

改进的密集连接网络遥感图像超分辨重建^①柏宇阳^② 朱福珍^③ 巫 红

(黑龙江大学电子工程学院 哈尔滨 150080)

摘 要 遥感图像超分辨增加了遥感图像的细节信息,在遥感图像处理中有重要的地位。为了进一步提高遥感图像超分辨的重建效果,本文提出一种改进的密集连接网络遥感图像超分辨重建算法。首先对基于残差网络的深度超分辨算法(VDSR)进行改进,结合密集连接网络(DenseNet),将残差网络中的残差块替换成密集块,并且添加一组密集层与瓶颈层,实现 DenseNet 网络结构的改进,同时,修改网络激活函数为 PReLU 函数,网络训练采用 L1 损失函数。为了使网络在遥感图像上具有更好的效果,训练网络时,数据集全部采用遥感图像作为训练样本。当训练的 epoch 达到了大约 35 次时网络已经收敛。实验结果表明,与 VDSR 算法相比,本文改进的算法对遥感图像的效果更优,峰值信噪比(PSNR)平均增加了 1.05 dB,结构相似度(SSIM)平均增加了 0.042。

关键词 遥感图像;超分辨率;密集连接网络(DenseNet);深度学习

0 引 言

遥感图像分辨率是判定遥感图像质量的重要指标之一,高分辨率的遥感图像可反映出更多的地面细节信息。图像超分辨重建(super-resolution reconstruction, SRR)是一种利用单帧或多帧的低分辨率图像(low resolution, LR)通过专用的算法,在不改变成像硬件设备条件的情况下,生成较清楚的高分辨率图像(high resolution, HR)的技术,它使重建后的高分辨率图像的视觉效果远远超过参与重建的低分辨率图像。

图像超分辨算法总体上可分为基于重构的超分辨^[1-5]和基于学习的超分辨^[6-13]两大类。基于学习的超分辨算法相比于基于重构的超分辨算法在视觉效果上更好,因为此方法直接学习了低分辨率图像与高分辨率图像之间的映射关系。常见的基于学习

的超分辨率算法包括基于样本的方法、基于邻域嵌入的方法、基于稀疏表示的方法和基于卷积神经网络的方法。本文研究的方法属于使用基于学习超分辨率算法。基于学习的超分辨算法是 2002 年 Freeman 等人^[14]提出的,该算法中高分辨率图像块和低分辨率图像块的关系可通过机器学习实现,他们将此映射关系应用于马尔科夫网络,并利用贝叶斯置信传播理论来进行超分辨率重建。随后,在 2004 年,Chang 等人^[15]将局部线性嵌入引入到超分辨率重建方法中,提出一种基于邻域嵌入的超分辨率重建算法。在 2010 年,Yang 等人^[16]认为图像具有稀疏性,因此将稀疏表示应用到图像超分辨中。

卷积神经网络在 20 世纪 80 到 90 年代间出现,但在最近几年中得到了广泛的应用。卷积神经网络图像超分辨的基本思路是通过给定的训练集,其中包括低分辨率图像和高分辨率图像,将低分辨率图像作为网络的输入,网络的输出即为高分辨率图像。

① 国家自然科学基金(61601174),黑龙江省自然科学基金(F2018026),黑龙江省博士后科研启动金(LBH-Q17150),黑龙江省普通高等学校电子工程重点实验室开放课题及省高校科技创新团队(2012TD007)资助项目。

② 男,1993 年生,硕士生;研究方向:遥感图像处理;E-mail:307486164@qq.com

③ 通信作者,E-mail:zhufuzhen@hlju.edu.cn

(收稿日期:2021-02-04)

卷积神经网络的图像超分辨中最具有代表性的是超分辨卷积神经网络 (super-resolution convolutional neural network, SRCNN) 算法^[17] 与快速超分辨卷积神经网络 (accelerating the super-resolution convolutional neural network, FSRCNN) 算法^[18], 其中 SRCNN 是基于卷积神经网络的图像超分辨的首作, 其通过构建 3 层卷积层来学习 LR-HR 的映射关系。后来人们发现卷积网络深度越深, 网络可以拟合出的函数越准确, 结果也越好。但是随着网络越来越深, 网络训练出现了梯度消失现象, 也导致网络训练时难收敛的问题。He 等人^[19] 在 2016 年提出了残差卷积神经网络 (residual network, ResNet) 解决了此问题, 网络泛化性能也得到了提升。Kim 等人^[20] 提出了深度超分辨 (very deep super-resolution, VDSR) 算法, 该算法将残差网络用在了图像超分辨问题上, 也解决了超分辨网络较深时训练难收敛的问题。最近研究表明, 如果卷积网络包含接近输入的层和接近

输出的层之间的较短连接, 则网络可以更深入、精确、有效地进行训练。密集连接网络^[21] (densely connected network, DenseNet) 就是这样减轻了梯度消失的问题, 也加强了特征的传递。其算法与残差网络相似, 残差网络是若干个残差块进行级联, 每个残差块由若干层组成, 每一层的输出都是由本层特征与前一层特征之和。而 DenseNet 是由若干个密集块级联, 每个密集块由若干个层组成, 但每一层的输出都是由本层特征与前面所有层的特征之和。由于有更多的短连接, DenseNet 算法的效果也比残差网络要好一些。

1 改进的密集连接神经网络

1.1 遥感图像退化模型

本文的超分辨网络针对的是遥感图像, 遥感图像退化模型如图 1 所示。

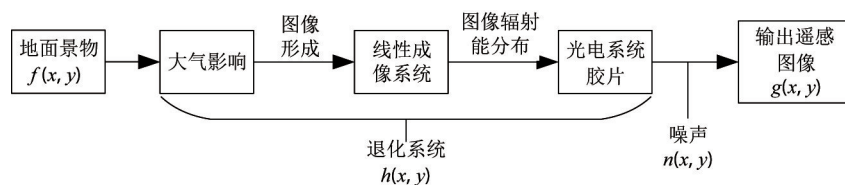


图 1 遥感图像退化模型

地面景物图像受到大气影响、遥感成像系统、运动模糊、噪声等退化因素的综合作用, 形成实际获取的降质遥感图像。遥感图像的退化系统 $h(x, y)$ 较复杂, 地面景物通过大气后, 形成了大气退化图像, 再由线性成像系统, 通过图像辐射能分布情况形成退化图像, 其数学模型为

$$g(x, y) = f(x, y) * h(x, y) + n(x, y) \quad (1)$$

式中, $f(x, y)$ 是高分辨率场景图像, $h(x, y)$ 是遥感退化系统, $n(x, y)$ 是噪声, $g(x, y)$ 是实际获取的遥感图像, “*” 代表卷积运算。

1.2 网络结构的改进

目前的 VDSR 和 DenseNet 超分辨网络都选用了 91-images 或 291-images 常规自然景物的彩色图像作为数据集, 这些图像并不包含任何遥感图像。遥感图像的退化系统与常规彩色图像不同, 且遥感图像包含了大量的土地、林地、湿地等各种的地面,

这些图像的纹理和高频细节信息等, 也与 91-images 或 291-images 彩色图像存在一些差异, 导致现有算法对遥感图像的超分辨效果一般。本文将训练测试样本改为遥感图像, 在 VDSR 网络的基础上, 结合了密集连接网络, 对网络进行了优化。并在网络结构、激活函数、学习规则等方面进行了改进, 使其客观参数与主观视觉上都好于其他的算法。本文算法流程分为训练集的构造、超分辨模型训练和测试过程三部分。

由于当网络结构较深时, 网络层拥有越大的感受野, 效果就越好, 深层网络在训练时较困难, 而密集连接网络在层数较深时也容易收敛, 本文算法仍使用密集连接网络。改进的网络结构如图 2 所示, 网络中将所有的残差部分更换为密集部分, 同时为了扩展网络深度构建了两组密集连接部分, 如图中虚线所示。输入为低分辨率图像, 通过一个卷积层

部分,输出一个低分辨率特征图。然后将此特征图通向了密集部分 A,此部分包括了 8 个密集块 (A1

~ A8), 再之后连接着瓶颈层 A, 这个层是将密集部分 A 中的密集块 A1 ~ A8 的输出进行特征叠加。

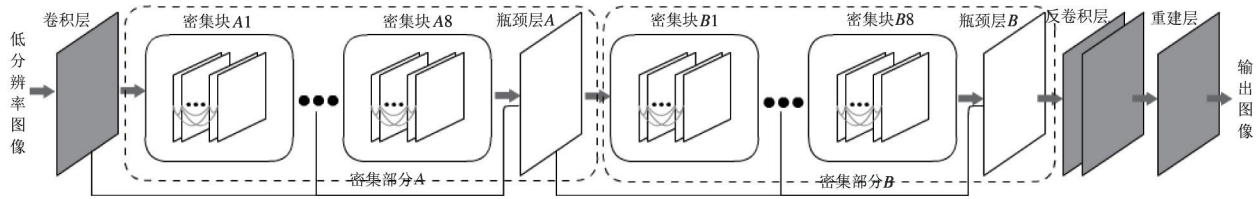


图 2 改进网络的结构

瓶颈层 A 的表达式为

$$bottleneck = f(\sum_i block(i) * W_i) \quad (2)$$

式中, $bottleneck$ 为瓶颈层 A 的输出, $block(i)$ 为密集块的输出, W_i 为第 i 块对应的卷积核, “*” 代表卷积运算, $f(\cdot)$ 为激活函数。

之后是密集部分 B, 也由 8 个密集块组成 (B1 ~ B8), 再之后再通过瓶颈层 B, 这个层是将密集部分 B 中的密集块 B1 ~ B8 的输出进行特征叠加。叠加后的特征图通过反卷积层增加图像尺寸放大图像, 最后经过重建层得到输出图像高分辨率图像, 重建层也是一个卷积层。

网络包含两组密集部分, 分别为密集部分 A 和密集部分 B。每部分都是由 8 个密集块组成, 每个密集块也都含有 8 个卷积层, 且这 8 个卷积层的每个层都与其前面的所有层相互连接, 其表达式为

$$layer(k) = f(\sum_{i=1}^{k-1} layer(i) \times W_i) \quad (3)$$

式中, $layer(k)$ 为第 k 层的输出, W_i 为卷积核, $f(\cdot)$ 为激活函数。

这些卷积层的卷积核尺寸为 3×3 , 特征个数为 16, 为了使卷积前后的图像尺寸保持不变, 卷积前图像边缘进行补 0 填充, 填充尺寸 (Padding) 为 1。

输入为由双三次插值为退化过程的低分辨率遥感图像, 经过两组密集部分后通过反卷积增大图像尺寸, 反卷积层的输出为 256 个高分辨率特征图。最后的重建层由一个卷积层组成, 尺寸为 3×3 , Padding 为 1, 通过这一层将 256 个高分辨率特征图重建得到最终的高分辨率输出图像。

1.3 激活函数的改进

尽管使用了残差网络, 网络在训练时仍然可能

出现训练不易收敛的问题, ReLU 函数是斜坡函数, 表达式为

$$\text{ReLU}(x) = \max(x, 0) \quad (4)$$

ReLU 函数将输入矩阵中所有负数都置为 0, 正值不变。此激活函数有一个缺点, 就是训练很脆弱, 很容易死亡。例如, 当一个较大的梯度流过一个 ReLU 神经元, 更新后参数这个神经元就再也不会对任何数据有激活现象。那么在这之后此神经元的梯度永远都是 0, 也就是说 ReLU 函数有时会在训练过程中不可逆地死亡, 即在训练过程中有些神经元没有激活。因此, 本文对 ReLU 函数进行了改进, 使用了 PReLU 作为激活函数, PReLU 函数可以解决 ReLU 函数的死亡问题, 其表达式为

$$\text{PReLU}(x) = \max(x, 0) + \min(ax, 0) \quad (5)$$

式中, a 为较小的数, 可取 0.01。当 $x \geq 0$ 时, PReLU 函数与 ReLU 函数值都为 x 。但 PReLU 函数比 ReLU 函数多了一项 $\min(ax, 0)$, 当 $x < 0$ 时, ReLU 函数为 0, 而 PReLU 函数是一条斜率为 a 且过原点的直线。PReLU 函数只增加了极少的参数, 网络计算量只增加了一点点, 但网络在训练时健壮了很多。

1.4 学习规则的改进

DenseNet 网络训练的损失函数常用有 L_1 范数和 L_2 函数。 L_1 范数损失函数也叫最小绝对值误差, 总的来说是把目标值与估计值的绝对差值之和最小化。 L_2 范数损失函数也叫最小平方误差, 总的来说是使目标值与估计值差值的平方之和最小化。由于 L_1 范数能处理数据中的异常值, 其鲁棒性比 L_2 更好, 本文训练使用 L_1 范数作为损失函数, 训练的最终目标是使 L_1 范数达到最小值。取实际高分辨率图像与重建后的图像的误差绝对值之和, 其表达式为

$$L_1(y, f) = \sum_{i,j} |y(i, j) - f(i, j)| \quad (6)$$

式中, $y(i, j)$ 为实际高分辨率图像, $f(i, j)$ 为重建后的图像。当 L_1 范数达到最小值时, 重建后的图像最接近于实际高分辨率图像。

网络优化器仍然采用 Adam 优化器, 它是一种可以替代传统随机梯度下降 (stochastic gradient descent, SGD) 过程的算法, Adam 算法的表达式如下:

$$\begin{aligned} m_t &= \frac{\beta_1 m_{t-1} + (1 - \beta_1) g_t}{1 - \beta_1^t} \\ v_t &= \frac{\beta_2 v_{t-1} + (1 - \beta_2) g_t^2}{1 - \beta_2^t} \\ \theta_{t+1} &= \theta_t - \frac{\alpha \cdot m_t}{\sqrt{v_t} + \varepsilon} \end{aligned} \quad (7)$$

式中, β_1 、 β_2 为参数, 分别取 0.9 和 0.999。 g_t 是 t 时间步的梯度, m_t 是对梯度的一阶矩估计, v_t 是对梯度的二阶矩估计, ε 取 10^{-8} 防止除数为 0, α 为初始学习率。

和传统的随机梯度下降不同, Adam 算法在训练过程中通过计算梯度的一阶矩估计和二阶矩估计, 动态地调整学习率。而传统的随机梯度下降一直保持固定的学习率更新所有的权重。Adam 算法可以使每一次迭代的学习率都在一个较好的数值, 使训练质量提高。Adam 的初始学习率 α 如果设得过大, 学习速度则会非常快, 但其收敛结果不好; α 设得过小会导致训练时间过长, 且对结果优化不大。因此, 本文采用默认的初始学习率 0.001, 此数值既可以保证学习速度, 又可以优化收敛结果。

2 实验与结果分析

2.1 网络训练实验过程

本文算法的实验环境为 Intel Core I7-7700K CPU@4.20 GHz, 16 GB 内存, 64 位 Ubuntu 系统, GPU 为 GTX1060, 深度学习框架使用了 TensorFlow 1.10。

训练样本使用合作单位提供的尺寸为 $20\,000 \times 20\,000$ 的遥感图像。为了提高网络训练速度和训练样本数量, 对每个小尺寸的遥感图像有交叉地截取为小图像, 每相隔 20 像素截取一张 100×100 尺寸的小图像, 最终得到超过 8 万张尺寸为 100×100 的

小图像作为网络训练数据集。为保证数据集的数量, 再将子图像分别旋转 90° 、 180° 、 270° , 使数据集的数目增加了 3 倍, 最终得到了约 2.5 万张 100×100 的小图像作为训练集。将训练集图像本身作为网络的输出, 训练集图像 (2 倍欠采样) 缩小 2 倍到 50×50 , 作为网络输入的低分辨率图像。

网络训练执行一次 epoch 大约 14 min, 当训练 epoch 达到 35 次时, 网络已经收敛, 此时累计用时大约 8 h。网络训练过程中误差与迭代次数的关系如图 3 所示, 网络的测试集峰值信噪比 (peak signal to noise ratio, PSNR) 与迭代次数的关系如图 4 所示, 每次 epoch 训练后进行测试, 约在 35 次训练时 PSNR 值接近了稳定。

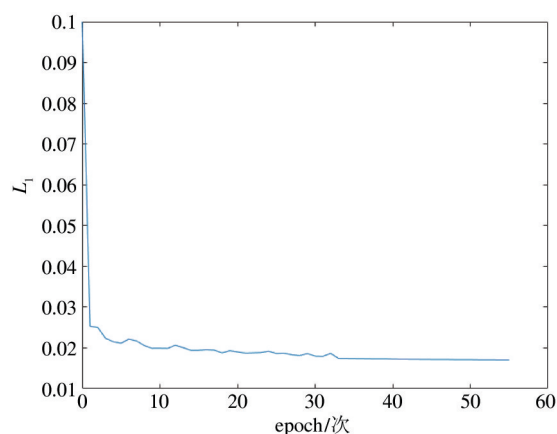


图3 网络训练误差与 epoch 的关系

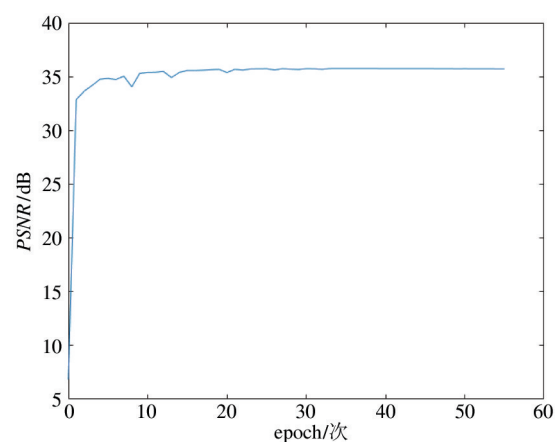


图4 测试集 PSNR 与 epoch 的关系

2.2 网络泛化实验结果与分析

取测试集 PSNR 最大时的网络作为最终的结果进行泛化实验, 此时 epoch 为第 35 次。在实验过程

中,选取 8 张遥感图像进行实验,如图 5 所示。实验选择低分辨率图像、双三次插值、VDSR 算法、本文算法作对比。为了实验对比的合理性,实验中缩放

因子都取 2,训练与测试图像均使用遥感图像,在此列出 4 组实验结果进行对比,遥感图像超分辨对比结果如图 6 ~ 图 9 所示。

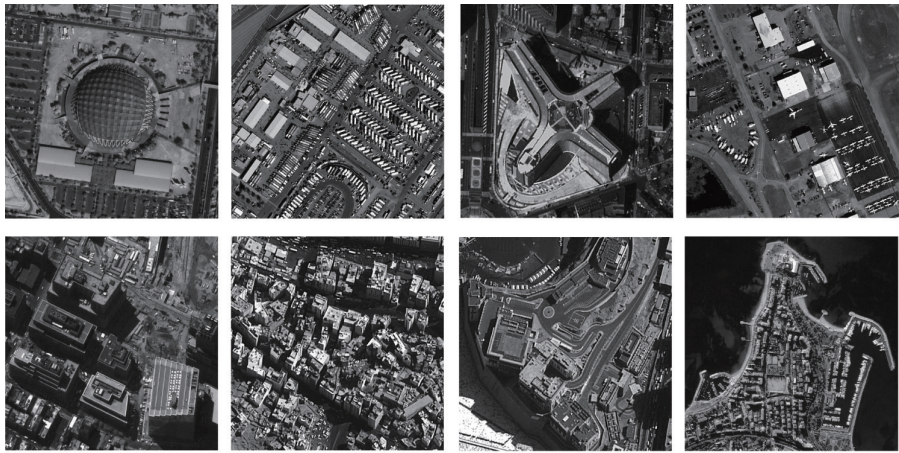


图 5 测试图像

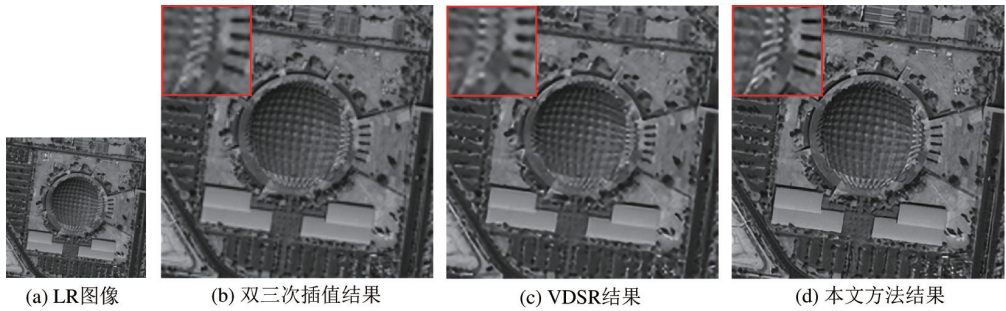


图 6 第 1 组图像超分辨效果比较

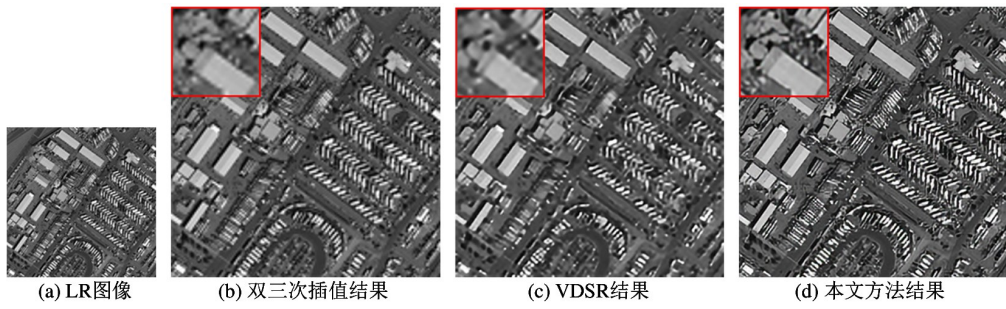


图 7 第 2 组图像超分辨效果比较



图 8 第 3 组图像超分辨效果比较

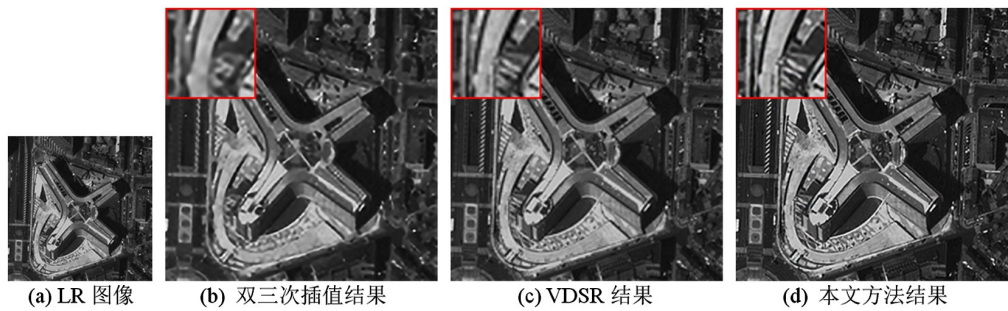


图9 第4组图像超分辨率效果比较

由图6~图9可以看出,与双三次差值与VDSR结果相比较,本文方法得到的超分辨率结果视觉上更清晰,细节更丰富,地物细节更容易分辨出来。

为了客观地评价遥感图像超分辨率重建效果,泛化实验中采用峰值信噪比(PSNR)和结构相似度(structural similarity, SSIM)作为超分辨率算法的质量评价指标,其定义如式(8)和式(9)所示。

$$PSNR = 10 \lg \frac{255^2 M \cdot N}{\sum_{i,j} [y(i,j) - f(i,j)]^2} \quad (8)$$

式中, $y(i,j)$ 为实际高分辨率图像, $f(i,j)$ 为重建后的图像。超分辨率图像与实际高分辨率图像存在误差,其误差越小 PSNR 越高,说明超分辨率结果图像的失真度越低,重建效果越好。

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (9)$$

式中, μ_x 、 μ_y 为均值, σ_x 、 σ_y 为标准差, σ_{xy} 为协方差, $c_1 = (0.01L)^2$, $c_2 = (0.03L)^2$, $L = 255$ 。SSIM 表示超分辨率图像与实际高分辨率图像的结构相似程度,其取值为范围为[0,1],SSIM 越接近 1 则说明它们的结构越相似,重建效果越好。

泛化实验统计的 PSNR 和 SSIM 结果分别如表 1 和表 2 所示。综上所述,本文改进的密集连接网络遥感图像超分辨率重建算法不论是主观视觉上还

表1 3种算法的 PSNR 对比(dB)

图像	双三次插值	VDSR	本文方法
第1组	26.73	27.74	28.66
第2组	18.76	20.08	20.61
第3组	23.05	24.72	26.03
第4组	25.07	26.30	27.73
平均值	23.40	24.71	25.76

表2 3种算法的 SSIM 对比

图像	双三次插值	VDSR	本文方法
第1组	0.821	0.839	0.895
第2组	0.692	0.768	0.813
第3组	0.794	0.853	0.883
第4组	0.829	0.869	0.904
平均值	0.784	0.832	0.874

是在客观评价参数上,超分辨率效果都优于其他算法。相比当前先进的 VDSR 算法,本文算法其 PSNR 平均提高约了 1.05 dB,SSIM 平均提高了 0.042。

3 结论

本文改进了 VDSR 算法,网络训练与测试使用的数据集采用了遥感图像,结合了密集连接网络,将残差网络中的残差块替换成密集块,并增加了一组密集层与瓶颈层完成了网络结构的改进。同时,在激活函数、学习规则方面也进行了改进,修改网络激活函数为 PReLU 函数,网络训练采用了 L_1 损失函数。当训练的 epoch 达到了大约 35 次时网络已经收敛。

网络泛化实验结果表明,与 VDSR 算法相比,本文改进的算法对遥感图像的效果更优。图像的超分辨率结果在主观视觉上更清楚,细节信息更丰富,地物细节更容易分辨出来。客观评价指标 PSNR 平均增加了 1.05 dB,SSIM 平均增加了 0.042。

参考文献

- [1] Schultz R R, Stevenson R L. Extraction of high-resolution frames from video sequences[J]. *IEEE Transactions on Image Processing*, 1996, 5(6): 996-1011
- [2] Patti A J, Altunbasak Y. Artifact reduction for set theoretic super resolution image reconstruction with edge

- adaptive constraints and higher-order interpolants [J]. *IEEE Transactions on Image Processing*, 2001, 10(1): 179-186
- [3] Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution [C] // European Conference on Computer Vision, Amsterdam, Netherlands, 2016: 694-711
- [4] Castro E B, Nakano M, Perez G S, et al. Improvement of image super-resolution algorithms using iterative back projection[J]. *IEEE Latin America Transactions*, 2017, 15(11): 2214-2219
- [5] Dun Y J, Da Z Y, Yang S, et al. Kernel-attended residual network for single image super-resolution[J] *Knowledge-Based Systems*, 2021, 213:106663
- [6] Chang S G, Cvetkovic Z, Vetterli M. Locally adaptive wavelet-based image interpolation[J]. *IEEE Transactions on Image Processing*, 2006, 15(6): 1471-1485
- [7] 王威,张彤,王新.用于图像超分辨率重构的深度学习方法综述[J]. 小型微型计算机系统, 2019, 40(9): 1891-1896
- [8] 罗喜君,陶婷婷,王丹.基于学习的数字图像超分辨率技术概述[J]. 计算机光盘软件与应用, 2014, 17(3): 295, 297
- [9] Freeman W T, Jones T R, Pasztor E C. Example-based super-resolution[J]. *IEEE Computer Graphics and Applications*, 2002, 22(2): 56-65
- [10] 贺瑜飞,高宏伟.基于多层连接卷积神经网络的单帧图像超分辨率重建[J]. 计算机应用与软件, 2019, 36(5): 220-224
- [11] Lai W S, Huang J B, Narendra A, et al. Deep laplacian pyramid networks for fast and accurate super-resolution [C] // IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA, 2017: 5835-5843
- [12] 王容,张永辉,张健,等.基于卷积神经网络的人脸超分辨率重建[J]. 计算机工程与设计, 2019, 40(9): 2614-2619
- [13] 李素梅,雷国庆,范如.基于卷积神经网络的深度图超分辨率重建[J]. 光学学报, 2017, 37(12): 132-140
- [14] Freeman W T, Jones T R, Pasztor E C. Example-based super-resolution[J]. *IEEE Computer Graphics and Applications*, 2002, 22(2): 56-65
- [15] Chang H, Yeung D Y, Xiong Y. Super-resolution through neighbor embedding [C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, USA, 2004: 275-282
- [16] 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
- [17] 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): 295-306
- [18] Dong C, Loy C C, Tang X. Accelerating the super-resolution convolutional neural network [C] // European Conference on Computer Vision, Amsterdam, Netherlands, 2016: 391-405
- [19] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C] // IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 770-778
- [20] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks [C] // IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 1646-1653
- [21] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks [C] // IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA, 2017: 4700-4708

An improved remote sensing images super-resolution method based on densely connected network

Bai Yuyang, Zhu Fuzhen, Wu Hong

(College of Electrical Engineering, Heilongjiang University, Harbin 150080)

Abstract

Remote sensing image super-resolution improves the detail information of remote sensing image, which is of great significance in remote sensing image processing. In order to further improve the super-resolution reconstructed effect of remote sensing image, an improved remote sensing image super-resolution algorithm based on densely connected network is proposed. First, the very deep super-resolution algorithm (VDSR) based on residual network is improved. Combined with densely connected network (DenseNet), the residual blocks in the residual network are replaced by densely blocks, and a group of densely layers and bottleneck layers are added to improve the DenseNet network structure. Second, the activation function is modified to PReLU function, and the L1 loss function is adopted in training. In order to make the super-resolution network have better effect on remote sensing images, remote sensing images are used as the training data set in training. The network converges when the epoch reaches about 35 times. Finally, the experimental results show that the effect of the improved algorithm is better than the effect of VDSR algorithm on remote sensing images. The PSNR increases by 1.05 dB and the SSIM increases by 0.042 on average.

Key words: remote sensing image, super-resolution, densely connected network (DenseNet), deep learning