

doi:10.3788/gzxb20184704.0410004

基于小递归卷积神经网络的图像超分辨算法

马昊宇,徐之海,冯华君,李奇,陈跃庭

(浙江大学 光电工程与信息学院 现代光学仪器国家重点实验室, 杭州 310027)

摘 要:针对现有软件实现超分辨算法通常过于复杂、运算开销大、模型复杂度高的问题,本文从成像过程中图像退化的物理原理出发,提出一套基于小递归卷积神经网络的单帧图像超分辨模型.将物理模型的约束融入到模型中,与现有的基于统计学习的图像超分辨算法相比,本文提出的模型的模型复杂度和计算量几乎可以忽略不计,同时内部的参数也有着更加明确的物理意义,并且引入了外部数据辅助对相应的模型参数进行学习.使用运行速度、峰值信噪比的数值方法对结果进行评价,结果表明:本文提出的算法消耗时间只有传统反向投影算法的 75%,而精度比反向投影算法提高了 0.2 dB,比双线性插值提高了 1.2 dB.本文提出的算法可以取得比迭代反投影算法更快、重建精度更高的超分辨重建效果.

关键词:超分辨成像;卷积神经网络;递归神经网络;图像恢复;底层视觉

中图分类号:TP751.1

文献标识码:A

文章编号:1004-4213(2018)04-0410004-9

Image Super-resolution Based on Tiny Recurrent Convolutional Neural Network

MA Hao-yu, XU Zhi-hai, FENG Hua-jun, LI Qi, CHEN Yue-ting

(State Key Laboratory of Modern Optical Instrumentation, College of Optical Engineering,
Zhejiang University, Hangzhou 310027, China)

Abstract: A super-resolution algorithm via tiny recurrent convolutional neural network was proposed on the basis of the principles of image degradation. The proposed model has very few parameters when compared to super-resolution algorithms based on naive statistical learning. Model parameters of the proposed model have their specific physical meanings because corresponding image degradation model is introduced and regularizes the proposed model implicitly. This paper also provides an inner view of the related parameters of the algorithm and how these parameters influence the performance of the algorithm. As a result, the proposed model can achieve better performance in terms of running speed and peak signal noise ratio, comparing to current iterative backprojection algorithm. The result illustrates that the proposed algorithm only takes about 75% time consumption, but improves the peak signal noise ratio by 0.2 dB comparing to conventional backprojection algorithm and 1.2 dB improvement comparing to bilinear interpolation respectively.

Key words: Super-resolution; Convolutional neural network; Recurrent neural network; Image restoration; Low-level vision

OCIS Codes: 100.2980; 100.3010; 100.3190

0 引言

随着信息技术的发展和智能设备的推广,用户对于图像质量的要求也越来越高.然而由于光学成像系统的像差、图像采集设备本身的带宽限制以及传输过程中的带宽限制,图像的像质不可能无限制的提高.因此如何采用合适的超分辨算法,提升图像质量,克服以上提到的种种问题,成为一个热门的研究课题.硬件实现

基金项目:国家自然科学基金(No.61475135)和浙江省科技计划项目(No.2017C01033)资助

第一作者:马昊宇(1993-),男,硕士研究生,主要研究方向为数字图像处理. Email: 21530059@zju.edu.cn

导师(通讯作者):徐之海(1964-),男,教授,博士,主要研究方向为数字图像处理. Email: xuzh@zju.edu.cn

收稿日期:2017-11-15; **录用日期:**2017-12-19

<http://www.photon.ac.cn>

超分辨常常精度较高性能较好^[1-2],但是常常应用较为局限,且成本很高,也无法应用于其它的成像环境.而现有的软件实现超分辨方案主要有两种实现方法:第一种是基于多帧的超分辨重建,从图像退化模型出发,对多张低分辨图像进行亚像素级配准,并利用配准信息重建出原始的高分辨图像^[3].然而,该方法要求多张低分辨图像之间的曝光等因素完全相同,并且对配准的精度要求极高,限制了该方法的应用.所以现在大量的研究都是针对单帧图像超分辨算法展开的研究,本文亦是主要研究单帧图像的超分辨技术.几乎所有现有的单帧超分辨算法都依赖于离线数据库,并且利用特定的算子提取图像中的信息作为模型的输入特征^[4-9]. Yang 等用稀疏表达的方法^[4],提取低分辨和高分辨训练样本对中的高频信息并且将其编为字典,之后用同样的方法提取测试图片中的高频信息,重建出对应的高分辨图像. Zeyde 等在此基础上对高频信息进行主成分分析(Principal Component Analysis, PCA)降维后再进行稀疏编码^[5],提升了算法的训练速度,降低了噪声对算法的影响.然而利用字典的方法对自然图像所在的流形空间进行描述依然是不够精确的,所以在 Timofte 等^[6-7]的方法中,作者选取每一个字典中的原子作为“锚点”,对周围的区域进行 L2 约束的线性回归,取得了比字典更好的效果. Schuler 等^[8]使用随机森林对该流形进行刻画,作为一种非线性回归器也可以取得不错的效果. Romano 等^[9]提出的 RAISR 模型将这一编码过程进行简化,同时加入许多工程技巧,在略微牺牲了重建精度的前提下大大提升了重建速度.另一些方法基于深度神经网络的模型可以达到最好的重建质量^[10-11],但是由于模型参数较多,其训练极度依赖于良好的硬件平台和样本集.因此并不适用于手机等一些低端成像设备.需要注意的是, Yang 等提出的迭代反投影算法^[4],被广泛应用于上述超分辨重建算法的后处理过程并普遍可以提升重建效果^[4,9,12],而且不需要训练,可以说本身就是一种简单的超分辨算法策略. Ledig 提出的基于对抗学习的算法可以取得更加接近于自然图片的超分辨图像^[13],但是由于存在大量的编造高频信息,所以峰值信噪比(Peak Signal to Noise Ratio, PSNR)较低,更类似于图像生成问题而非超分辨问题.

本文详细介绍了图像退化模型和基于此模型设计的超分辨算法,并与已有的算法进行比对,详细解释了本文采用的实验设计,在本文提出的小递归卷积神经网络框架下,对最终实现效果进行数值分析和主观效果的评价,得出结论是本文的方法可以获得更快和更高峰值信噪比的重建结果.

1 算法原理

本文从成像过程中的图像退化模型出发,推导单帧图像超分辨成像的原理,以及对应的数学模型,并且根据这套原理提出相应的算法实现,进而引出本文提出的递归卷积神经网络.单从本文网络的结构出发,本文提出的网络和已有的残差网络和反向投影算法都有一定的相似性,但是原理、实现都不尽相同,因此还对本文网络与这两种算法的联系与区别进行更深层的解读.

1.1 基于迭代的单帧图像超分辨模型

为了更好地解决单帧图像超分辨的问题,直接将成像过程中的图像退化模型作为提出模型中的约束,本文首先讨论成像过程中的图像退化模型,在获取图像的过程中,由于探测器噪声、较低的采样频率、光学系统的像差等因素的存在,必然导致图像退化,假设图像退化算子为 D ,未知的原始高分辨图像为 x ,在经过成像系统后,得到的低分辨图像为 b 可以表示为

$$b = Dx \quad (1)$$

式中,假设原始的高分辨图像,其采样倍率为低分辨图像的 k 倍,列向量化后的低分辨图像 $b \in \mathbb{R}^{M \times 1}$,那么可得 $x \in \mathbb{R}^{kM \times 1}$, $D \in \mathbb{R}^{M \times kM}$.在成像的过程中,真实的高分辨图像 x 已经无法获得,只能得到退化后的低分辨图像 b ,与此同时,往往也无法准确估计下采样算子 D ,因此只能利用近似的下采样算子 \hat{D} 来大致估算出退化前的图像 x ,将推算出的结果记为 \hat{x} ,那么需要保证估算的高分辨图像 \hat{x} 在经过对应的图像退化模型后,可以得到和获得的低分辨图像 b 一致的结果,这个求解高分辨图像的过程可以表示为

$$\hat{x} = \arg \min_x \|\hat{D}\hat{x} - b\|_2^2 \quad (2)$$

但是由于下采样算子 $\hat{D} \in \mathbb{R}^{M \times kM}$ 将一个高维空间的向量映射到低维空间上,本身是一个一对多的问题,因此 $\hat{D}^T \hat{D} \in \mathbb{R}^{kM \times kM}$ 本身的秩只有 M ,不可逆,也无法得到 \hat{x} 的解析解.所以只能通过迭代计算近似计算出 \hat{x}

的系数.根据梯度下降法,求解可以表示为

$$\hat{x}_{j+1} = \hat{x}_j - \lambda \nabla F(\hat{x}_j) \quad (3)$$

式中 \hat{x}_j 为第 j 步得到的近似高分辨图像, \hat{x}_{j+1} 为第 $j+1$ 步迭代得到的近似高分辨图像, λ 为迭代步长, $F(\hat{x}_j)$ 为重建的高分辨率图像 \hat{x}_j 经过图像退化后获得的低分辨率图像与实际获得的低分辨率图像之间的残差.该式可以进一步化简得到

$$\hat{x}_{j+1} = \hat{x}_j - \lambda \hat{D}^T (\hat{D} \hat{x}_j - b) \quad (4)$$

式(4)即为实现单帧图像超分辨率的迭代模型.其中 \hat{x}_j 的变化 $\Delta \hat{x} = -\lambda \hat{D}^T (\hat{D} \hat{x}_j - b)$ 由两部分组成:第一部分是 $a = (\hat{D} \hat{x}_j - b)$,即重建的高分辨率图像 \hat{x}_j 经过图像退化后获得的低分辨率图像与实际获得的低分辨率图像的残差;第二部分为 $\hat{D}^T a$,即将该残差由对应的上采样算子上采样所得到的残差图像.整个迭代计算实现超分辨率的过程可为

算法 1:	基于迭代的超分辨率算法
第 1 步:	初始化高分辨图像估计 \hat{x}_0 ,令 $\hat{x}_j = \hat{x}_0$
第 2 步:	计算经过退化模型后的残差 $a = \hat{D} \hat{x}_j - b$
第 3 步:	计算上采样后的残差 $\Delta \hat{x} = -\lambda \hat{D}^T a$
第 4 步:	更新高分辨图像估计 $\hat{x}_{j+1} = \hat{x}_j + \Delta \hat{x}$
第 5 步:	如果达到迭代终止条件,退出流程,否则返回第 2 步

1.2 基于卷积神经网络的递归超分辨率模型

在实际计算过程中,算法 1 的实现存在两个困难:首先,由于 \hat{D} 未知, $\hat{D}^T \hat{D}$ 不可逆,这意味着只能对 \hat{D} 和 \hat{D}^T 进行估算和近似表示;其次,对于一张尺寸为 $M \times N$ 的低分辨率图片来说,列向量化后的 \hat{x}_j 维度将为 $MN \times 1$,如果上采样倍率为 k ,对应的 \hat{D} 和 \hat{D}^T 维度为 $MN \times k^2 MN$ 和 $k^2 MN \times MN$,在图像维度较大时,这么大的矩阵几乎是无法进行计算和优化的.

为了解决这两个问题,本文提出的算法中针对退化模型 \hat{D} 做出假设:即在图像退化后,得到的低分辨率图像特定位置像素的像素值只受到与其对应的高分辨率图像对应位置像素以及周围较小范围内几个像素的影响,并且该影响可以表示为周围像素值以固定的权重进行加权.基于这个假设,本文用一个简单的下采样算法来近似表示退化过程 \hat{D} .同时也可以得出结论:重建高分辨图像某个特定位置像素的信息,也只保留在若干个对应的临近像素中,所以可以用一个简单的插值+卷积的组合,实现对 \hat{D}^T 的近似.通过实验,本文实验证明了做出的假设是合理的.通过以上讨论,本文提出的算法解决问题的关键就是计算卷积中的每个参数,从而完成对 \hat{D}^T 的近似.

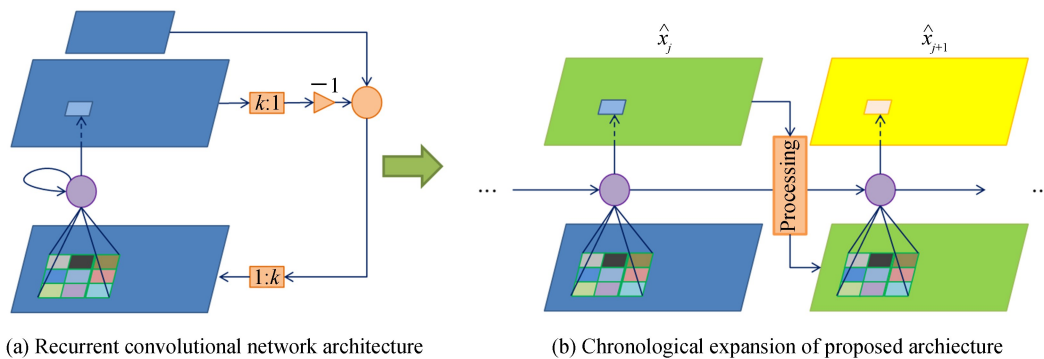


图 1 基于递归卷积神经网络的超分辨率算法流程图

Fig.1 The pipeline of our proposed super-resolution algorithm based on recurrent convolutional neural network

基于以上假设和先前提出的算法 1, 本文设计了以下线性激活函数的递归卷积神经网络模型^[14]来实现超分辨成像. 为了方便, 令卷积核的大小为 $(2S+1) \times (2S+1)$, 那么在 j 时刻, 如果需要对复原图像坐标为 (x, y) 的位置的像素需要重建, 那么选取以该位置为中心、大小为 $(2S+1) \times (2S+1)$ 的矩形区域作为输入, 与卷积核做内积后加上原先该位置内存中的数据, 即得到该点的新重建像素值. 完成整张图像所有位置像素值的重建以后, 这些像素值被存储在内存中, 即完成了一轮迭代. 由于采用线性激活函数, 所以输出值即为内存中的数据. 输出图像经过对应的退化模型、计算与实际观测低分辨图像的残差和上采样后, 被作为下一轮迭代的输入. 在 j 时刻, 令神经网络在图像坐标 (x, y) 位置的输入为 $i_j(x, y)$, 神经元内存为 $c_j(x, y)$, 输出为 $o_j(x, y)$. 那么该网络的数学定义为

$$i_j(x, y) = \begin{cases} [X_0(x+m, y+n)] \Big|_{m=-S}^S \Big|_{n=-S}^S & \text{if } j=0 \\ f[o_{j-1}(m, n)] & \text{if } j \neq 0 \end{cases} \quad (5)$$

$$c_j(x, y) = \begin{cases} \sum_{m=-S}^S \sum_{n=-S}^S W(m, n) i_j(m, n; x, y) & \text{if } j=0 \\ \sum_{m=-S}^S \sum_{n=-S}^S W(m, n) i_j(m, n; x, y) + c_{j-1}(x, y) & \text{if } j \neq 0 \end{cases} \quad (6)$$

$$o_j(x, y) = c_j(x, y) \quad (7)$$

假设待重建高分辨图像大小为 $M \times N$, 则 $X_0 \in \mathbb{R}^{M \times N}$ 表示初始高分辨图像估计, $i_j(m, n; x, y)$ 表示输入矩阵 $i_j(x, y) \in \mathbb{R}^{(2S+1) \times (2S+1)}$ 在 (m, n) 的数值. $f(x)$ 为图 1 所示的固定操作. 该循环神经网络 (Recurrent Neural Networks, RNN) 定义的流程完全对应了算法 1, 在完成一定次数迭代后, 模型将输出最终的重建结果, 如果是在训练阶段, 该结果与实际高分辨图像之间的均方差将被用于计算开销函数和反向传递更新网络参数.

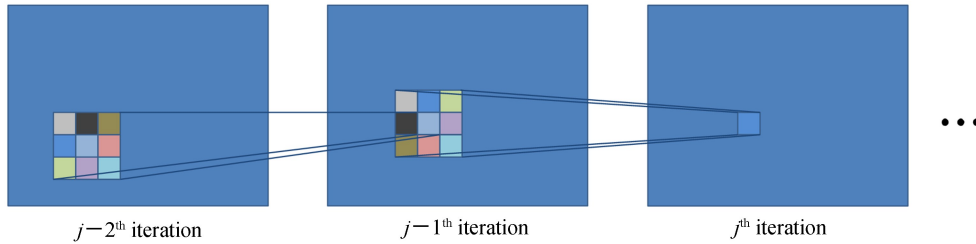


图 2 随着迭代次数的增加, 实际重建某个像素时的感知域不断加大

Fig.2 The actual perceptual field used to reconstruct a pixel in a certain location increases when more iterations are used

本文设计的递归卷积神经网络结构带来了两个好处: 首先, 虽然采用了固定大小的卷积核, 每次迭代只使用了周围有限的像素信息, 但是随着迭代次数的增加, 重建某个像素需要的感知域是在不断增加的, 例如对于 $(2S+1) \times (2S+1)$ 的卷积核大小和 K 次迭代, 总共用到的像素数量为 $(2KS+1) \times (2KS+1)$, 不断增大的感知域意味着网络可以更多地考虑全局信息, 得到更精确的重建结果. 另一方面, 由于从实际物理模型出发设计的算法只使用了一个神经元, 所以需要确定的参数只有该神经元的卷积核, 这样的参数量是非常小的, 允许用很少的样本以及少量的计算得到较好的效果, 相比较传统深度学习算法^[8,11], 需要的计算开销和训练样本数量非常少, 用十几张图和 CPU 就可以完成.

1.3 与残差网络的联系

本文提出的递归卷积神经网络在时域上的展开类似于传统的残差网络^[15]的构造 (如图 3 所示): 传统的残差网络可以理解为不同层的卷积核之间引入了残差连接, 而本文提出的模型的这一残差链接则是通过 RNN 神经元更新内存时叠加前一时刻的内存, 实现了不同时刻的残差连接, 这样带来了类似的好处, 即本文提出的网络也具有较好的收敛性.

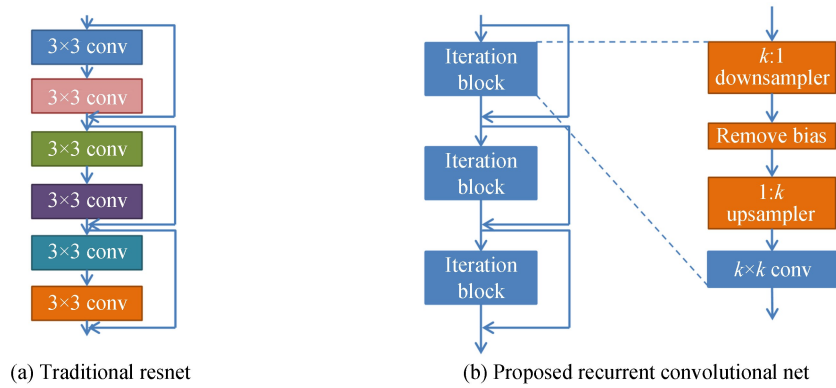


图3 传统残差网络和本文提出的递归卷积神经网络的比较

Fig.3 Comparison between traditional ResNet and our proposed recurrent convolutional neural network

但是本文提出的算法和传统的残差网络存在两个很大的不同:首先,本文提出的网络结构是从物理模型出发,推导出的超分辨成像迭代模型,而利用残差网络算法的超分辨算法则多是将超分辨成像的物理过程本身看作是一个黑盒子,利用深层残差网络优秀的泛化能力对这一黑盒子进行近似而不对物理过程建模;其次,传统的残差网络采用多个卷积层,每个卷积层都是待定的(图3(a)),所以引入了大量的未知参数,为了避免出现过拟合,对样本数量和计算机性能的要求都非常高,而本文提出的网络基于物理模型,则是在时域上实现了一个残差网络,通过参数共享,只有极少的参数需要优化,极大地减小了对样本数量和计算机性能的要求。

1.4 与反向投影算法的联系

从原理上来看,本文提出的算法和文献[2]中使用的反向投影算法具有很高的相似性,都是从减小重构图像退化后与实际观测得到的低分辨率图像的残差出发,对重构图像进行优化,但是本文提出的算法选择了递归网络的结构,并且具有更好的泛用性,尽管为了方便,实验中使用了简单的插值下采样方法和插值+卷积相结合的方法去近似表示超分辨成像中的物理模型,实际上这些都可以根据实际情况中图像表现出来的特性进行修改,通过训练过程可以对待定参数进行优化,而在迭代反投影算法中,研究人员使用了插值+高斯模糊的方式去实现上采样,而高斯模糊核的相关参数完全由研究人员设定,算法的性能则非常依赖于研究人员的经验。

2 实验结果与分析

在实验中,为了方便比较,本实验采用了超分辨算法最常见 Set5 数据集作为测试集,Set14 作为训练集。主要在模型的复杂度和重建效果上,将本文提出的算法与现有算法进行比较,并且对批数量大小、卷积核大小等一系列可能会对模型性能产生较大影响的参数进行深层次的探究和分析。

2.1 模型复杂度

尽管 Set14 只有 14 张图片,几乎不可能被用于传统深度神经网络的训练^[13],甚至在很多传统的超分辨算法中这样的训练样本数量也太小了^[9],但是网络依然表现出了良好的收敛性,这主要得益于模型的小体积,对于高斯核大小为 $K \times K$ 的迭代反投影模型,总共有 K^2 个参数需要存储下来,而在本实验中采用了 Tucker 分解,所以只需要存储 $2K$ 个参数,考虑到 K 一般不会超过 20,这样的参数数量是大多数设备都可以接受的。在实验中,一个典型的模型通常使用 CPU 就可以利用十几分钟完成训练,需要注意的是,由于提出的算法中含有大量的卷积计算,这些运算都是可以并行的,所以只要简单的优化就可以达到更快的训练速度。

而对于论文中提出的稀疏字典模型^[4],假设图像块大小为 $K \times K$,字典原子数量为 N ,上采样倍率为 k ,那么字典将包含大约 $(k+1)NK^2$ 个参数,而为了保证字典重建的质量,通常原子数超过 1 000 个,这样庞大的参数数量将会对设备提出很高的要求,通常训练需要若干个小时。而基于传统算法的模型^[7]为了追求更好的重建精度,进一步增加了模型的复杂度。基于深度学习的模型^[13],由于含有较多的卷积层,每个卷积层又有很多通道,所以需要存储的参数将是所有的卷积层参数之和,模型^[13]需要确定的参数在 1 万以上,使用一

块顶级 GPU 也需要几周时间才能完成训练.无论是传统超分辨算法还是深度学习算法,其模型复杂度都是远远高于本文提出的网络的.本实验中对一块大小为 168×168 大小的图像进行上采样,采样倍率为 3,与当前较优秀的超分辨算法^[5]进行对比,不同算法的时间开销如表 2,可见本文的算法经过优化,速度比文献[2, 5]中的算法更为优秀.

表 1 不同超分辨重建方法速度对比

Table 1 Reconstruction speed comparison between different super-resolution algorithms

Algorithms	BP[2]	A+[5]	Proposed
Elapsed time	0.08	0.87	0.06

2.2 方法评价指标

本文实验主要关注视觉指标和量化指标.视觉指标即研究者人为评价图像质量好坏,采用超分辨算法中常用的峰值信噪比(Peak Signal Noise Ratio, PSNR)作为量化指标.PSNR 其实是计算两张图像的相似性(均方差),相似性越高(均方差越低)则 PSNR 越高.将重建图像与真实高分辨图像进行比较,如果具有较高的 PSNR,则重建效果较好.

2.3 实验参数设置

本实验在一台联想 ideapad 700 上展开实验,使用 CPU 版本的 tensorflow 完成网络框架的搭建和测试.虽然本文提出的模型可以采用任意的上采样倍率,但是为了方便讨论,下列的实验都针对上采样倍率为 3 时不同参数的表现情况进行分析.通过实验发现,对模型性能影响较大有批数量大小,卷积核大小,迭代轮次三个参数.利用 Set14 对模型进行训练,用 Set5 对模型进行测试,并记录测试得到的 PSNR,同时将文献[4]中提出的迭代反投影算法作为基准与本文提出的算法性能进行对比,迭代反投影的高斯模糊核选用大小为 5×5 ,方差为 1 的高斯核,为了公平起见,将该模糊核也作为本文网络采用模糊核的初始估计,将双线性插值获得的结果作为高分辨图像的初始估计.接下来将逐一对这些参数进行分析.

2.3.1 批数量大小

在训练过程中,可以一次前馈一张图片,按照算法 1 更新内部参数,这是一种随机梯度下降的方法,也可以一次前馈多张图片,对其计算得到的梯度取均值,更新内部参数,这是一种批量随机梯度下降的方法.考虑到在训练超分辨网络过程中,超分辨区域是随机选择的,而图像的大多数区域都为低频信息,所以如果一次前馈一张图片,很有可能图像区域非常平坦导致计算得到的残差为零,容易为卷积核权重引入很大的噪声,所以开始时随机将前馈的一张图片裁剪为 N 个相同大小的区域并且输入网络,计算梯度均值并对网络进行优化,一方面优化了算法的收敛性,一方面引入随机性,较少过拟合.这个 N 的数量就被记为批数量大小,理论上批数量并非越大越好,过大的批数量将导致算法更容易在局部极值处收敛.

图 4 验证了先前的假设,尽管所有提出的算法性能都要优于原始算法,但是在批数量大时,网络表现得收敛速度更快,也更容易导致网络在局部极小值处收敛,而批数量较小时,训练网络时表现得收敛性质较差.不难从图 4 中发现,选取批数量为 8 是一个较好的选择.

2.3.2 卷积核大小

卷积核直接决定了每次迭代计算时,用于重建像素的感知域的大小,理论上卷积核越大则模型的表达能力越强,但是较大的卷积核将会导致模型在样本较少的情况下更容易出现过拟合,而且根据前文的讨论,随着迭代次数的增加,实际感知域也是不断增大的,所以过大的卷积核并没有必要.实验中发现卷积核对算法的性能影响更为明显(图 5).在卷积核较大时,模型表达能力更强,但是非常容易由于小样本出现过拟合,训练一开始可以迅速达到很高的重建精度,但是随着训练轮次的增多,在测试集上的准确率反而越来越低(卷

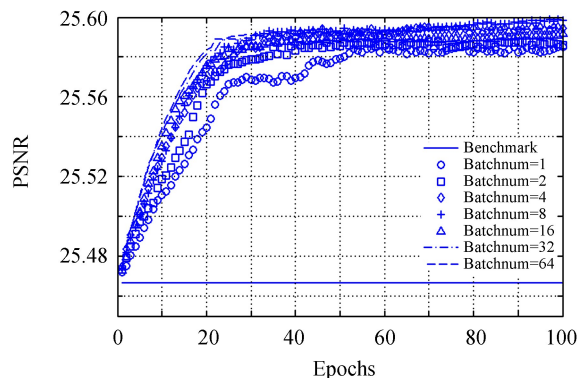


图 4 不同批数量大小下测试图像重建精度与训练轮次

Fig.4 Relationship between reconstruction accuracy and number of training epochs under different batch number settings

积核大小为 9 和 11 时).这个问题可以通过两种工程方法解决:一种是增大样本,一种是增加交叉检验集作为中止训练的指标.但是考虑尽量让模型的复杂度和训练样本数量相匹配,所以最终采用卷积核大小为 5.这也佐证了在大多数深度学习网络中卷积核大小都比较小^[13],以及迭代反投影中研究人员选用了大小为 5 的卷积核^[11].

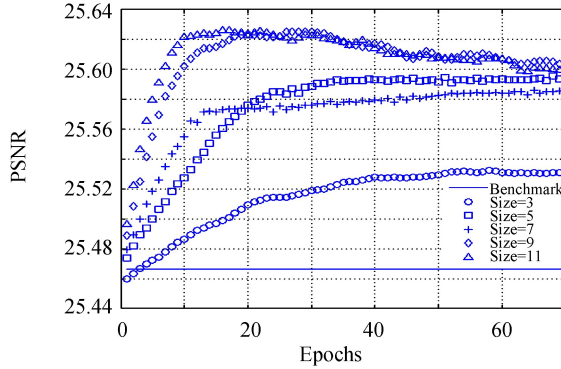


图 5 不同卷积核大小下测试图像重建精度与训练轮次图

Fig.5 Relationship between reconstruction accuracy and number of training epochs under different kernel sizes

2.3.3 迭代计算

由前面的讨论可知:迭代次数越多,重建单个像素时的感知域也越大,所以迭代次数越多,重建性理论上越好,这和传统的残差网络层数越多网络表达能力越强的特性是一致的.但是单纯地增加迭代次数也会导致计算时间的增加,因为迭代计算是无法通过并行计算加速的.在实验中发现迭代次数为 15 时,已经基本达到所能达到的最好效果,见图 6.

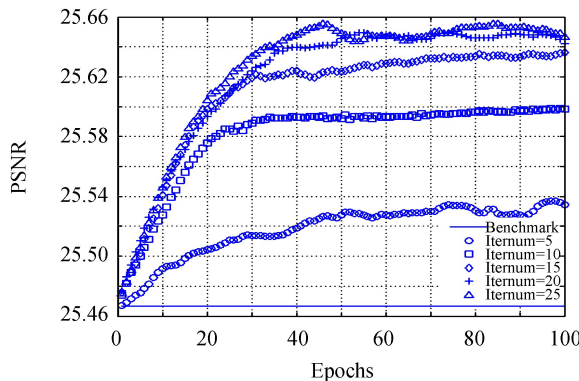


图 6 不同迭代次数下测试图像重建精度与训练轮次图






Fig.6 Relationship between reconstruction accuracy and number of training epochs under different iteration settings

2.4 整体效果

该部分主要比对算法实现后的最终效果,由于原理上和迭代反投影较为接近,但是实现上差别较大,所

表 2 不同重建方法精度对比

Table 2 Reconstruction accuracy comparison between different algorithms

						Mean
Bilinear	33.0 dB	31.4 dB	23.2 dB	32.4 dB	27.6 dB	29.5 dB
BP[2]	34.0 dB	32.6 dB	24.2 dB	33.1 dB	28.7 dB	30.5 dB
A+[5]	34.0 dB	34.1 dB	25.3 dB	32.5 dB	29.6 dB	31.3 dB
Proposed	34.2 dB	32.9 dB	24.4 dB	33.2 dB	29.0 dB	30.7 dB

以选择了迭代反投影的模型^[4]进行比较.同时,还选择了目前较为复杂,但是重建性能最好的模型之一^[7]进行比较,该模型^[7]利用了数百张图片 and 50 万个样本完成了训练,其模型复杂度和需要的训练样本都远远大于本文提出的算法,最后的重建效果确实比本文提出的算法更胜一筹.同时选择双线性插值得到的图像的 PSNR 作为基准,所有的超分辨算法都可以取得比普通插值方法更好的重建效果.

本文还将典型输出图片最终显示出来,进行人眼主观比较.本文主要比较和提出的方法原理相似的迭代反投影算法和双线性插值,结果如图 7.对比表明本文提出的算法比普通的插值效果更为优秀,普通的插值方法更容易将图像平滑化.并且本文提出的算法比文献^[4]中提出的迭代反投影算法具有更少的振铃,这是因为在迭代计算时没有采用固定的卷积核,而是由实际的退化模型计算得到的.而相比基于图像块的算法^[7],本文提出的算法还存在较多的振铃,这可能是为了减小样本的数量和模型的复杂度,训练样本较少,模型的表达能力不如大模型好导致的.



图 7 不同算法进行高分辨重建效果
Fig.7 Reconstruction results from different reconstruction algorithms

4 结论

针对现有超分辨算法通常过于复杂的问题,本文从成像过程中的图像退化模型出发,提出一套基于小递归卷积神经网络的单帧图像超分辨模型.由于利用了成像过程中的物理模型作为约束,本文提出的模型可以实现相比较于重建效果较好,但将图像退化作为黑箱模型去考虑的超分辨模型几乎可以忽略不计的模型复杂度和计算量,取得比迭代反投影算法更好的超分辨重建效果.并且对实验参数的分析和讨论,验证了其他研究人员用于相关模型的工程经验.本算法相比较于其它较大的超分辨算法模型,有极低的运算复杂度,但是在有大样本集,并且允许大量的计算开销时,高复杂度算法在结果上更胜一筹.

参考文献

- [1] YAO Bao-li, LEI Ming, XUE Bin, *et al.* Progress and applications of high-resolution and super-resolution optical imaging in space and biology[J]. *Acta Photonica Sinica*, 2011, **40**(11):1607-1618.
姚保利, 雷铭, 薛彬, 等. 高分辨和超分辨光学成像技术在空间和生物中的应用[J]. 光子学报, 2011, **40**(11):1607-1618.
- [2] LIU Hai-ying, LI Yun-song, WU Cheng-ke. A method for compressive sensing of images based on zone control of digital micromirror device and super-resolution[J]. *Acta Photonica Sinica*, 2014, **43**(5):0510002.
刘海英, 李云松, 吴成柯. 一种数字微镜阵列分区控制和超分辨重建的压缩感知成像法[J]. 光子学报, 2014, **43**(5):0510002.
- [3] SCHOLEFIELD A, DRAGOTTI P L. Accurate image registration using approximate Strang-Fix and an application in super-resolution[C]. *Signal Processing Conference, IEEE*, 2010:1063-1067.

- [4] YANG J, WRIGHT J, HUANG T S,*et al.* Image super-resolution via sparse representation[J]. *IEEE Transactions on Image Processing*, 2010, **19**(11):2861-2873.
- [5] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations[M]. *Curves and Surfaces*. Springer Berlin Heidelberg, 2010:711-730.
- [6] TIMOFTE R, DE V, GOOL L V. Anchored neighborhood regression for fast example-based super-resolution[C]. *IEEE International Conference on Computer Vision*, IEEE, 2014:1920-1927.
- [7] TIMOFTE R, SMET V D, GOOL L V. A+: Adjusted anchored neighborhood regression for fast super-resolution[C]. *Asian Conference on Computer Vision*. Springer, Cham, 2014:111-126.
- [8] SCHULTER S, LEISTNER C, BISCHOF H. Fast and accurate image upscaling with super-resolution forests[C]. *Computer Vision and Pattern Recognition*, IEEE, 2015:3791-3799.
- [9] ROMANO Y, ISIDORO J, MILANFAR P. RAISR: rapid and accurate image super resolution[J]. *IEEE Transactions on Computational Imaging*, 2017, **3**(1):110-125.
- [10] DONG C, CHEN C L, HE K,*et al.* Learning a deep convolutional network for image super-resolution[C]. *European Conference on Computer Vision*, Springer, Cham, 2014:184-199.
- [11] DONG C, CHEN C L, HE K,*et al.* Image super-resolution using deep convolutional networks[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2016, **38**(2):295-307.
- [12] TIMOFTE R, ROTHE R, GOOL L V. Seven ways to improve example-based single image super resolution[C]. *Computer Vision and Pattern Recognition*, IEEE, 2016:1865-1873.
- [13] LEDIG C, THEIS L, HUSZAR F,*et al.* Photo-realistic single image super-resolution using a generative adversarial network[J]. *Arxiv Preprint Arxiv*, 2016:1609.04802.
- [14] BASE L T. *neural networks for pattern recognition*[M]. Oxford University Press, 1995.
- [15] HE K, Z X, REN S, *et al.* Deep residual learning for image recognition[C]. *Computer Vision and Pattern Recognition*, IEEE, 2016:770-778.