) \quad (7)$$

式中: $F^{-1}(\cdot)$ 表示逆傅里叶变换; $F(\cdot)$ 表示傅里叶变换; $\overline{F(\cdot)}$ 表示傅里叶变换后取共轭。由于式(7)中包含隐式先验项,因此使用神经网络来求解。

1.2 展开网络

本文提出一种基于多尺度特征提取的图像超分辨深度展开网络,网络的整体结构如图 1 所示。

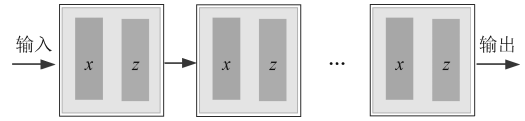
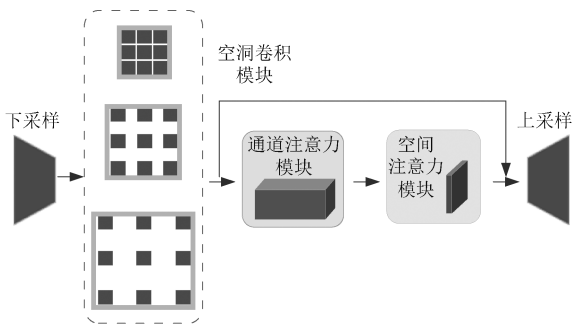


图 1 深度展开网络的整体结构

该网络包括 x 模块和 z 模块,每一模块都有明确的定义,并且分别执行不同的功能,因此网络是可解释的。 x 模块对应式(5)子问题的求解,而式(7)是式(5)的闭式解,主要完成去卷积操作。 x 模块可以充分利用模糊核的先验知识,可以完成图像的尺度变换和去模糊的任务。 z 模块对应式(6)子问题的求解,主要完成去噪的任务。噪声包含两部分,一部分是图像获取时引入的噪声,另一部分是 x 模块去卷积时引入的噪声。

Unet 是一种经典的网络结构,可以有效地提取目标图像的多尺度特征。 z 模块在原始 Unet 网络的基础上,在最小尺度特征图转换时,引入了多尺度空洞卷积模块、通道及空间注意力模块,完成对特征图的多尺度信息的提取。 z 模块的主体结构如图 2 所示。

首先设计多个尺度的空洞卷积模块组,对特征图进行信息的提取。不同尺度的空洞卷积采取不同的空洞率,可以产生不同感受野的卷积。较大的空洞率卷积可以更加关注图像的非局部信息,较小的空洞率卷积可以更加关注图像的局部

图 2 z 模块的主体结构

信息。因此,不同空洞率的空洞卷积组可以更好地提取图像的局部信息和全局信息。

然后对空洞卷积组提取到的特征图进行注意力权重的分配。先进行通道注意力的生成,通过最大池化操作将每个通道的特征图映射为一个权重,给每个通道施加不同的权重,使得网络更加关注重要的特征通道。再使特征图通过空间注意力模块,通过最大池化操作将空间特征映射为不同的权重,使得网络可以自适应地学习特征图的空间信息。对当前重建像素比较重要的区域施加较大的权重,对不太重要的区域施加较小的权重,从而使得网络具有更好的重建能力。

在注意力模块的前后采取跳连接方式相连,将得到的特征图进行多次上采样,得到与原始图像大小一致的特征图。

综上,该网络迭代执行 x 模块和 z 模块,从而交替进行去模糊和去噪的操作。迭代流程与传统优化算法相同,每一步骤都有明确的定义并且

完成相应的功能。最终实现图像的超分辨重建,并使用 L1 准则来计算最终的损失函数。

2 实验验证

通过实验来验证所提算法的有效性。UC Merced Land Use 数据集^[17]是一个常用的公共遥感图像数据集,包含 21 种场景,每种场景有 100 张图片,图片大小为 256×256 像素。从中挑选了 10 种场景作为算法评估的数据集。每种场景的前 40 个样本用来训练,再随机选择 5 个样本用来测试,最终得到 400 个训练样本和 50 个测试样本。选择文献[10]中的 6 种模糊核来验证算法的有效性,包括 3 个各向同性高斯模糊核和 3 个各向异性高斯模糊核,模糊核为 21×21 的矩阵。

实验在 PyTorch 框架下进行,使用 Titan RTX 进行训练。设置优化器为 Adam、学习率为 0.0005、网络迭代次数为 3。本文中的空洞率设置为 2,4,6。由于在去卷积的过程中已经引入了噪声,为了实验的简便性,未再加入额外噪声。

用峰值信噪比(PSNR)作为评价指标,不同算法的图像重建结果如表 1 所示。可以看到,采用不同的模糊核,算法都表现出了较好的性能。本文所提算法与 DNCNN、IMDN、IRCNN、DPSR 等算法相比,重建图像的峰值信噪比分别高 1.15,1.50,1.98,0.29 dB。由此可以证明本文所提算法的有效性。

表 1 不同重建算法的图像峰值信噪比

算法	峰值信噪比/dB					
	模糊核1	模糊核2	模糊核3	模糊核4	模糊核5	模糊核6
DNCNN ^[14]	31.46	31.38	30.44	30.12	29.74	29.93
IMDN ^[8]	31.20	31.04	30.17	29.69	29.31	29.59
IRCNN ^[12]	30.25	30.30	29.70	29.38	29.06	29.37
DPSR ^[13]	33.46	32.51	31.07	30.50	30.12	30.55
本文算法	33.64	32.75	31.29	30.86	30.51	30.93

图 3 是不同算法的超分辨重建结果,场景为飞机场。选择模糊核 6,2 倍下采样。其中 LR 是由低分辨率图像进行最近邻插值得到的图像,OURS 是采用本文所提算法重建的图像。可知,

经过降质的图像已经变得不清楚了,无法分辨飞机的细节结构信息,对后续目标检测非常不利。经过超分辨复原后,所有超分辨算法都可以使降质图像变得清晰。其中 IRCNN 的图像重建效果

优于其他几种对比算法。而本文所提算法恢复的图像最清晰,最有利于进一步的目标检测。其他算法重建的图像能大致分辨物体的结构,但是对于较为精细的结构不能很好地分辨。

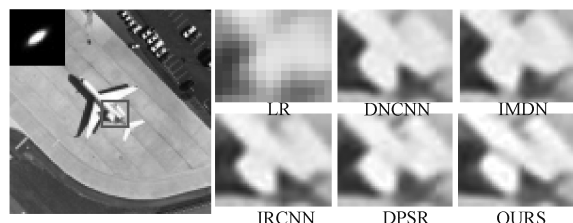


图 3 不同算法的超分辨重建结果

综上所述,本文所提算法无论是重建图像的视觉效果还是峰值信噪比都是最好的,证明了算法的有效性。

3 结论

本文提出了一种基于多特征提取的图像超分辨深度展开网络,网络的每一部分都有明确的定义和特定的功能,因此是可解释的。该网络通过半二次分裂法求解优化问题,并将求解过程展开成一个网络,使得网络同时具有基于优化方法的可解释性和神经网络强大的拟合性。为了使网络可以充分提取目标图像的信息,进一步提出了一种多尺度空洞卷积注意力模块。该模块通过设计不同空洞率的空洞卷积组,使得网络可以充分挖掘图像的全局信息和局部信息,并且关注图像中更重要的特征。实验证明所提算法可以对低分辨率图像进行有效的超分辨重建。

参考文献

- [1] LI X, ORCHARD M T. Newedge-directed interpolation[J]. *IEEE Transactions on Image Processing*, 2001, 10(10): 1521-1527.
- [2] BLU T, THEVENAZ P, UNSER M. Linear interpolation revitalized[J]. *IEEE Transactions on Image Processing*, 2004, 13(5): 710-719.
- [3] LI F, JIA X P, FRASER D, et al. Super resolution for remote sensing images based on a universal hidden Markov tree model[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, 48(3): 1270-1278.
- [4] SCHULTZ R R, STEVENSON R L. A Bayesian

approach to image expansion for improved definition[J]. *IEEE Transactions on Image Processing*, 1994, 3(3): 233-242.

- [5] YANG J C, WRIGHT J, HUANG T, et al. Image super-resolution as sparse representation of raw image patches[C]// 26th IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2008: 4587647.
- [6] ZHENG Z H, WANG B, SUN K. Single remote sensing image super-resolution and denoising via sparse representation[C]// 2011 International Workshop on Multi-platform/Multi-sensor Remote Sensing and Mapping. Piscataway, NJ: IEEE Press, 2011: 5697420.
- [7] DONG C, LOY C C, HE K M, et al. Image super-resolution using deep convolutional networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295-307.
- [8] ZHENG H, GAO X B, YANG Y C, et al. Lightweight image super-resolution with information multi-distillation network[C]// Proceedings of the 27th ACM International Conference on Multimedia. NY: ACM, 2019: 2024-2032.
- [9] LIEBEL L, KORNOR M. Single-image super resolution for multispectral remote sensing data using convolutional neural networks[J]. *The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*, 2016, XLI-B3: 883-890.
- [10] LEI S, SHI Z W, ZOU Z X. Super-resolution for remote sensing images via local-global combined network[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(8): 1243-1247.
- [11] HAUT J M, FERNANDEZ-BELTRAN R, PAOLETTI M E, et al. Remote sensing image super-resolution using deep residual channel attention[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(11): 9277-9289.
- [12] ZHANG K, ZUO W M, GU S H, et al. Learning deep CNN denoiser prior for image restoration [C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2019: 1671-1681.