

DOI: 10.13482/j.issn1001-7011.2019.12.035

投稿网址: <http://hdzrxb.cbpt.cnki.net>

# 改进的残差卷积神经网络遥感图像超分辨重建

柏宇阳, 朱福珍

(黑龙江大学 电子工程学院, 哈尔滨 150080)

**摘 要:** 为了提高遥感图像超分辨效果, 对深度超分辨( Very deep super-resolution, VDSR) 算法进行了深入研究, 提出了一种改进的残差卷积神经网络超分辨算法。在 VDSR 基础上, 将神经网络的卷积层数目由 20 层增加到 24 层, 改进激活函数为 PReLU 函数, 并以 80 000 张遥感图像作为训练集, 训练迭代次数到 40 000 次网络收敛, 最终针对遥感图像进行了超分辨的网络训练。结果表明, 与当前较先进的 VDSR 算法相比, 重建结果图像中的细节信息明显增多, PSNR 提高 0.45 dB, SSIM 提高 0.023。

**关键词:** 遥感图像; 超分辨率; 残差网络; 深度学习

中图分类号: TP394.1

文献标志码: A

文章编号: 1001-7011(2020)03-0372-07

## An improved remote sensing images super-resolution method based on residual convolution neural network

BAI Yuyang, ZHU Fuzhen

(College of Electronic Engineering, Heilongjiang University, Harbin 150080, China)

**Abstract:** In order to improve the super-resolution effect of remote sensing image, an improved residual convolution neural network super-resolution algorithm is proposed based on VDSR, through which the number of the convolution layers of neural network is increased from 20 to 24, activation function is improved to PReLU function, and the number of training iterations reaches 40 000 with 80 000 remote sensing images being used as training set. Finally, a super-resolution network is trained for remote sensing images. Network generalization experiments show that the subjective visual effect and objective evaluation parameters are improved and the PSNR is increased by 0.45 dB compared with the existing convolution neural network super-resolution algorithms.

**Keywords:** remote sensing image; super-resolution; residual neural network; deep learning

## 0 引 言

遥感图像的分辨率是判定遥感图像质量的重要指标之一, 高分辨率图像可见更多的细节信息。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

收稿日期: 2019-09-29

基金项目: 国家自然科学基金资助项目( 61601174); 黑龙江省博士后科研启动金资助项目( LBH-Q17150); 黑龙江省普通高等学校电子工程重点实验室( 黑龙江大学) 开放课题资助及省高校科技创新团队资助项目( 2012TD007); 黑龙江省省属高等学校基本科研业务费基础研究项目( KJCXZD201703); 黑龙江省自然科学基金资助项目( F2018026)

通讯作者: 朱福珍( 1978-), 女, 副教授, 博士, 主要研究方向: 图像超分辨、压缩感知、神经网络、图像融合和图像处理, E-mail: zhufuzhen@hlju.edu.cn

引文格式: 柏宇阳, 朱福珍. 改进的残差卷积神经网络遥感图像超分辨重建 [J]. 黑龙江大学自然科学学报, 2020, 37( 3): 372-378.

image, LRI)在不改变成像硬件设备条件的情况下,通过相应的算法,生成较清楚的高分辨率图像(High resolution image, HRI)的技术。它可使重建后的高分辨率图像的视觉效果超过任何帧的低分辨率图像,通常分为基于重构<sup>[1-4]</sup>和基于学习<sup>[5-13]</sup>两种类型。本文使用基于学习的超分辨率算法,直接学习低分辨率图像与高分辨率图像之间的映射关系。当将低分辨率图像传入到神经网络的输入时,输出的是预测的高分辨率图像。基于学习的超分辨算法相比于基于重构的超分辨算法效果更好,其相关研究也更加具有实用意义,其中包括基于样本、邻域嵌入、稀疏表示和基于卷积神经网络的方法。2002年,Freeman等提出了一种基于样本学习的图像重建算法,算法中说明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<sup>[6]</sup>;Chang等将局部线性嵌入引入到超分辨率重建方法中,提出了一种基于邻域嵌入的超分辨率重建算法;Yang等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。

卷积神经网络的出现可以追溯到几十年前,但其最近几年才被广泛应用,主要原因也是它在图像处理方面的巨大成功,如物体检测、人脸识别、行人检测等。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,如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。卷积神经网络的图像超分辨中最具有代表性的是Dong等提出的超分辨卷积神经网络(Super-resolution convolutional neural network, SRCNN)算法<sup>[14]</sup>,它通过构建出三层卷积层来学习LR-HR的映射关系。后来Dong等提出了快速超分辨卷积神经网络(Accelerating the super-resolution convolutional neural network, FSRCNN)算法<sup>[15]</sup>,最后一层使用反卷积层放大尺寸,因此不需要将输入图像通过双三次插值放大尺寸。何恺明等提出了残差卷积神经网络(ResNet),解决了之前网络结构较深时无法训练的问题,性能也得到了提升<sup>[16]</sup>。Kim等提出了VDSR算法,该算法使用了残差网络,解决了网络较深时训练难收敛的问题<sup>[17]</sup>。由于输入的低分辨率图像和输出的高分辨率图像在很大程度上是相似的,也就是指低分辨率图像携带的低频信息与高分辨率图像的低频信息相近,训练时会多花费大量的时间,实际上只需要学习高分辨率图像和低分辨率图像之间的高频残差即可,残差网络结构的思想特别适合用来解决超分辨率问题,可以说影响了之后的深度学习超分辨率方法,VDSR是最直接明显的学习残差的结构。VDSR将低分辨率图像经过双三次插值后,得到与目标尺寸相同的图像作为网络的输入,再将这个图像与网络学到的残差相加得到最终的输出,其残差网络为20个卷积层与20个激活函数相互级联,每层卷积层由 $3 \times 3$ 的卷积核组成,其激活函数使用了ReLU函数。

## 1 改进的残差卷积神经网络

由于VDSR算法以彩色图像作为训练集,而对遥感图像的超分辨效果没有其对彩色图像的超分辨效果好,本文在VDSR的基础上,对网络的结构与激活函数进行了优化,使用遥感图像作为训练集,客观参数与主观视觉上都好于其他算法。

### 1.1 网络结构的改进

由于当网络结构较深时,网络层拥有更大的感受野,其效果更好。但网络较深时网络训练会更困难,而使用残差学习,残差图像比较稀疏,大多数的值都为0或者比较小,因此收敛会更快,可以解决网络结构较深时无法训练的问题,本文算法使用了残差网络。改进的网络结构如图1所示,输入层为低分辨率图像,经过若干个卷积层后得到一张残差图像,最终输出为残差图像与低分辨率图像的和,即为高分辨率图像。为了提高超分辨效果,本文卷积层深度由20层扩展到24层,激活函数由ReLU函数修改为PReLU函数。很明显,本文算法与VDSR、SRCNN算法都没有使用池化层,因为下采样层通常应用于图像的分类,保证了图像特征。同时,可以减少参数增加网络性能,但图像超分辨的目的是为了获取图像更多的细节信息,与池化层的作用恰好相反,所以在图像超分辨算法中并不需要池化层。

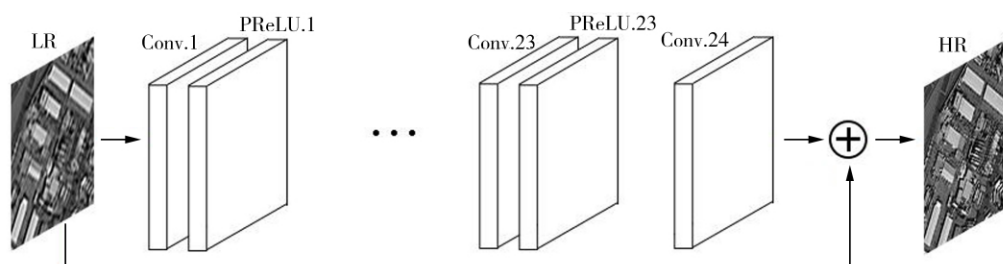


图1 改进的网络结构

Fig.1 Improved network structure

### 1.1.1 隐含层

本文算法使用了 24 个卷积层,第 1 层为 64 个卷积核尺寸为  $5 \times 5$  的卷积层。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,补 0 填充尺寸( Padding) 为 2。作为整个神经网络的第一层,输入为由双三次插值为退化过程的低分辨率遥感图像,输出为 64 个特征图。中间的 22 层,每一层都是 64 个卷积核的卷积层。为了降低计算量,卷积核尺寸为  $3 \times 3$ ,Padding 为 1,每个卷积层的输入为 64 个特征图,输出也为 64 个特征图,图像尺寸在卷积前后仍然保持不变。最后一层是一个卷积核尺寸为  $3 \times 3$  的卷积层,Padding 为 1,输入为 64 个特征图,输出为 1 张高分辨率的残差图像。

深度为  $D$  的网络拥有  $(2D + 1)^2$  个感受野,相比于 VDSR,本文算法的深度增加到 24 个。感受野增多,实验效果也比 VDSR 好。

### 1.1.2 输出层

由于深度卷积网络不容易收敛,且当深度网络的后几层都是恒等映射时,网络就会退化成浅层网络,残差网络具有可以使深度网络更容易收敛的优点,本文仍然使用残差网络。本文的网络中,输出层是由输入的低分辨率图像  $X_0$  与残差图像  $X_{24}$  相加,最终得到高分辨率图像  $Y = X_0 + X_{24}$ 。由于低分辨率图像  $X_0$  与高分辨率图像  $Y$  的低频成分基本相近,残差图像  $X_{24}$  为低分辨率图像  $X_0$  与高分辨率图像  $Y$  的高频残差,其大部分数值很小,比较稀疏,收敛速度快。

### 1.2 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,SRCNN 与 VDSR 的激活函数都使用了 ReLU 函数,ReLU 函数是修正线性单元,其表达式为

$$\text{ReLU}(x) = \max(x, 0) \quad (1)$$

ReLU 函数将矩阵中所有负数都设为 0,其余的值不变。与 Sigmoid 函数和 Tanh 函数相比,ReLU 函数可以大大地加快随机梯度下降算法的收敛。但它有一个缺点,就是训练很脆弱,很容易死亡,如一个很大的梯度流过一个 ReLU 神经元,参数更新后这个神经元就再也不会对任何数据有激活现象了,那么这个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。例如,当学习率很大时,很有可能网络中 40% 的神经元都死亡了,即在整个的训练过程中神经元都没有激活。本文对 ReLU 函数进行了改进,使用 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:

$$\text{PReLU}(x) = \begin{cases} x, & x > 0 \\ ax, & x \leq 0 \end{cases} \quad (2)$$

可以看出,当  $x > 0$  时,PReLU 函数与 ReLU 函数相同;当  $x \leq 0$  时,ReLU 函数为 0,而 PReLU 函数是一条斜率  $a$  且过原点的直线, $a$  是参数,且其值较小,可以取 0.01 等。为方便与 ReLU 函数作对比,PReLU 函数也可写为:

$$\text{PReLU}(x) = \max(x, 0) + \min(ax, 0) \quad (3)$$

在 24 个卷积层中,每个卷积层中都使用 PReLU 激活函数作为修正单元,使卷积神经网络具有良好的非线性映射,其中第  $i$  层的表达式为:

$$X_i = \max\{W_i * X_{i-1} + B_i, 0\} + \min\{a(W_i * X_{i-1} + B_i), 0\}, 1 \leq i \leq 24 \quad (4)$$

式中, $X_{i-1}$  和  $X_i$  分别为第  $i$  层网络的输入和输出,\* 表示卷积运算; $W_i$  为第  $i$  层的卷积核; $B_i$  为第  $i$  层的偏差; $X_0$  为低分辨率图像输入; $X_{24}$  为网络学习的残差。

### 1.3 学习规则的改进

使用均方误差(MSE)作为损失函数,其表达式为:

$$\text{MSE} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [y(i, j) - f(i, j)]^2 \quad (5)$$

式中: $y(i, j)$  为实际高分辨率图像; $f(i, j)$  为重建后的图像; $M$ 、 $N$  分别为图像的宽度和高度。

网络中采用 Adam 优化器替代传统随机梯度下降(SGD)算法作为误差学习规则。Adam 算法与传统的随机梯度下降不同,随机梯度下降保持单一的学习率,更新所有的权重,学习率在训练过程中并不会改变,而 Adam 通过计算梯度的一阶矩估计和二阶矩估计,在训练过程中动态地调整学习率,可以使每一次迭代的学习率都有一个较好的数值,并使训练质量提高,Adam 具体表达式为:

$$m_t = \frac{\beta_1 m_{t-1} + (1 - \beta_1) g_t}{1 - \beta_1^t}$$

$$v_t = \frac{\beta_2 v_{t-1} + (1 - \beta_2) g_t^2}{1 - \beta_2^t}$$



$$\theta_{t+1} = \theta_t - \frac{\alpha \cdot m_t}{\sqrt{v_t} + \varepsilon} \quad (6)$$

式中:  $g_t$  为  $t$  时间步的梯度;  $m_t$  为对梯度  $g_t$  的一阶矩估计;  $v_t$  为对梯度  $g_t$  的二阶矩估计;  $\beta_1$  和  $\beta_2$  为参数, 通常分别取 0.9 和 0.999;  $\varepsilon$  取  $10^{-8}$  以防止除数为 0;  $\alpha$  为初始学习率。

学习率较大表明学习速度非常快, 但其收敛结果不好, 因此, 不能盲目缩小学习率。本文采用适中的学习率, 既可以保证学习速度, 又可以优化收敛结果, 训练中设置 Adam 的初始学习率为 0.001。

## 2 实验与结果分析

### 2.1 网络训练过程

本算法的实验环境为 Intel Core i7-7700K CPU @ 4.20 GHz, 16 GB 内存, 64 位 Ubuntu 系统, 双路 RTX 2080 Ti GPU 加速训练, 深度学习框架使用 Caffe。

为了提高超分辨图像的效果, 训练样本使用实拍的遥感图像。为保证训练集的数量, 每个遥感图像有交叉地截取子图像, 每相邻 21 像素截取一张  $33 \times 33$  尺寸的子图片, 最终得到了超过 80 000 张尺寸为  $33 \times 33$  的小图像作为训练集。将训练集图像本身作为网络的输出, 训练集图像进行双三次插值缩小后再放大到原尺寸的低分辨率图像作为网络的输入。网络一共进行 40 000 次训练迭代, 累计耗时 40 min, 并且每 500 次迭代进行一次测试, 并取测试过程中 PSNR 最大的网络作为最终的模型。网络训练误差与迭代次数的关系如图 2 所示。

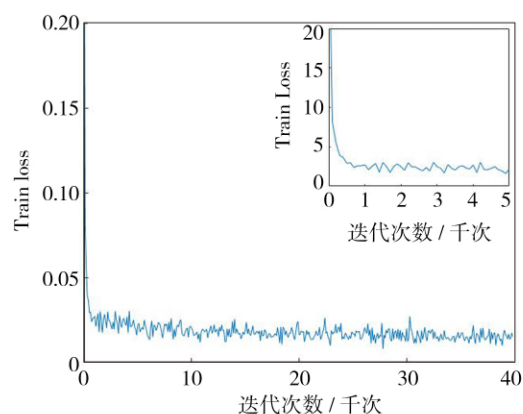


图2 训练误差与迭代次数的关系

Fig. 2 Relationship between training error and iteration times

### 2.2 实验结果与分析

选取 8 张遥感图像进行了测试, 结果如图 3 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 3, 训练与测试图像均使用遥感图像, 迭代次数均取 10 000 次。在此仅列出 4 组实验对比结果, 如图 4 ~ 图 7 所示。



图3 测试遥感图像

Fig. 3 Test remote sensing images

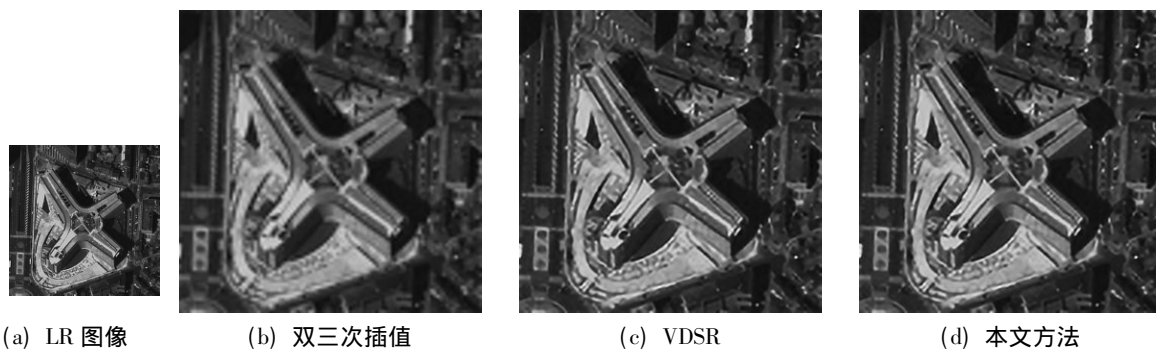


图 4 第一组遥感图像超分辨效果比较

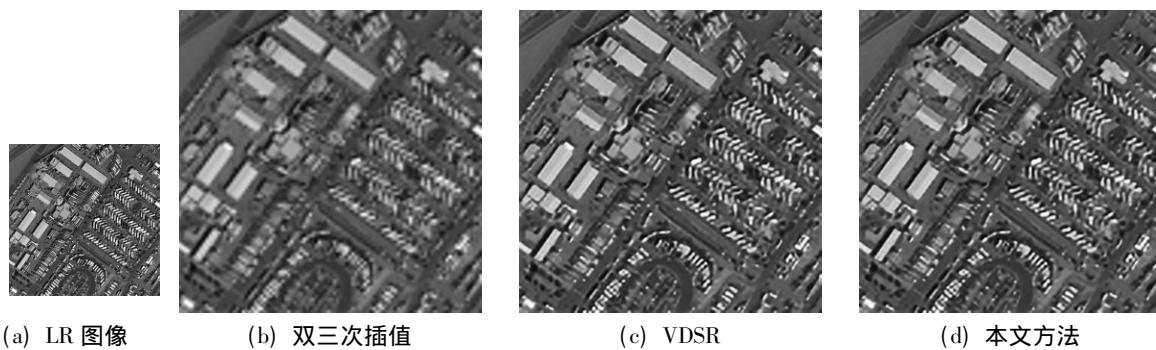
Fig. 4 Comparison of the 1<sup>st</sup> group remote sensing images SR with different algorithms

图 5 第二组遥感图像超分辨效果比较

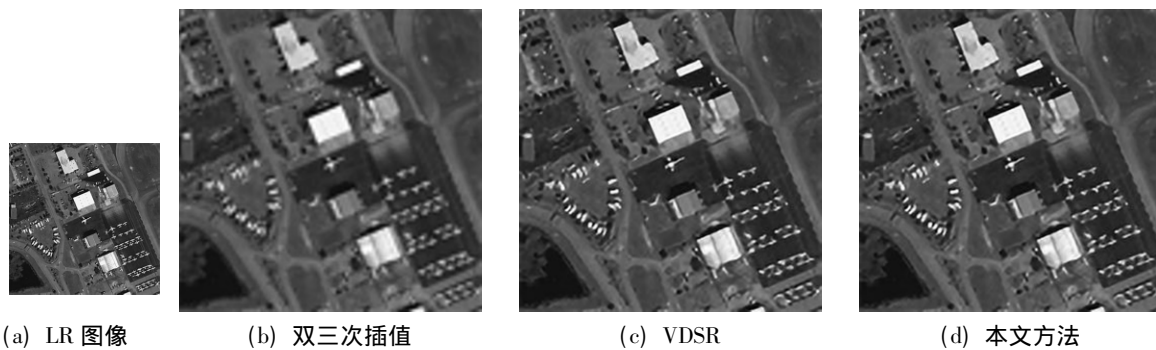
Fig. 5 Comparison of the 2<sup>nd</sup> group remote sensing images SR with different algorithms

图 6 第三组遥感图像超分辨效果比较

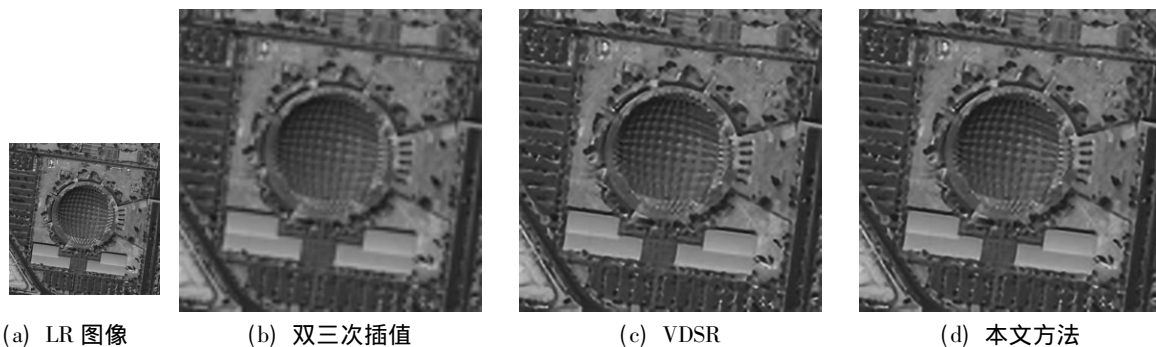
Fig. 6 Comparison of the 3<sup>rd</sup> group remote sensing images SR with different algorithms

图 7 第四组遥感图像超分辨效果比较

Fig. 7 Comparison of the 4<sup>th</sup> group remote sensing images SR with different algorithms

从以上结果可以清楚地看出,与双三次插值和 VDSR 结果相比,本文改进算法重建图像的视觉效果更好,细节信息更丰富。

为了直观地评价各种超分辨图像的效果,本文选用峰值信噪比(PSNR)和结构相似度(SSIM)作为图像

质量评价的指标 ,其定义如式 ( 7) 和式 ( 8) 所示:

$$\text{PSNR} = 10\lg \frac{255^2 M \cdot N}{\sum_{i=1}^M \sum_{j=1}^N (y(i,j) - f(i,j))^2}$$

( 7)

式中:  $y(i,j)$  为实际高分辨率图像;  $f(i,j)$  为重建后的图像。

超分辨图像与原图存在误差 ,其误差越小 ,PSNR 越高 ,说明超分辨结果图像的失真度越低 ,重建效果越好。

$$\text{SSIM} = \frac{4\mu_x\mu_y\sigma_{xy}}{(\mu_x^2 + \mu_y^2)(\sigma_x^2 + \sigma_y^2)}$$

( 8)

式中:  $\mu_x$  和  $\mu_y$  为均值;  $\sigma_x$  和  $\sigma_y$  为方差;  $\sigma_{xy}$  为协方差。SSIM 表示超分辨图像与原图的结构相似程度 ,SSIM 越接近 1 ,说明它们的结构越相似 ,重建效果越好。

本文将改进的算法与双三次插值、VDSR 算法作对比 ,8 张测试遥感图像三种算法的 PSNR 对比如表 1 所示 ,SSIM 对比如表 2 所示 ,三种算法的 PSNR 与 SSIM 趋势对比如图 8 所示。可以看出 ,本文算法无论在主观视觉上 ,还是在客观参数上 ,对图像的超分辨效果都优于其他算法。相比 VDSR 算法 ,PSNR 平均提高约 0.45 dB ,SSIM 平均提高了 0.023。

表 1 8 组测试遥感图像 SR 的 PSNR 对比 / dB

Table 1 PSNR comparison of 8 groups of remote sensing images SR / dB

图像	双三次插值	VDSR	本文方法
第 1 组	20.65	21.94	22.40
第 2 组	16.31	17.37	17.82
第 3 组	20.98	22.41	22.85
第 4 组	23.99	24.96	25.41
第 5 组	24.90	27.25	27.70
第 6 组	19.85	21.95	22.40
第 7 组	22.37	24.75	25.20
第 8 组	22.80	23.82	24.28
平均值	21.48	23.06	23.51

表 2 8 组遥感图像 SR 的 SSIM 对比

Table 2 SSIM comparison of 8 groups of remote sensing images SR

图像	双三次插值	VDSR	本文方法
第 1 组	0.594	0.701	0.729
第 2 组	0.456	0.590	0.630
第 3 组	0.649	0.735	0.760
第 4 组	0.654	0.738	0.763
第 5 组	0.775	0.858	0.872
第 6 组	0.712	0.828	0.845
第 7 组	0.730	0.838	0.854
第 8 组	0.779	0.834	0.850
平均值	0.669	0.765	0.788

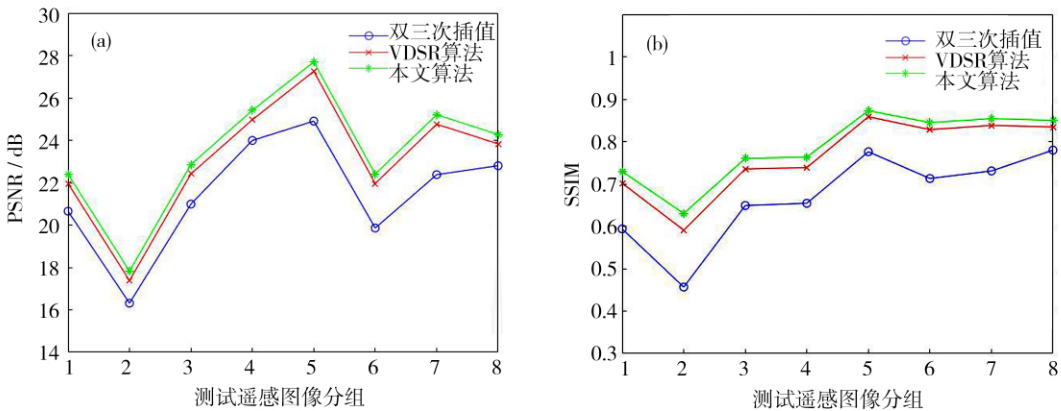


图 8 三种算法下 PSNR ( a) 与 SSIM ( b) 趋势对比

Fig. 8 Trend comparison of PSNR ( a) and SSIM ( b) with three algorithms

### 3 结 论

采用增加网络深度的方法 ,从将卷积层个数扩展到 24 层、改进激活函数、优化网络误差规则等几方面对 VDSR 算法进行了优化 ,并使用遥感图像作为训练集 ,使此卷积神经网络在遥感图像的超分辨效果更好。结果表明 ,本方法可以生成更高质量的超分辨图像 ,图像纹理细节信息更丰富。相较于现有的卷积神经网络超

分辨,其主观视觉效果更好,客观参数 PSNR 提高约 0.45 dB,SSIM 提高了约 0.023。

## 参考文献

- [1] 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.
- [2] 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.
- [3] HARDIE R C, BARNARD K J, ARMSTRONG E E. Joint map registration and high-resolution image estimation using a sequence of undersampled images [J]. IEEE Transactions on Image Processing, 1997, 6(12): 1621–1633.
- [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] 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.
- [6] FREEMAN W T, JONES T R, PASZTOR E C. Example-based super-resolution [J]. IEEE Computer Graphics and Applications, 2002, 22(2): 56–65.
- [7] 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. IEEE, 2017: 5835–5843.
- [8] CHANG S G, CVETKOVIC Z, VETTERLI M. Locally adaptive wavelet-based image interpolation [J]. IEEE Transactions on Image Processing, 2006, 15(6): 1471–1485.
- [9] 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.
- [10] 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.
- [11] 成云凤, 汪伟. 基于医学图像的超分辨率重建算法综述 [J]. 北京生物医学工程, 2019, 38(5): 535–543.
- [12] 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重建 [J]. 计算机应用与软件, 2019, 36(5): 220–224.
- [13] 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.
- [14] 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.
- [15] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network [C]. European Conference on Computer Vision. 2016: 391–405.
- [16] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE Computer Society, 2016, 770–778.
- [17] 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. IEEE Computer Society, 2016: 1646–1653.