

# 基于深度学习的高分辨率卫星遥感影像条带噪声去除

高浩博<sup>1</sup>, 卜桐<sup>3</sup>, 李欣<sup>1</sup>, 陆世东<sup>2</sup>, 钟慧敏<sup>3</sup>, 崔林<sup>3</sup>

1. 武汉大学 遥感信息工程学院, 武汉 430079;

2. 湖北省国土资源研究院, 武汉 430071;

3. 中国资源卫星应用中心, 北京 100094

**摘要:** 受到成像环境、硬件条件等因素的限制, 高分辨率卫星遥感影像上普遍存在条带噪声的现象, 其严重影响了影像的辐射质量和可用性。本文针对传统条带去除方法存在的适应性差、去噪效率低、依靠先验知识等不足, 提出了一种基于深度学习卷积神经网络的条带噪声去除方法。本方法首先利用不同尺度的卷积层进行特征提取, 然后对多尺度的特征图进行特征融合得到去噪底图, 通过残差学习的方法在底图上预测存在的噪声分量, 最后用噪声影像减去条带噪声分量实现噪声的去除。以模拟和真实获取的噪声影像为实验数据, 将本文提出的方法与一些经典的去噪方法进行实验结果对比分析, 实验结果表明本文提出的基于深度学习的条带噪声去除方法能够在保留影像地物细节的情况下, 能以优异的速度达到最高的定量指标和最好的视觉效果, 充分证明了本文方法的优越性。

**关键词:** 高分影像, 深度学习, 条带噪声, 卷积神经网络, 特征融合

**中图分类号:** P2

**引用格式:** 高浩博, 卜桐, 李欣, 陆世东, 钟慧敏, 崔林. 2023. 基于深度学习的高分辨率卫星遥感影像条带噪声去除. 遥感学报, 27(3): 610-622

Gao H B, Bu T, Li X, Lu S D, Zhong H M and Cui L. 2023. Stripe noise removal in high resolution satellite remote sensing images based on deep learning. National Remote Sensing Bulletin, 27(3): 610-622 [DOI: 10.11834/jrs.20221054]

## 1 引言

受卫星 CCD 成像系统性能的限制或影像拍摄过程中外界环境的影响, 不同 CCD 探元在扫描成像时, 容易出现对同一辐射能量响应不一致的现象 (孙斌等, 2014), 导致影像上出现条带噪声。条带噪声的存在会极大地影响卫星遥感影像的目视效果和后续使用, 因此高分辨率卫星遥感影像 (下文简称高分影像) 条带噪声去除是重要的研究内容。目前条带噪声去除方法主要分为四类。一是基于统计的方法, 其中主要包括直方图匹配 (Horn 和 Woodham, 1979; Carfantan 和 Idier, 2010) 和矩匹配 (Fischel, 1984; Rakwatin 等, 2007), 其前提是假设不同卫星传感器探测的地物像素灰度

值具有相同的分布, 利用无噪声影像上统计的结果实现噪声影像上条带噪声的改正, 该方法实现简单、效率高, 但当前提不满足时, 容易出现条带噪声去除不彻底或过度去除的问题。二是基于滤波的方法, 主要包括空间域-频率域转换 (Chen 等, 2003) 和小波变换 (赵变红等, 2012), 其主要是把影像转换到频率域, 对影像进行滤波后, 再把影像恢复到空间域得到去噪后的影像。这种方法虽然可以较为有效地去除影像中的噪声, 但同时也会造成影像信息丢失。三是基于最优化的方法, 其把影像去噪看作为一个病态的优化任务, 通过引入先验知识, 采用正则化方法计算能量函数的最小值得到去噪影像。Shen 和 Zhang (2009) 提出了一种基于 Huber-Markov 模型的最大后验方

收稿日期: 2021-02-20; 预印本: 2021-10-15

基金项目: 国家重点研发计划 (编号: 2018YFA0605500); 湖北省国土资源科研计划“省级国土资源卫星应用示范关键技术研究及系统研制” (编号: 844-11)

第一作者简介: 高浩博, 研究方向为摄影测量与遥感。E-mail: 934741099@qq.com

通信作者简介: 陆世东, 研究方向为自然资源大数据。E-mail: 76235431@qq.com

法实现了影像的去噪。考虑到条带噪声具有明显的单向性，利用条带方向的信息，Bouali 和 Ladjal (2011) 提出了一种用于 MODIS 影像去噪的单向全变分模型 UTV (Unidirectional Variational Model)，并在此基础上，一些改进的 UTV 模型也相继被提出 (Chang 等, 2013; Wang 等, 2016; Song 等, 2018; 王昶等, 2019)。Chang 等通过详细分析条带噪声的结构特征结合影像的先验知识，提出了一种基于低秩的单图像分解模型 LRSID (low-rank-based single-image decomposition model) (Chang 等, 2016)，将原始噪声图像上条带噪声分量进行了分离。基于最优化的方法虽然可以达到较好的去噪效果，但其过度依赖模型输入的先验知识，且模型较为复杂，效率较差，在实际应用中受限。

上述3种方法都具有一定的特点，且在特定的环境下都具有较好的去噪效果，但都具有一定的局限性，容易受到各自前提假设的影响。基于统计的方法中，影像需要满足像素灰度值在同一分布，基于滤波和优化的方法往往依赖于先验知识和人工参数的选择，且去噪效率较差，不能解决海量遥感影像带来的问题和挑战，因此去噪方法需要在效果和速度上进行平衡。而第四种基于深度学习的去噪方法无需任何先验知识和假设条件，同时依赖于 GPU 的计算能力，深度学习卷积神经网络 (CNN) 可以快速而准确的实现条带噪声的去除。Kuang 等 (2017) 提出了一个用于红外影像的单波段红外影像噪声去除网络 (SNRCNN)。Xiao 等 (2018) 提出了一种局部-全局组合的结构模型 (ICSRN) 实现了红外云图上的条带噪声去除。在不破坏光谱空间结构的情况下，HSI-DeNet 网络可以对高光谱影像中每个通道的各种类型噪声进行去除，并且通过修改第一层和最后一层的通道数，可以满足不同任务的需要 (Chang 等, 2019)。

综上所述，在高分影像上进行条带噪声去除具有重要意义，并且在各类去噪方法中，基于深度学习的条带噪声去除方法表现出了较大的潜力，可以对噪声进行快速而有效的去除。因此，本文提出了一种基于深度学习的高分遥感影像条带噪声去除方法。通过端对端的网络学习，利用多尺度特征融合及残差学习的方法实现高分影像上条带噪声的去除。

## 2 原理与方法

### 2.1 条带噪声影像模型

高分卫星为推扫式成像，影像每一列均为 CCD 的同一探元沿轨方向推扫获得，因此影像的同一列可以看作受到同一强度噪声的影响，所以可以把影像某一列存在的条带噪声看作为一个附加的噪声分量，从而得到噪声影像模型，如式 (1) 所示：

$$Y = X + N \quad (1)$$

式中， $Y \in \mathbf{R}^{W \times H \times B}$  表示噪声影像，由  $X \in \mathbf{R}^{W \times H \times B}$  无噪声影像与噪声分量  $N \in \mathbf{R}^{W \times H \times B}$  叠加而成，参数  $W$ 、 $H$ 、 $B$  分别代表影像的列宽、行宽和波段数。去噪网络的任务是恢复出无噪声影像  $X$ ，而根据噪声模型公式可知，只要能够预测出噪声分量  $N$ ，就可以通过噪声影像反推出无噪声影像  $X$ ，因此去噪任务也可以转为对噪声分量的预测任务 (Zhong 等, 2020)。

### 2.2 残差学习

与部分 CNN 去噪网络不同，本文受到残差学习理论的启发 (He 等, 2016)，采用残差学习的策略预测条带噪声分量  $N$ ，通过噪声影像减去条带噪声分量的方式实现噪声的去除。采用残差学习具有以下原因，首先，利用卷积神经网络直接得到去噪影像是较为困难的，而利用网络预测条带噪声分量较为简单，因此可以利用残差学习降低去噪的难度和复杂性。其次通过多层网络学习条带噪声的残差映射要比直接学习去噪影像的未定义的映射要快，大量的研究证明，残差映射的方法可以极大的提高低级图像处理任务的能力 (Zhang 等, 2017)。最后在基于最优化的方法中，也是通过将条带特征利用不同的正则化参数表示，迭代预测出条带噪声分量  $N$ ，从而得到去噪影像，Zhong 等 (2020) 也证明了网络预测条带分量的可行性。同时 3.3 (1) 节中的消融实验也证明了残差学习对于网络去噪的有效性。

### 2.3 数据模拟

与深度学习中场景分类、目标识别等常用任务相比，真实的噪声影像样本和其对应的无噪声标签较难获得，因此需要利用无噪声影像模拟仿真的方法生成噪声影像样本和对应标签。依据 2.1

中所述的卫星成像特性, 本文根据 (Zhong 等, 2020) 中条带噪声的模拟方式, 通过对无噪声影像的每一列添加零均值的高斯白噪声获得噪声影像, 同时考虑到不同探元对于辐射亮度的响应不

同, 每一列添加的噪声强度总体满足标准差为  $\sigma$  的正态分布, 具体模拟流程如图 1 所示, 从图 1 中可以看出, 添加噪声后的影像影响了目视判读, 破坏了影像的视觉效果。

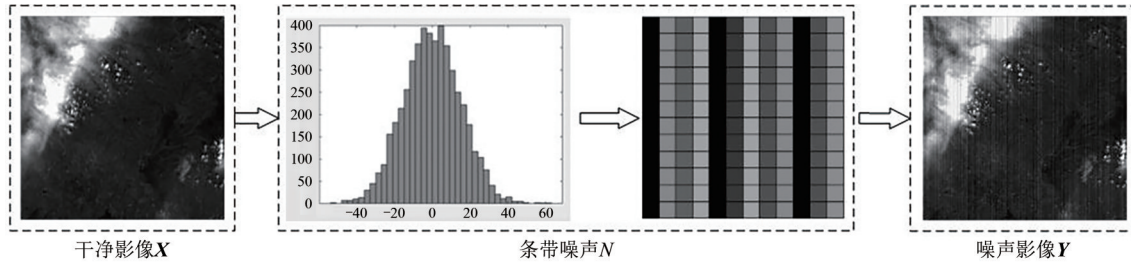


图 1 模拟条带噪声流程

Fig. 1 The process of simulating stripe noise

## 2.4 网络结构

本文提出的网络总体结构如图 2 所示, 采用端对端的方式进行学习, 主要包括特征提取模块, 特征融合模块和影像去噪模块。首先利用多层卷积层对噪声影像进行不同尺度的特征提取, 然后利用反卷积层对不同尺度的特征图进行上采样, 采用逐元素相加的方法融合上采样后的特征图, 随后利用卷积层在融合特征图上对条带噪声分量进行预测, 最后把噪声图像减去条带噪声分量得到最终的去噪影像。网络各层的具体细节如表 1 所示, 其中  $p$  单位为像素。

特征提取模块层数从 Conv1 到 ReLU8, 由 8 个卷积层组成, 其中包括 6 个相同尺度的卷积层、2 个下采样 2 倍的卷积层, 其中卷积层通道数均为 64, 采用 ReLU 函数作为激活函数避免过拟合, 计算公式如式 (2):

$$Y_i(X_i) = f_{\text{ReLU}}(\mathbf{W}_i X_i + \mathbf{B}_i) \quad (2)$$

式中,  $Y_i(X_i)$  为卷积层输出,  $X_i$  为卷积层输入,  $\mathbf{W}_i$  和  $\mathbf{B}_i$  分别为卷积核的权重值和偏差值,  $f_{\text{ReLU}}$  为 ReLU 激活函数。

表 1 网络每层具体实现表

Table 1 Details of our network in each layer

层数	输入大小	卷积核大小	输出大小
Conv1	$p \times p \times 1$	$7 \times 7 \times 1 \times 64$	$p \times p \times 64$
ReLU1	$p \times p \times 64$	—	$p \times p \times 64$
Conv2-5	$p \times p \times 64$	$7 \times 7 \times 64 \times 64$	$p \times p \times 64$
ReLU2-5	$p \times p \times 64$	—	$p \times p \times 64$
Fus1	$p \times p \times 64$	$3 \times 3 \times 64 \times 64$	$p \times p \times 64$
ReLU6	$p \times p \times 64$	—	$p \times p \times 64$
Fus2	$p \times p \times 64$	$3 \times 3 \times 64 \times 64$	$p/2 \times p/2 \times 64$
ReLU7	$p/2 \times p/2 \times 64$	—	$p/2 \times p/2 \times 64$
Fus3	$p/2 \times p/2 \times 64$	$3 \times 3 \times 64 \times 64$	$p/4 \times p/4 \times 64$
ReLU8	$p/4 \times p/4 \times 64$	—	$p/4 \times p/4 \times 64$
Fus2_1	$p/2 \times p/2 \times 64$	$4 \times 4 \times 64 \times 64$	$p \times p \times 64$
ReLU9	$p \times p \times 64$	—	$p \times p \times 64$
Fus3_1	$p/4 \times p/4 \times 64$	$4 \times 4 \times 64 \times 64$	$p/2 \times p/2 \times 64$
ReLU10	$p/2 \times p/2 \times 64$	—	$p/2 \times p/2 \times 64$
Fus3_2	$p/2 \times p/2 \times 64$	$4 \times 4 \times 64 \times 64$	$p \times p \times 64$
ReLU11	$p \times p \times 64$	—	$p \times p \times 64$
Fus4	$p \times p \times 64 \times 3$	—	$p \times p \times 64$
Fus5	$p \times p \times 64$	$3 \times 3 \times 64 \times 64$	$p \times p \times 64$
ReLU12	$p \times p \times 64$	—	$p \times p \times 64$
Fus6	$p \times p \times 64$	$3 \times 3 \times 64 \times 64$	$p \times p \times 64$

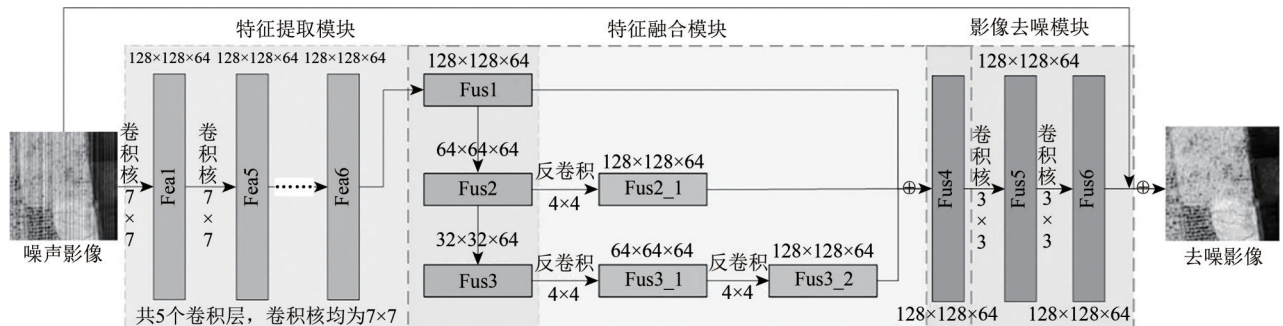


图 2 本文网络结构图

Fig. 2 The structure of our network

受到宽度参考网络 (Wide Inference Network) 的启发 (Liu 和 Fang, 2017), 为了获得更多的特征信息, 网络在前 5 层卷积层中采用  $7 \times 7$  卷积核进行特征的提取, 而在不同尺度的采样阶段采用  $3 \times 3$  卷积核, 从而减少内存消耗, 3.3 节 (1) 中的消融实验验证了大卷积核对于图像去噪任务的有效性。考虑到网络输入和输出的影像大小应保持一致, 故卷积层的填充 (Padding) 需随着卷积核的大小发生变化, 同时去除了池化层。

特征融合模块层数从 Fus2\_1 到 Fus4, , 主要由 3 个反卷积层和 1 个特征融合层组成, 各个卷积层的通道数均为 64。网络通过特征提取模块共得到了 3 个不同尺度的特征图, 然后利用反卷积层得到 3 个尺度相同大小和相同通道数的特征图。同时为了保证反卷积层的输入与输出大小相同, 卷积核大小设置为  $4 \times 4$ 、步长 (stride) 为 2。最后利用逐元素相加的方法对不同尺度的相同大小的特征图进行信息融合, 得到最终的去噪底图。具体公式如式 (3) 所示:

$$F_{ij}^c = X_{ij}^c + f_{\text{ReLU}}(W_y Y_{ij}^c + B_y) + f_{\text{ReLU}}(W_{z1} f_{\text{ReLU}}(W_{z2} Z_{ij}^c + B_{z2}) + B_{z1}) \quad (3)$$

式中,  $F_{i,j}^c$  为融合底图上第  $c$  通道第  $i, j$  处的值,  $X_{i,j}^c$  为原始尺度特征图上第  $c$  通道第  $i, j$  处的值,  $Y_{i,j}^c$  为缩小 2 倍尺度特征图上第  $c$  通道第  $i, j$  处的

值,  $Z_{i,j}^c$  为缩小 4 倍尺度特征图上第  $c$  通道第  $i, j$  处的值,  $W_y, W_{z1}, W_{z2}$  为不同尺度下反卷积层的权重值,  $B_y, B_{z1}, B_{z2}$  为不同尺度下反卷积层的偏差值,  $f_{\text{ReLU}}$  为 ReLU 激活函数。3.3 节中的消融实验证明了不同尺度特征融合对于网络去噪的有效性。

影像去噪模块层数从 Fus5 到 Fus6, 包含两个卷积层来实现融合底图上条带噪声分量的预测, 受到超分辨率网络在影像恢复重建上的启发 (Kim 等, 2016), 采用  $3 \times 3$  的小卷积核能够更好的预测条带噪声分量, 最后噪声图像减去预测的噪声分量就可得到最终的去噪影像。

### 3 实验结果与分析

#### 3.1 实验数据

利用 GF-1、GF-2 共 16 景影像模拟训练数据, 根据 2.3 节叙述的方法, 分别对影像添加标准差为 5、15、30、50 的高斯白噪声, 考虑到网络输入, 把噪声影像和其对应的标签裁剪为  $128 \times 128$  像素大小的影像块, 每个噪声强度生成 4000 个样本对, 因为本次实验为盲去噪, 所以把不同噪声强度的样本进行混合, 共计 16000 个样本对用于训练, 不同噪声强度下的部分样本对示例如图 3 所示。

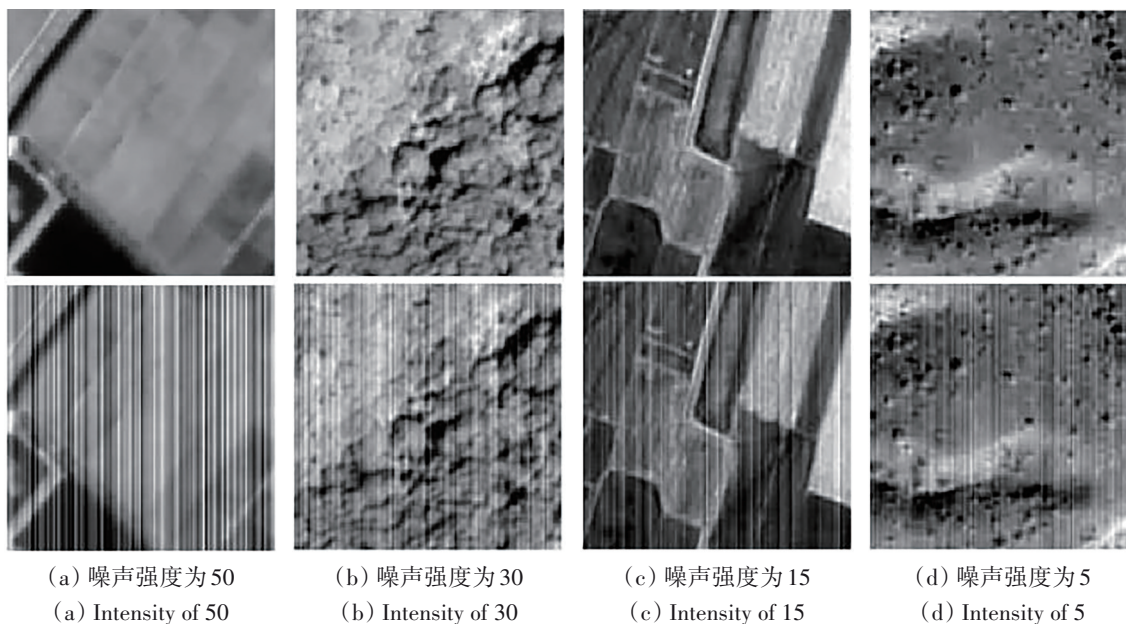


图3 不同噪声强度样本图

Fig. 3 Sample images of different noise identity

### 3.2 实验方法

在网络训练过程中，网络输入均为单波段影像，16000个影像样本对按照9:1的方式划分为训练集和测试集。噪声影像的灰度值在预处理阶段被归一化到[0, 1]，这样可以使影像灰度值具有相同的分布，从而加快网络的训练速度。在网络后向传播过程中，使用Adam作为网络优化器，训练周期(epoch)为80，学习率初始值为0.002，在70个周期后减小为0.0002，输入批大小(Batch Size)为96。同时考虑到影像去噪可以看作影像上每一个像素灰度值的预测问题，即尽量得到与无噪声影像对应位置处相同的灰度值，而均方误差(MSE)函数通过计算去噪影像与无噪声影像的方差，可以很好的描述两者之间的差异，因此采用MSE函数作为损失函数，其具体计算公式如式(4)：

$$MSE = \frac{1}{n \times m} \sum_{i=1}^n \sum_{j=1}^m (y_{ij} - \bar{y}_{ij})^2 \quad (4)$$

式中， $n$ ， $m$ 为影像的长和宽， $y_{i,j}$ 为无噪声影像上(i, j)位置处的灰度值， $\bar{y}_{i,j}$ 为去噪影像上(i, j)位置处的灰度值。本文提出的网络基于Pytorch库实现，训练所使用的GPU型号为NVIDIA 1080ti。

实验选择了一些去噪方法进行对比，其中包括基于统计的方法SLD (Statistical Linear Destriping) (Carfantan 和 Idier, 2010)，基于滤波的方法WFAF (the wavelet-Fourier adaptive filter) (Pande-Chhetri 和 Abd-Elrahman, 2011)，基于最优化的方法LRSID (low-rank-based single-image decomposition model) (Chang等, 2016)和基于深度学习的方法(ICSRN) (Xiao等, 2018)，传统算法均通过matlab实现，深度学习方法基于Pytorch实现，所有的参数被设置为默认值，并按照相关文献的建议进行微调。

实验采用主观目视判断和定量指标评价相结合的方式对去噪效果进行综合评价。采用峰值信噪比PSNR (peak signal-to-noise ratio)和结构相似性SSIM (structure similarity)指标 (Wang等, 2004)对模拟的噪声影像进行定量评价。而在实际获取的噪声影像中，由于没有实际参考数据，利用两个无参考评价指标：反向变异系数ICV (inverse coefficient of variation)，平均相对偏差MRD (mean relative deviation)进行评价 (Shen和

Zhang, 2009)。ICV、MRD指标分别是在两个20×20、30×30的均匀地物区域计算得到的。一般来说，较高的PSNR、SSIM、ICV和较低的MRD代表着较好的去噪效果。

### 3.3 实验结果

(1) 消融实验。根据2.4节网络设计中考虑的因素，本文设计了一些消融实验来证明网络的有效性，所有的网络均在3.1节介绍的模拟噪声数据集上训练得到，验证数据为包含4种不同噪声强度的共1600张模拟噪声影像，定量验证指标为PSNR和SSIM，结果为1600张影像的平均值。除特别说明的参数不同外，其余训练结构和超参数都相同，具体的实验结果如表2所示。

表2 消融实验定量指标结果

Table 2 Quantitative evaluation results for the ablation experiment

数据	残差学习	卷积核大小	特征提取层数	特征融合	PSNR	SSIM
模拟噪声影像	√	7×7	5	√	<b>48.7283</b>	<b>0.9954</b>
	√	7×7	4	√	45.8115	0.9903
	√	7×7	6	√	46.0421	0.9910
	√	5×5	5	√	47.7294	0.9941
	√	3×3	5	√	46.8788	0.9927
	√	7×7	5	×	45.5947	0.9897
	×	7×7	5	√	41.3391	0.9733

注：粗体下划线标记的为最佳去噪结果。

从表2中可以看出，在特征提取模块，考虑到网络训练的效率 and 参数复杂度，实验选用了3组卷积核3×3，5×5，7×7进行比较，实验结果表明正如2.4节中说明的，当网络的卷积核越大时，其去噪效果越好，证明了大卷积核对于去噪的有效性。同时也比较了不同卷积层数对于去噪效果的影响，实验表明当特征提取层数为5时，网络去噪效果最佳。

为了验证本文提出的特征融合模块的作用，把反卷积层和下采样层均换成了卷积核为3×3、步长为1、通道数为64的卷积层，从而保证网络具有相同的层数，实验结果表明，该模块能够有效地提高网络的去噪性能。

实验也证明了残差学习在条带噪声去除中的优越性，比较了两种不同的网络结构，一种为直接预测去噪影像，一种为预测条带噪声分量，即

本文运用的策略。从表2中可以看出，直接预测去噪影像的定量指标与预测条带噪声分量相比具有较大的差距，证明了该预测方法的困难性。因此采用残差学习的策略可以大大的提升网络的去噪性能，提高网络收敛速度。

(2) 基于模拟影像的实验结果。利用不同方法对不同噪声强度的模拟影像的去噪结果如表3所示，用于定量评价的指标为PSNR、SSIM，结果为1600张测试样本的平均值，其中粗体下划线标记

的为最佳去噪结果，粗体标记的为第二好的去噪结果。从表中可以看出，本文的方法在不同的噪声强度中均表现出稳定的性能，并且在条带噪声严重的情况下取得了最好的去噪效果，而SLD、WFAF、LRSID3种传统方法的去噪效果随着噪声强度的增大逐渐变差，ICSRN虽然在高强度噪声影像上仍具有较好的性能，但去噪效果不如本文方法。

表3 模拟实验定量指标结果

Table 3 Quantitative evaluation results for the simulated experiment

噪声强度	指标	噪声影像	SLD	WFAF	LRSID	ICSRN	本文方法
$\sigma=5$	PSNR	46.2213	51.0048	44.529	41.7781	<u>51.3083</u>	<u>51.1719</u>
	SSIM	0.9914	<b>0.9973</b>	0.9845	0.9765	<b>0.9976</b>	<b>0.9976</b>
$\sigma=15$	PSNR	36.6836	<b>46.0334</b>	42.5235	40.6307	45.5371	<u>47.2636</u>
	SSIM	0.9313	<b>0.9924</b>	0.9789	0.9723	0.9915	<u>0.9941</u>
$\sigma=30$	PSNR	30.6309	41.3019	40.5494	39.3575	<b>44.4468</b>	<u>47.8046</u>
	SSIM	0.7829	0.977	0.9722	0.965	<b>0.9888</b>	<u>0.9945</u>
$\sigma=50$	PSNR	26.2737	37.2528	38.2532	37.6255	<b>45.5178</b>	<u>48.227</u>
	SSIM	0.5908	0.9441	0.9575	0.9511	<b>0.9904</b>	<u>0.9939</u>

注：粗体下划线标记的为最佳去噪结果，粗体标记的为第二好结果。

图4、图5、图6、图7展示了不同噪声强度下，不同方法的去噪结果图，图下方为定量指标值（PSNR、SSIM），从图4（c）、5（c）、6（c）、7（c）中可以看出，SLD在噪声强度较低时具有较好的去噪效果，但当噪声强度逐渐升高时，去噪效果变差，尤其是在水体等地物均匀处，例如在图6（c）、图7（c）中均出现了较为明显的条纹，说明了SLD方法在高强度噪声去除中的局限性。WFAF在低强度的噪声影像上容易引入新的条带噪声，如图4（d）中出现了明显的暗纹，而在高强度的噪声图像上的地物均匀处去噪效果不好，从而造成整体的指标评价较差。LRSID在不同强度的噪声影像中，目视均具有较好的效果，但其定量指标较差，这是因为LRSID会使影像变得更加平滑，对影像的原始信息破坏较为严重，从图中可以看出，LRSID虽然去除了噪声，同时也去除了影像的细节信息，使影像变得模糊。与前几种方法相比，ICSRN在不同噪声强度下均具有较好的去噪效果，但与本文提出的方法相比，如图6（f）、图7（f）所示，在高强度噪声情况下，ICSRN在水体处效果较差，出现了明显的亮条纹。而本文提

出的方法在不同噪声强度下均具有较高的定量指标，且目视效果良好，无明显的细节信息丢失和残留条带。综上所述，本文提出的方法在目视效果和定量指标上均优于其他方法。

(3) 基于实际获取的噪声影像实验结果。为了进一步验证本方法的性能，选择在实际获取的高分影像上进行实验，同时为了验证网络的泛化能力，也选取了训练样本之外的卫星数据进行测试，包括1景GF-1B影像，2景GF-2影像，从中裁剪了600×600像素大小的子图像用于噪声去除，影像的灰度值在预处理阶段线性归一化为[0-1]。GF-1B的实际图像如图8（a）所示，GF-2影像如图9（a）、图10（a）所示，实际影像均存在随机的竖状条带噪声，由于实际获取的影像并不存在无噪声的参考值，每张影像分别选取两块20×20，30×30像素的地物均匀区域用来计算ICV、MRD评估去噪影像的质量。考虑到去噪中内存大小的限制，无法把整个影像一次性输入，需要对影像进行分块。为避免分块去噪产生的边界效应，影像每块之间各有30个像素的重叠区，重叠大小为经验值，图11说明了重叠区对于消除边界效应的有效性。

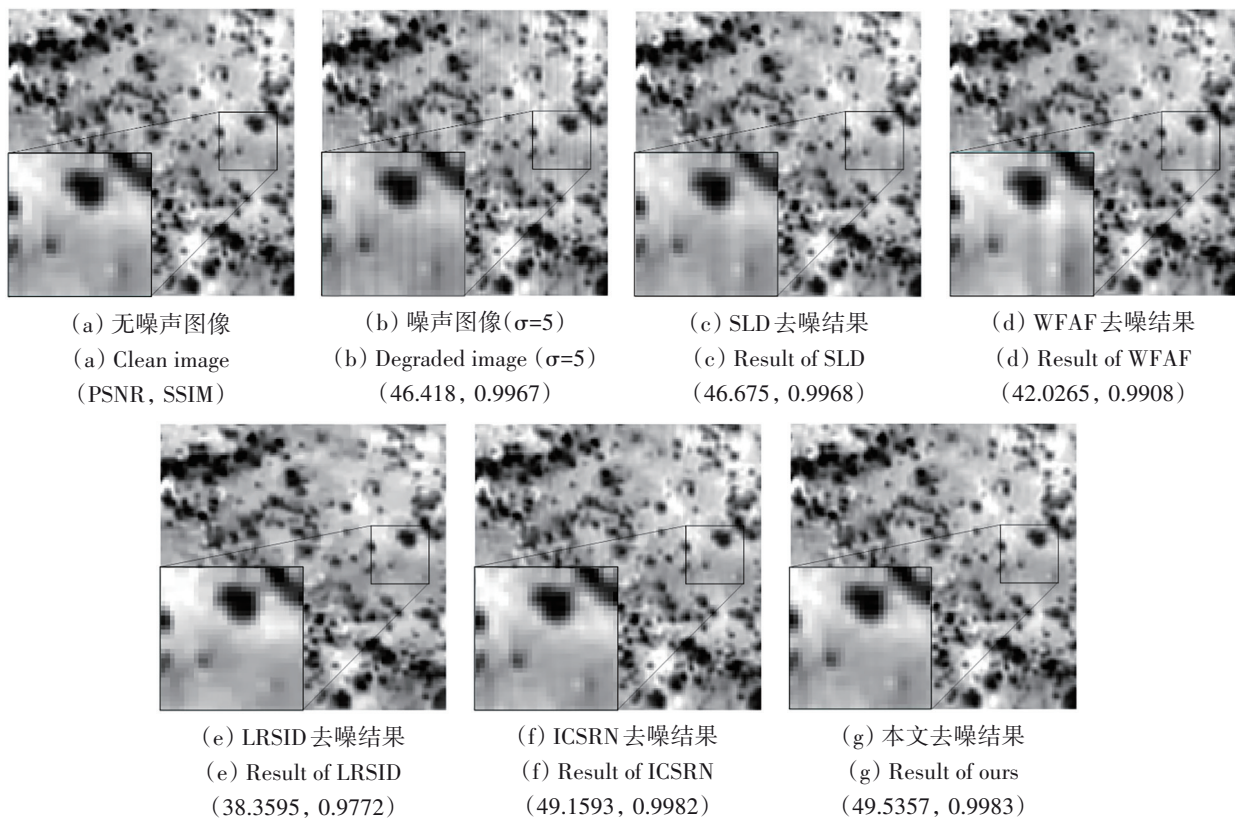


图4 标准差为5的模拟影像去噪结果

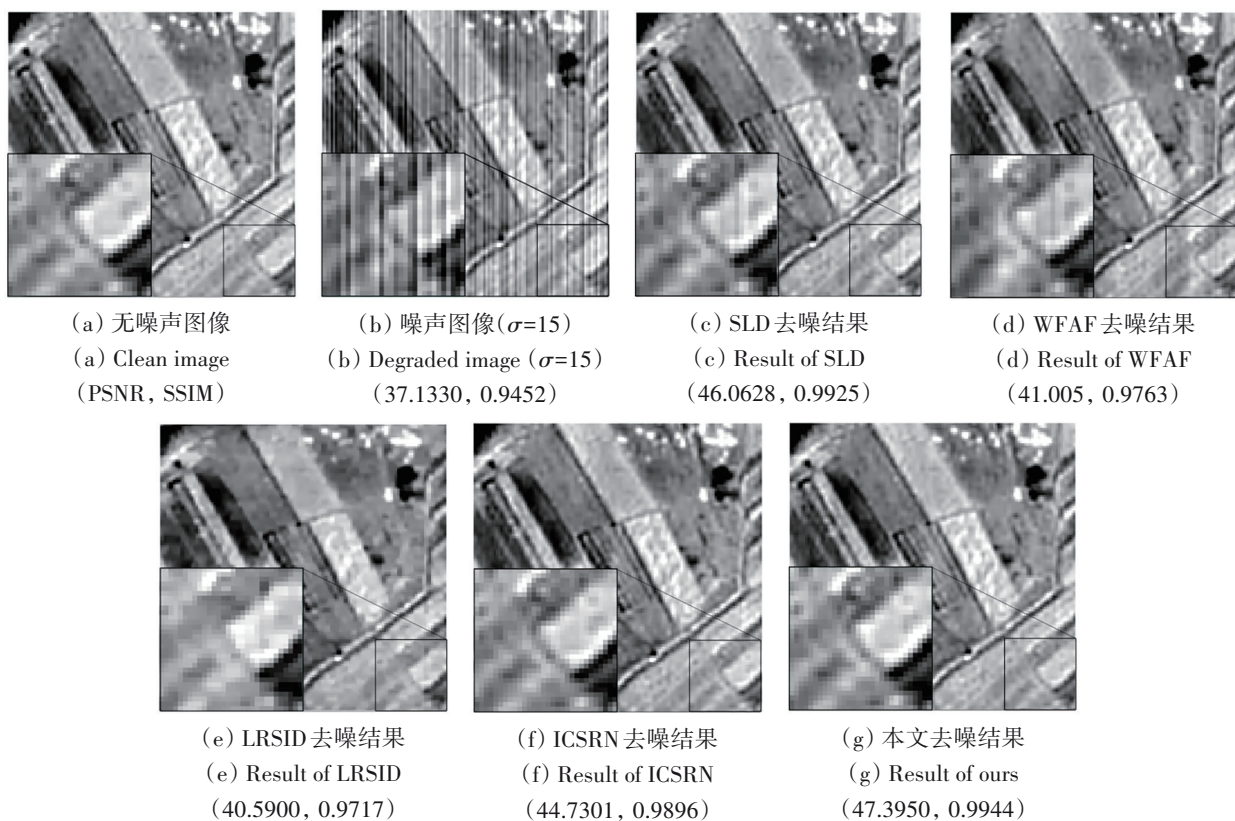
Fig. 4 Denoising results of simulated image with  $\sigma=5$ 

图5 标准差为15的模拟影像去噪结果

Fig. 5 Denoising results of simulated image with  $\sigma=15$

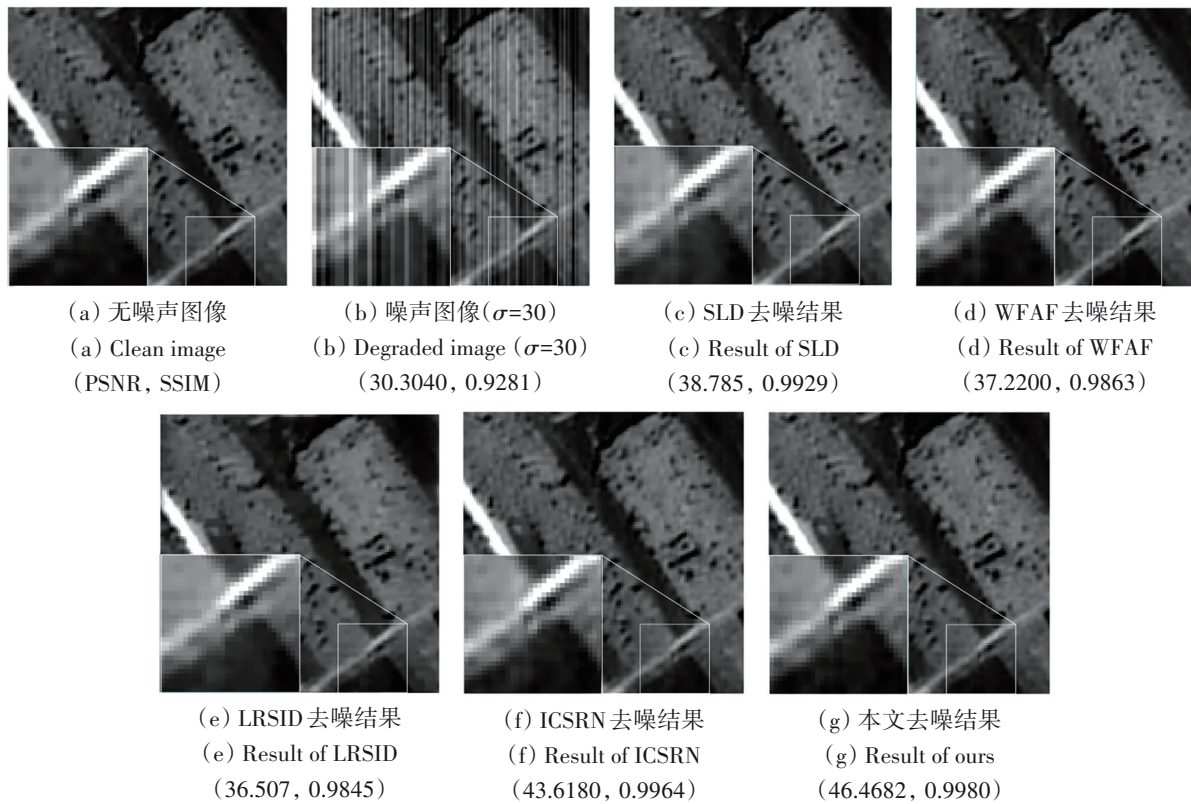


图6 标准差为30的模拟影像去噪结果

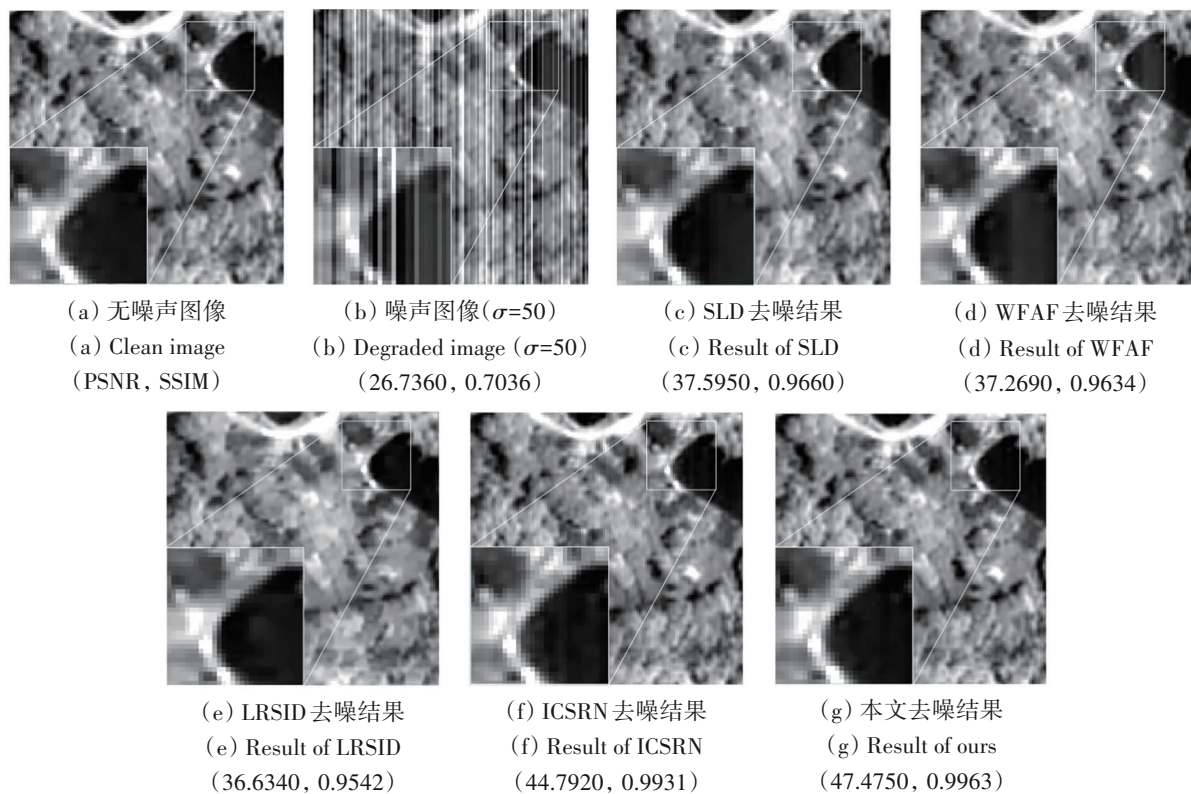
Fig. 6 Denoising results of simulated image with  $\sigma=30$ 

图7 标准差为50的模拟影像去噪结果

Fig. 7 Denoising results of simulated image with  $\sigma=50$



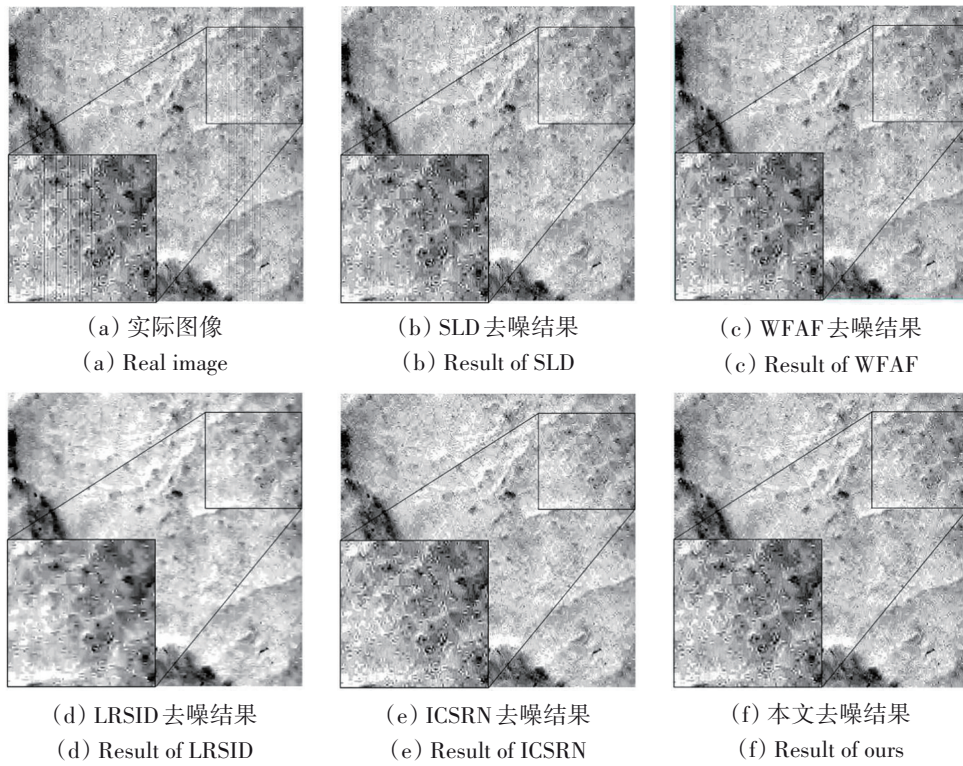


图8 GF-1B影像去噪结果

Fig. 8 Result of GF-1B

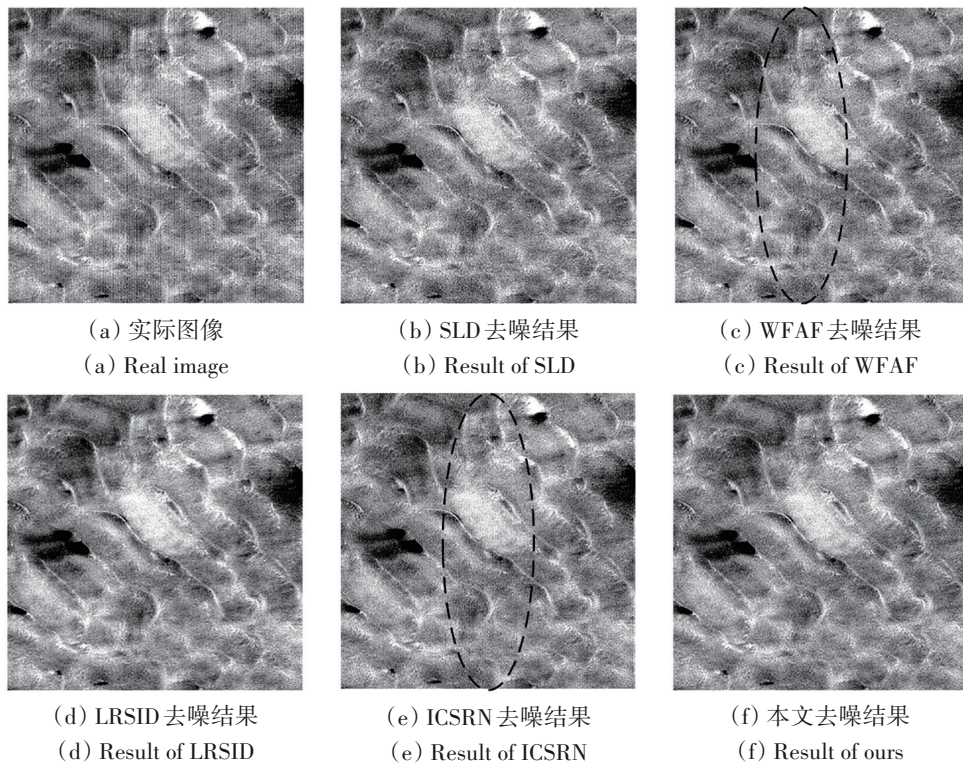


图9 GF-2影像去噪结果

Fig. 9 The first denoising result of GF-2

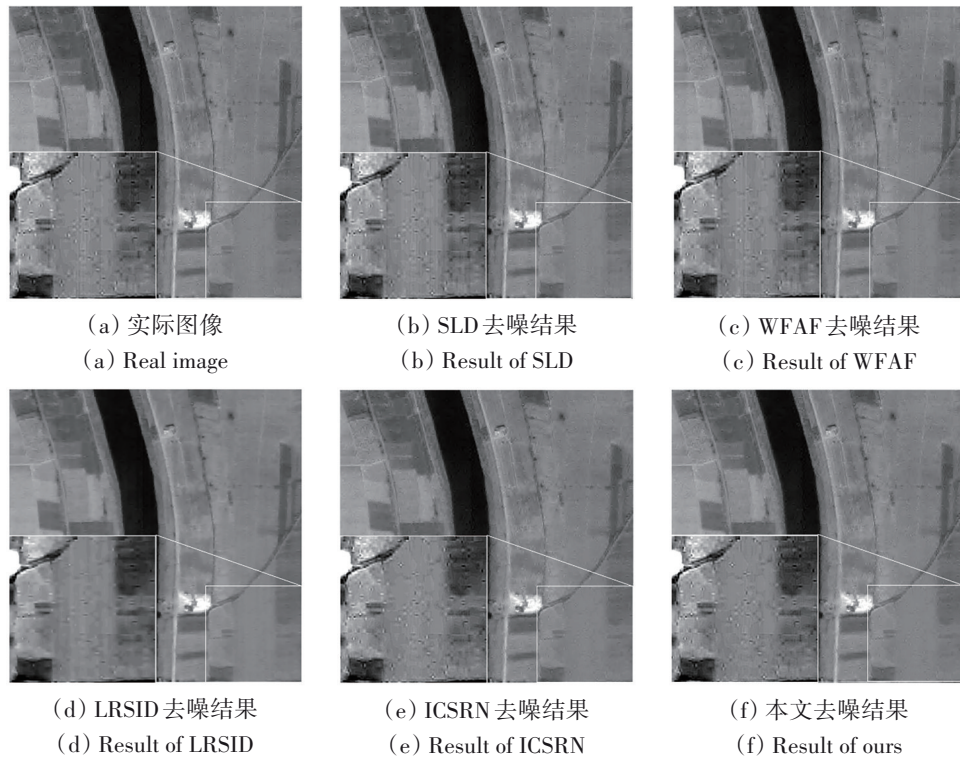


图 10 GF-2 第二张影像去噪结果  
Fig. 10 The second denoising result of GF-2

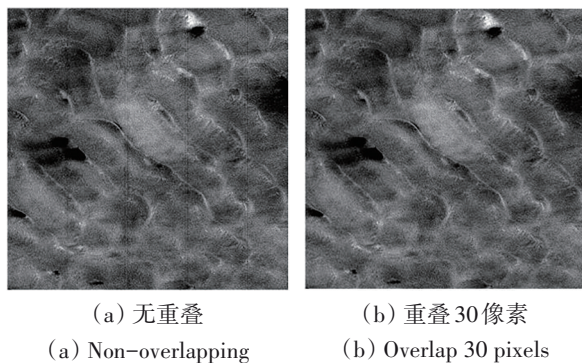


图 11 有无重叠范围的去噪结果  
Fig. 11 Results without or with overlapping range

从图 8、图 9、图 10 中 (a), (b) 子图可以看出, SLD 和 WFAF 方法并不能去除影像中所有的条带噪声, 具有一定的残留, 但其对影像的细节信息的破坏较小, 因此具有较低的 MRD 值。而 LRSID 方法虽然去除了所有的噪声, 目视效果较好, 但与模拟噪声影像去噪中存在的问题一样, 该方法使影像变得平滑和模糊, 破坏了影像原本的地物信息, 如图 8 (b) 中的地物均匀区域所示。而 ICSRN 方法在图 8 (e)、图 10 (e) 中去噪效果较好, 条带噪声没有明显的残留, 但在图 9 (e) 中黄框所标记的区域内, 可以明显的看出残留的

噪声。而本文的方法在 3 幅影像中均具有较好的去噪效果, 不仅能够去除影像中存在的条带噪声, 而且还能最大限度的保留影像自身的细节。表 4 去噪结果的定量指标表明, 本文方法获得了较高的 ICV 和最低的 MRD 指标, 而其中 LRSID 最大的原因在于, 该方法使影像变得过度平滑, 减少了选取范围内灰度值的方差, 增大了 ICV 的值, 造成了目视效果差但指标较高的现象, 从 MRD 也可以看出, LRSID 对于影像原始信息的破坏是最大的。

### 3.4 效率测试

为了比较本文提出的方法与其他去噪方法的计算效率, 我们在 3.2 节描述的硬件环境下对不同强度的 1600 张影像进行了测试, 不同方法的一张影像平均去噪时间如表 5 所示, 单位为秒。从表中可以看出, LRSID 的方法耗时最长, 而本文提出的方法能够在保持优异去噪速度的同时, 达到了最高的去噪指标, 具有最好的去噪性能。同时, 本文网络的去噪时间并不会随着影像大小的增大而迅速增加, 充分说明本网络在实际大尺寸遥感影像去噪任务中更具有实用价值。

表4 真实影像定量评估结果

Table 4 Quantitative evaluation results for the real images

指标	实际噪声图像	SLD	WFAF	LRSID	ICSRN	Our
ICV	295.4542	296.1574	296.4846	<u>467.3558</u>	309.8651	317.5125
MRD/%	—	0.02289	<b>0.01828</b>	0.06260	0.03292	<b><u>0.01621</u></b>

注: 粗体下划线标记的为最佳去噪结果, 粗体标记的为第二好结果。

表5 不同去噪方法的平均运行时间

Table 5 Average run times for the different destriping methods

图像大小/pixel	SLD	WFAF	LRSID	ICSRN	Our
128×128	0.014	0.088	1.645	<b>0.011</b>	<b><u>0.012</u></b>
256×256	0.035	0.125	4.049	<b>0.020</b>	<b><u>0.023</u></b>

注: 粗体下划线标记的为最佳去噪时间, 粗体标记的结果次之。

## 4 结论

本文以卷积神经网络为基础, 提出了一种基于多尺度特征融合的卷积神经网络去噪方法, 该方法结合条带噪声的特点, 采用残差学习的策略预测噪声影像上的条带噪声分量, 通过噪声影像减去条带分量实现高分辨率卫星遥感影像上条带噪声的去除。研究表明:

(1) SLD、WFAF、LRSID传统方法自适应能力较差。在模拟样本上, 去噪效果随着噪声强度的增大而逐渐变差。而在真实影像上, 这些方法会造成影像细节的丢失或者条带噪声的残留。

(2) 基于深度学习的去噪方法能够充分学习到不同强度噪声的特征, 能够自适应的对不同强度的噪声进行去除, 因此两种方法在模拟和真实的影像上均达到了较好的目视效果和定量指标。

(3) 融合不同尺度的特征图能够进一步提高网络的特性学习能力; 采用残差学习的策略能够大大提高网络的训练精度和速度。

(4) 得益于GPU的强大矩阵计算能力, 基于深度学习的条带噪声去除方法能够更高效的对条带噪声进行去除, 速度较SLD、WFAF、LRSID传统方法有明显的提升。

本文提出的方法还有一定的改进空间, 后续考虑采用跳跃连接、增加批归一化层等来优化网络的训练速度和提高网络的去噪性能, 同时进一步研究网络的迁移能力, 尝试将网络框架扩展到其他类型的遥感影像中, 例如航空影像和高光谱影像。

**志 谢** 本文所用高分卫星遥感数据由中国资源卫星应用中心提供, 传统去噪方法的源代码由昌毅博士提供, 在此表示衷心感谢!

## 参考文献(References)

- Bouali M and Ladjal S. 2011. Toward optimal destriping of MODIS data using a unidirectional variational model. *IEEE Transactions on Geoscience and Remote Sensing*, 49(8): 2924-2935 [DOI: 10.1109/TGRS.2011.2119399]
- Carfantan H and Idier J. 2010. Statistical linear destriping of satellite-based pushbroom-type images. *IEEE Transactions on Geoscience and Remote Sensing*, 48(4): 1860-1871 [DOI: 10.1109/TGRS.2009.2033587]
- Chang Y, Fang H Z, Yan L X and Liu H. 2013. Robust destriping method with unidirectional total variation and framelet regularization. *Optics Express*, 21(20): 23307-23323 [DOI: 10.1364/OE.21.023307]
- Chang Y, Yan L X, Fang H Z, Zhong S and Liao W S. 2019. HSI-DeNet: hyperspectral image restoration via convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2): 667-682 [DOI: 10.1109/TGRS.2018.2859203]
- Chang Y, Yan L X, Tao W and Zhong S. 2016. Remote sensing image stripe noise removal: from image decomposition perspective. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12): 7018-7031 [DOI: 10.1109/TGRS.2016.2594080]
- Chen J S, Shao Y, Guo H D, Wang W M and Zhu B Q. 2003. Destriping CMODIS data by power filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9): 2119-2124 [DOI: 10.1109/TGRS.2003.817206]
- Fischel D. 1984. Validation of the thematic mapper radiometric and geometric correction algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, GE-22(3): 237-242 [DOI: 10.1109/TGRS.1984.350616]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Horn B K P and Woodham R J. 1979. Destriping LANDSAT MSS images by histogram modification. *Computer Graphics and Image Processing*, 10(1): 69-83 [DOI: 10.1016/0146-664X(79)90035-2]
- Kim J, Lee J K and Lee K M. 2016. Accurate image super-resolution using very deep convolutional networks//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA:

- IEEE: 1646-1654 [DOI: 10.1109/CVPR.2016.182]
- Kuang X D, Sui X B, Chen Q and Gu G H. 2017. Single infrared image stripe noise removal using deep convolutional networks. *IEEE Photonics Journal*, 9(4): 3900913 [DOI: 10.1109/JPHOT.2017.2717948]
- Liu P and Fang R G. 2017. Wide inference network for image denoising. CoRR, abs/1707.05414.
- Pande-Chhetri R and Abd-Elrahman A. 2011. De-stripping hyperspectral imagery using wavelet transform and adaptive frequency domain filtering. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(5): 620-636 [DOI: 10.1016/j.isprsjprs.2011.04.003]
- Rakwatin P, Takeuchi W and Yasuoka Y. 2007. Stripe noise reduction in MODIS data by combining histogram matching with facet filter. *IEEE Transactions on Geoscience and Remote Sensing*, 45(6): 1844-1856 [DOI: 10.1109/TGRS.2007.895841]
- Shen H F and Zhang L P. 2009. A MAP-based algorithm for destripping and inpainting of remotely sensed images. *IEEE Transactions on Geoscience and Remote Sensing*, 47(5): 1492-1502 [DOI: 10.1109/TGRS.2008.2005780]
- Song Q, Wang Y H, Yan X Y and Gu H G. 2018. Remote sensing images stripe noise removal by double sparse regulation and region separation. *Remote Sensing*, 10(7): 998 [DOI: 10.3390/rs10070998]
- Sun B, Li J L, Zhang X X and Ren J Y. 2014. Interleaving assembly of TDICCDs on 600 mm focal plane. *Optics and Precision Engineering*, 22(11): 2908-2913 (孙斌, 李景林, 张星祥, 任建岳. 2014. 600 mm 长焦平面时间延迟积分 CCD 的交错拼接. *光学精密工程*, 22(11): 2908-2913) [DOI: 10.3788/OPE.20142211.2908]
- Wang C, Wang X and Ji S. 2019. Stripe noise removal of remote images based on variation. *Journal of Xi'an Jiaotong University*, 53(3): 143-149 (王昶, 王旭, 纪松. 2019. 采用变分法的遥感影像条带噪声去除. *西安交通大学学报*, 53(3): 143-149) [DOI: 10.7652/xjtub201903020]
- Wang M, Zheng X H, Pan J and Wang B. 2016. Unidirectional total variation destripping using difference curvature in MODIS emissive bands. *Infrared Physics and Technology*, 75: 1-11 [DOI: 10.1016/j.infrared.2015.12.004]
- Wang Z, Bovik A C, Sheikh H R and Simoncelli E P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600-612 [DOI: 10.1109/TIP.2003.819861]
- Xiao P F, Guo Y C and Zhuang P X. 2018. Removing stripe noise from infrared cloud images via deep convolutional networks. *IEEE Photonics Journal*, 10(4): 7801114 [DOI: 10.1109/JPHOT.2018.2854303]
- Zhang K, Zuo W M, Chen Y J, Meng D Y and Zhang L. 2017. Beyond a gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7): 3142-3155 [DOI: 10.1109/TIP.2017.2662206]
- Zhao B H, He B, Yang L H, Tao M H and Ren J Y. 2012. Destripping for TDI-CCD remote sensing images. *Chinese Journal of Space Science*, 32(2): 298-304 (赵变红, 何斌, 杨利红, 陶明慧, 任建岳. 2012. TDI-CCD 遥感图像条带噪声的消除. *空间科学学报*, 32(2): 298-304) [DOI: 10.11728/cjss2012.02.298]
- Zhong Y F, Li W Q, Wang X Y, Jin S Y and Zhang L P. 2020. Satellite-ground integrated destripping network: a new perspective for EO-1 Hyperion and Chinese hyperspectral satellite datasets. *Remote Sensing of Environment*, 237: 111416 [DOI: 10.1016/j.rse.2019.111416]

## Stripe noise removal in high resolution satellite remote sensing images based on deep learning

GAO Haobo<sup>1</sup>, BU Tong<sup>3</sup>, LI Xin<sup>1</sup>, LU Shidong<sup>2</sup>, ZHONG Huimin<sup>3</sup>, CUI Lin<sup>3</sup>

1. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China;

2. Hubei Provincial Research Institute of Land and Resources, Wuhan 430071, China;

3. China Centre for Resources Satellite Data and Application, Beijing 100094, China

**Abstract:** Affected by imaging conditions, data transmission, and other factors, stripe noise is common in satellite remote sensing images. It seriously restricts the quality and further use of images. In early studies, various denoising methods, such as statistics-based methods, filtering-based methods, and optimization-based methods, have been proposed to overcome the above problems. These proposed methods have achieved inspiring results in some aspects. However, they still suffer from poor adaptability, low denoising efficiency, and the need for prior knowledge. Therefore, stripe noise removal remains a challenging task.

In this study, we take advantage of the convolutional deep network while considering the characteristics of the stripe noise image itself. A deep-learning-based method is proposed, which includes three parts: a feature extraction module, a feature fusion module, and a stripe denoising module. The feature extraction module uses the convolutional layer of the same channel with different strides to extract features. As a result, different-scale feature maps of the noisy image are obtained for the following feature fusion module. The feature fusion module upsamples different-scale feature maps. It fuses these upsampled feature maps through the element-wise addition method. Finally, a

denoising network is used to predict the components of stripe noise. The stripe component is subtracted from the noise image based on predictions. Given the difficulties in obtaining real noise samples, the network is trained by simulation samples. Then, it is extended to denoise real images.

Experiments on simulation and real images show the excellent performance of our network. In the quantitative assessment, the PSNR and the SSIM of our network when simulated images are used are higher than those of the four methods. In the visual assessment, our network performs well on homogeneous and nonhomogeneous objects. Our network denoises more efficiently and retains more details of ground features than traditional methods and other denoising networks. In real noise images, our method achieves the best denoising performance with the highest ICV and the lowest MRD. Compared with traditional methods, our network has a fast denoising speed, approximately 100 times faster than the denoising speed of the optimization-based denoising method. The above experimental results demonstrate that our network has the best denoising performance in simulated and real images.

In this study, a convolutional neural network denoising method based on multiscale feature fusion is proposed based on the fully convolutional neural network. The method uses residual learning to predict the strip-noise components on images. It achieves clean images by subtracting the strip components from noisy images. Experiments demonstrate that compared with traditional methods, deep learning denoising methods are adaptive for removing stripe noise of different intensities without losing image details. The strategy of feature fusion and residual learning can effectively improve the training speed and denoising accuracy of the network. In the future, skip-connected and batch normalization layers will be included to optimize training speed and improve denoising performance. Further studies will be conducted in terms of the transfer ability of the network and the extension of its application in other types of remote sensing images, such as aerial images and hyperspectral images.

**Key words:** high-resolution image, deep learning, stripe noise, convolutional neural network, feature fusion

**Supported by** National Key Research and Development Program of China (No.2018YFA0605500); Land and Resources Research Plan of Hubei Province (No. 844-11)