

图像超分辨率全局残差递归网络

张雷¹ 胡博文¹ 张宁² 王茂森²

(沈阳航空航天大学电子信息工程学院 沈阳 110136)¹ (上海航天电子技术研究所 上海 201109)²

摘要 将深度网络模型应用在图像超分辨率上取得了很大的成功,并且已经证明了在将低分辨率图像重建成高分辨率图像的重建质量上深度网络模型普遍高于传统的算法。为了进一步提高图片的重建质量,文中提出了全局残差递归网络。通过优化经典的残差网络,提出全局残差块特征融合和局部残差块特征融合,让模型产生“自适应”更新权值的思想,改善信息流。结合 L1 代价函数,ADAM 优化器进一步提高了训练的稳定性,并通过 DIV2K 训练集来训练模型。通过 PSNR/SSIM 图像重建指标来评价图片重建质量,在 SSIM 指标中,所提模型最高可达 0.94,优于目前最新的深度学习模型(EDSR)的 0.92。全局残差递归网络模型有效地提高了图像的重建质量,减少了训练时间,避免了梯度衰减,提高了学习效率。

关键词 图像超分辨率, L1 代价函数, 全局残差递归网络, ADAM 优化器, DIV2K 训练集

中图分类号 TP394 文献标识码 A

Global Residual Recursive Network for Image Super-resolution

ZHANG Lei¹ HU Bo-wen¹ ZHANG Ning² WANG Mao-sen²

(College of Electronic Information Engineering, Shenyang Aerospace University, Shenyang 110136, China)¹

(Space Electronic Technology Research Institute in Shanghai, Shanghai 201109, China)²

Abstract The application of the deep network model has achieved great success in image super-resolution, and it has been proven that the reconstruction quality of low-resolution images reconstructed into high-resolution images is generally higher than traditional algorithms. In order to further improve the reconstruction quality of image, a global residual recursive network was proposed. By optimizing the classical residual network, the global residual block feature fusion and the local residual block feature fusion are proposed, which allows the model to generate the idea of adaptive updating weights, and it improves information flow. In combination with the L1 cost function, the ADAM optimizer further improves training stability and trains the model through the DIV2K training set. Through the PSNR/SSIM image reconstruction index, the quality of picture reconstruction is obtained. In the SSIM index, the maximum value is 0.94, which is superior to 0.92 of the current latest deep learning model(EDSR). The global residual recursive network model effectively improves the image reconstruction quality, reduces straining time, effectively avoids gradient attenuation, and improves learning efficiency.

Keywords Image super-resolution, L1 cost function, Global residual recursive network, ADAM optimizer, DIV2K training set

1 引言

图像超分辨率是指由一幅低分辨率图像或图像序列恢复出高分辨率图像。图像的分辨率对图像解译质量有着重要影响。与低分辨率图像相比,高分辨率图像具有更高的像素密度,能够提供更多的色调、形状和纹理信息。早期的主流方法使用最多的是基于采样理论的插值技术^[1,15,17-18],但是这些插值方法在预测纹理方面表现出局限性,会导致重建后的图像模糊不清,从而降低图片重建质量。SRCNN^[2-3]将深度学习与传统稀疏编码之间的关系作为依据,将 3 层网络划分为图像块提取、非线性映射以及图像重建层,所展现的重建能力都更具有优势。在图像预处理阶段, SRCNN 将原始低分辨率

图像通过双三次插值法处理,获得网络模型所输入图像的大小,因此在计算复杂度方面, SRCNN 与高清图像的空间大小成正比关系,对于放大系数 n ,插值后低分辨率图像的计算复杂度是原始低分辨率图像的 n^2 倍。在最近几年,计算更快、更深的网络模型被相继提出^[7,13]。

VDNR^[6]为了增加深度网络模型的层数,提出了自适应梯度裁剪,将梯度限制在某一范围,以减小训练网络模型的难度,避免梯度爆炸与衰减。通过建立有效的网络模型,使得深度学习网络越来越深,并且具有更好的重建性能。但是, VD-SR 将插值后得到的低分辨率图像作为网络的输出,导致参数多,计算复杂度高,占用内存多。虽然 SRResNet^[9]以良好的性能成功解决了训练时间和内存问题,但它延续了原始 Res-

本文受国家自然科学基金(61671037),上海航天科技创新基金(SAST2016090)资助。

张雷(1972-),男,博士,副教授,主要研究方向为图像压缩、图像识别等, E-mail: rd_zhangl@126.com; 胡博文(1992-),男,硕士生,主要研究方向为图像超分辨率、图像压缩等, E-mail: 767464760@qq.com(通信作者); 张宁(1982-),男,博士,高级工程师,主要研究方向为卫星数据智能处理、空间电子系统等; 王茂森(1985-),男,硕士,工程师,主要研究方向为卫星计算机等。

Net^[12]架构。ResNet架构的提出是为了解决高级计算机视觉问题,例如图像分类和检测。因此,将ResNet架构直接应用于超分辨率等低级视觉问题可能效果不太理想。DRCN^[11]提出使用递归神经网络,通过获得更多的卷积层来增加网络感受野,特征的非线性变换数据循环地通过该层多次,将这个递归网络进行展开,就等效于使用同一组参数在多个串联的卷积层进行学习训练。MemNet^[19]内存块网络可以存储信息,让信息在不同的记忆单元之间流动,具有较强的学习能力。DenseNet^[4]减轻了梯度消失问题,支持特征复用,且减少了参数数量。EDSR^[10]使用残差块思想建立了一个非常深的网络,优化了SRResNet多余的模块,在相同的计算资源下,可以堆叠更多的网络层或者使每层提取更多的特征,从而得到更好的性能。

但上述方法都忽略了来自原始图像的分层特性。从这一假定出发,本文创新性地提出了一种新的网络结构——全局残差递归网络,充分利用原始LR图像中的所有分层特征。利用残差块(RB)和扩展残差块(ERB)构建具有储存记忆功能的全局残差块(GRB)作为全局残差递归网络(GRRN)的构建模块。本文基于SRResNet模型进行改进,删除其不必要的模块,以简化网络架构;同时保留跳跃连接,用来进行全局残差学习(GRRL),防止网络层数增加而导致的梯度弥散问题与退化问题,使每一级的全局残差块都具备更新、保存状态和改善信息流的功能,类似得到一个“共享单元”与“自适应单元”。每一级的“共享单元”与“自适应单元”可以自适应地保留分层功能,融入新加入的信息,改善输入信息流,提高图像重建质量,降低计算复杂度,从而获得更深的网络层数。本文所提模型的核心思想主要包含以下4点。

(1)本文提出全局残差递归网络(GRRN),有效地融合了低分辨率图像空间中所有层的特征。

(2)增加Shrinking层。通过对输入图片提取特征进行压缩降维,降低计算复杂度;同时增加Expanding层,还原高分辨率图像损失的图片细节。

(3)本文提出的一种GRB,GRBN块,充分利用每一层网络中保存的信息,改善了每一级输入的信息流,实现了类似“自适应”的功能。本文创新性地提出了全局残差块特征融合、全局残差学习,优化了SRResNet架构;延续自适应梯度裁剪、特征复用等思想,提高了图片重建质量并减少了训练时间。

(4)在实际实验过程中发现EDSR的计算复杂度依然较

高,并且没有充分利用原始图像的特征提取,而且EDSR中的跳跃连接只关注了优化方面。在EDSR中,原文作者采用了在每一个残差块上都乘以固定数值0.1的方法,但是这个固定数值作者并没有给出一个明确的计算方法,存在不确定因素。本文要解决EDSR中的“不确定因素”,同时提高图片重建质量,降低计算复杂度。

2 全局残差递归网络

图1比较了SRResNet,EDSR和本文提出的GRRN。通过优化SRResNet结构,本文模型减少了计算参数,提高了学习效率。

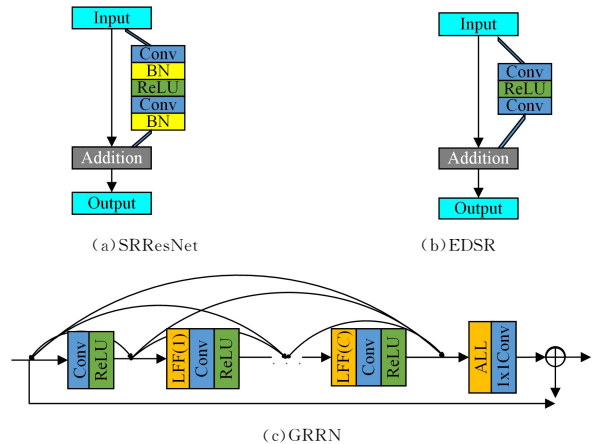


图1 SRResNet,EDSR和GRRN

2.1 图像特征提取单元

F_{00} 是第一层卷积特征提取的原始图像。

$$F_{00} = H_0(I_{LR}) \quad (1)$$

2.2 Shrinking层

本文在原始图像特征提取之后,增加了Shrinking层,用于减少原始图像的特征维度。实验证明,增加Shrinking层可以大幅度提升训练效率。删除Shrinking层,每次迭代的平均时间为7.56s;增加Shrinking层,每次迭代的平均时间仅为2.13s。

$$F_0 = H_{\text{Shrinking}}(F_{00}) \quad (2)$$

2.3 全局残差块

残差网络^[14]在计算机视觉问题中表现出优异的性能。图2给出GRBN的模型结构。

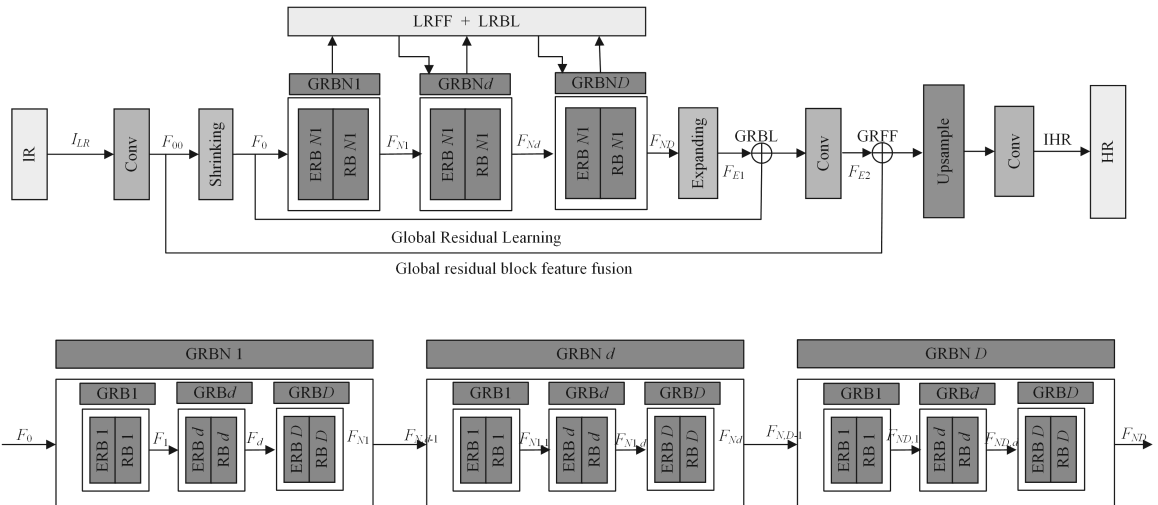


图2 GRRN的模型架构

在 GRB 块内,每一组 GRB 块由 ERB 与 RB 组成,用到了局部特征融合(LFF)的方法,用来自适应地融合 GRB 内的信息流向。本文最后用一个 1×1 卷积层来“自适应”地控制输出信息。在每级 GRBN 中的特征融合包括局部残差块特征融合(LRFF)和局部残差块学习(LRBL)。

LRFF 的主要功能为:保存每一级 GRBN 之间输出的信息,产生的“自适应”更新进入下一级 GRBN 的输入信息。

LRBL 的主要功能为:充分利用 LRFF 所更新的状态和信息,形成共享权值的功能。

GRBN1 为第一级 GRB 的总集合:

$$F_{N1} = H_{GRB,n}(H_{GRB,n-1} \cdots (H_{GRB,1}(F_0))) \quad (3)$$

GRBN d 为第 d 级 GRB 的总集合:

$$F_{Nd} = H_{GRB,n}(H_{GRB,n-1} \cdots (H_{GRB,1}(F_{N,d-1}))) \quad (4)$$

GRBND 为第 D 级的 GRB 总集合:

$$F_{ND} = H_{GRB,n}(H_{GRB,n-1} \cdots (H_{GRB,1}(F_{N,D-1}))) \quad (5)$$

其中, F_{Nd} 表示第 d 级 GRBN 的操作。假定通过 Shrinking 层的特征维数为 S ,则 D 级 GRBN 特征维数为: $F_{ND} = S + (D - 1) * S$ 。

2.4 Expanding 层

Expanding 层的作用类似于 Shrinking 层的逆过程。实验表明,删除 Expanding 层,图像在重建过程中会损失大量的图像细节,并且 L2 代价函数在客观评价指标上并不能保证更好的性能。在本文的实验中训练了深层网络,结果 L1 代价函数^[16]比 L2 代价函数的重建性能有所提高,而且 L1 代价函数相较于 L2 代价函数具有参数少的特点,减少了训练时间。

2.5 实验参数设置

本文提出的全局残差递归网络的实验参数设置如下:第一层卷积层的大小为 3×3 ;对于 Shrinking 层、Expanding 层,其大小为 1×1 的卷积核,改善了特征图像信息流;本文将零值填充到原始图像输入的每一侧,以保持图像大小不变,并保证图片边缘的细节处理;浅特征提取层 $F = 256$, $Shrinking = 128$, $Expanding = 256$;全局特征融合层中的第一层卷积核数量为 128,第二层卷积核数量为 256,则第 N 层卷积核数量为 $128 + (N - 1) * 128$; $M = 4$,每级 GRBN 由 4 组 GRB 构成,即 $GRBN = [GRB1, GRB2, GRB3, GRB4]$; $N = 5$,表示分为 5 级: $GRBNS = [GRBN1, GRBN2, GRBN3, GRBN4, GRBN5]$;因为本文输出的图像为 RGB 图像,所以最终的 Conv 层具有 3 个输出通道。

3 实验结果

最近, Timofte 发布了高质量(2K 分辨率)数据集 DIV2K 图像恢复应用。DIV2K^[15]是由 800 张高清图像组成的数据集,其中包含 100 张验证图像和 100 张测试图像。本文用 800 张图像集训练全局残差递归网络模型。图片重建质量用 PSNR 和 SSIM 进行评估。本文通过 ADAM^[8]优化器来提升网络模型的训练效率,从而提升网络的重建性能,其参数设置为 $\beta_1 = 0.9, \beta_2 = 0.999$ 。为了加速网络模型收敛速度,使其更快地找到全局最优解,本文采用小批梯度下降(mini-batch),大小设置为 16;学习率初始化为 1×10^{-5} 。本文采用 L1 代价函数来训练全局残差递归网络;使用 NVIDIA-GTX750Ti 对其进行训练,迭代 5000 次,验证了全局残差递归网络可以提

升图片重建质量并减少训练时间。在 RGB 图像中,EDSR 为最新的深度学习网络模型,但 EDSR 作者明确说明,使用 32 组残差块时必须要在每一组残差块后乘以固定数值 0.1 以避免梯度衰减,使其网络模型收敛。本文提出的 GRN 与 EDSR 一同进行 5000 次迭代,并比较在训练过程中记录的 PSNR 和 SSIM 值,得到客观评价重建质量指标。首先对比全局残差递归网络,表 1 列出不同的 GRN 网络结构对重建质量的影响。

表 1 不同 GRN 结构下的 PSNR 值

方法	PSNR
EDSR	26.62
5-5-5-5	27.91
8-4-6-2	27.33
2-4-6-8	26.69
4-6-6-4	27.87
6-6-6-6	28.71

(1)5-5-5-5 模型结构

$GRBNS = [GRBN1, GRBN2, GRBN3, GRBN4]$ ($N = 4$)

每 5 组 GRB 组成一级 GRBN,共 4 级 GRBN。

(2)8-6-4-2 模型结构

$GRBNS = [GRBN1, GRBN2, GRBN3, GRBN4]$ ($N = 4$)

分别为每 8 组、6 组、4 组、2 组 GRB 组成一级 GRBN,共 4 级 GRBN。

(3)2-4-6-8 模型结构

$GRBNS = [GRBN1, GRBN2, GRBN3, GRBN4]$ ($N = 4$)

分别为每 2 组、4 组、6 组、8 组 GRB 组成一级 GRBN,共 4 级 GRBN。

(4)4-6-6-4 模型结构

$GRBNS = [GRBN1, GRBN2, GRBN3, GRBN4]$ ($N = 4$)

分别为每 4 组、6 组、6 组、4 组 GRB 组成一级 GRBN,共 4 级 GRBN。

如表 1 所列,无论是 GRN 的哪种结构,PSNR 指标都优于 EDSR。本实验发现:同层结构中 5-5-5-5 结构的训练效率最高,而且本文并没有乘以固定数值 0.1。本文将延续这种 GRBNS 结构进行下面的加层实验。通过比较训练收集的 PSNR 数值可以得出:6-6-6-6 结构的性能指标比 5-5-5-5 结构的性能更加优异,并且 6-6-6-6 结构的深度学习网络模型层数高于 EDSR 的最深层数。实验表明,本文提出的 GRN 模型比 EDSR 模型更适合做深层的网络结构。本文先训练网络尺寸为 2 倍的模型,当模型收敛后,本文用其来预训练其他尺寸的网络。图 3 给出 Bicubic,EDSR 及本文方法的结果比较,并基于 PSNR/SSIM 比较 3 种算法的重建性能,如表 2 所列。

表 2 其他测试图片的 PSNR/SSIM 值

训练集(图片)	尺寸	Bicubic	EDSR	GRN
Urban100 (im062)	2 倍	19.82/0.72	21.32/0.825	22.01/0.832
	3 倍	18.13/0.70	20.64/0.802	21.16/0.823
	4 倍	17.34/0.66	19.36/0.783	19.51/0.797
DIV2K (0390)	2 倍	18.96/0.70	22.92/0.874	23.32/0.886
	3 倍	18.51/0.68	22.04/0.863	22.26/0.872
	4 倍	17.02/0.67	21.39/0.842	21.45/0.846
DIV2K (0724)	2 倍	17.96/0.69	23.43/0.842	23.76/0.856
	3 倍	17.03/0.68	22.97/0.828	23.01/0.836
	4 倍	16.78/0.66	22.43/0.803	22.44/0.812
DIV2K (0752)	2 倍	17.64/0.73	21.09/0.839	21.87/0.853
	3 倍	17.25/0.71	20.84/0.824	21.06/0.835
	4 倍	16.87/0.70	20.11/0.802	20.12/0.807

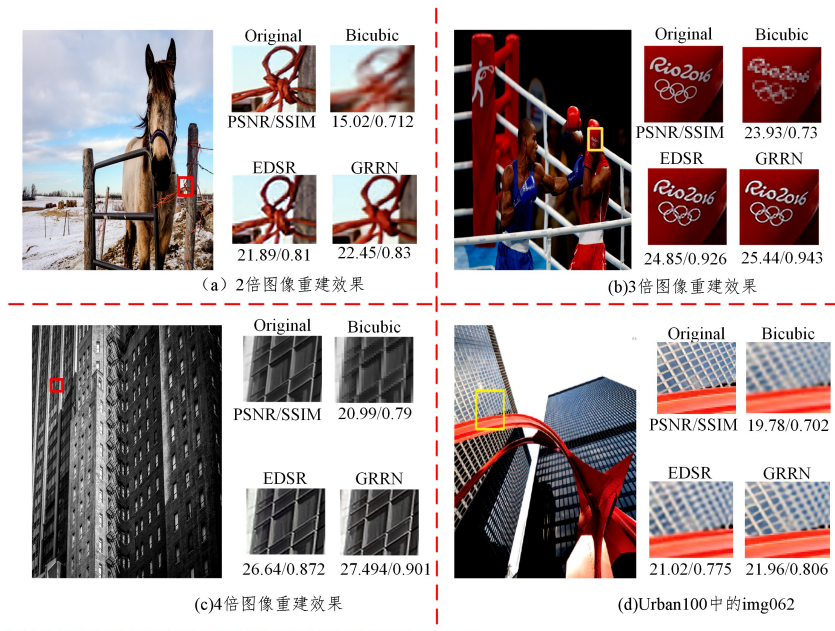


图3 Bicubic、EDSR及本文方法的结果比较

结束语 本文延续了跳跃连接、残差学习、梯度裁剪、递归残差块等优秀的思想,创新性地提出了残差块特征融合的思想。本文提出全新的模型架构,通过删除简化残差结构,让残差学习思想更适合图像超分辨这种低视觉领域;增加 Shrinking 层、Expanding 层,使得网络模型的计算复杂度大幅度降低;增加残差块融合思想,使网络模型产生“自适应”更新权重,以达到“自适应”改善信息流的目的。本文延续 ADAM 优化器的思想,抛弃常用的 L2 代价函数,采用 L1 代价函数。本文的网络模型不仅在训练时间上有明显的减少,而且在图像重建指标上相对于最新的 EDSR 模型有了进一步的改善和提升。在实验过程中发现本文提出的网络模型训练更稳定,而且可以很好地避免层数过多导致的梯度衰减和梯度爆炸问题。全局残差递归网络比 EDSR 的网络模型更深,探索性的创新 GRRN 网络模型比 EDSR 网络模型更适合做深层的网络模型。

参考文献

- [1] ALLEBACH J, WONG P W. Edge-directed interpolation[C]//ICIP. 1996.
- [2] DONG C, LOY C C, HE K, et al. Learning a deep convolutional network for image super-resolution[C]//ECCV. 2014.
- [3] DONG C, LOY C C, TANG X. Accelerating the superresolution convolutional neural network[C]//ECCV. 2016.
- [4] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//CVPR. 2016.
- [5] KINGMA D, BA J. Adam: A method for stochastic optimization [C]//ICLR. 2014.
- [6] KIM J, LEE J K, LEE K M. Accurate image super resolution using very deep convolutional networks[C]//CVPR. 2016.
- [7] DONG C, LOY C C, TANG X. Accelerating the superresolution convolutional neural network[C]//ECCV. 2016.
- [8] GLOROT X, BORDES A, BENGIO Y. Deep sparse rectifier neural networks[C]//AISTATS. 2011.
- [9] LEDIG C, THEIS L, HUSZ' A F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//CVPR. 2017.
- [10] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//CVPR. 2017.
- [11] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution[C]//CVPR. 2016.
- [12] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//CVPR. 2016.
- [13] SZEGEDY C, IOFFE S, VANHOUCKEET V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[C]//AAAI. 2017.
- [14] KARRAS T, AILA T, LAINE S, et al. Progressive growing of gans for improved quality, stability, and variation[C]//ICLR. 2017.
- [15] TIMOFTE R, AGUSTSSON E, GOOL L V, et al. Ntire 2017 challenge on single image super-resolution Methods and results [C]//CVPRW. 2017.
- [16] ZHAO H, GALLO O, FROSIO I, et al. Loss functions for neural networks for image processing[J]. arXiv:1511.08861, 2015.
- [17] 李键红, 吴亚, 吕巨建. 基于组稀疏表示的在线单帧图像超分辨率算法[J]. 计算机科学, 2018, 45(4): 312-318.
- [18] 刘甜甜, 包芳勋, 张云峰, 等. 有理分形曲面造型及其在图像超分辨率中的应用[J]. 计算机科学, 2018, 45(3): 35-45.
- [19] TAI Y, YANG J, LIU X, et al. Memnet: A persistent memory network for image restoration[C]//ICCV. 2017.