

结合残差编解码网络和边缘增强的遥感图像去噪

吴从中, 陈曦, 詹曙

合肥工业大学 计算机与信息学院, 合肥 230601

摘要: 高分辨率遥感图像去噪对于提高后续图像分析、识别等问题的准确性具有重要意义。目前的去噪算法普遍存在去噪结果边缘信息模糊、易产生视觉伪影导致遥感信息丢失的缺点, 针对以上问题本文提出了一个基于边缘增强的残差编解码去噪网络用于高分辨率遥感影像去噪。该方法首先将噪声图片通过低通滤波器分解成高频层和低频层, 然后将含噪声信息的高频层输入到带残差模块的编解码网络中, 通过采样运算在多尺度空间上学习残差映射生成残差图像, 最后使用跳跃连接得到完整的去噪结果。其损失函数由逐像素和感知损失两部分组成, 逐像素损失使用传统的均方根误差学习像素级信息, 感知损失学习语义特征上的差异可以保留更多边缘信息, 最终得到更清晰的结果, 其中感知损失是由级联在后的语义分割网络提取的特征图定义的。本文对不同测试数据做去噪实验并与几个经典方法对比证明本文模型的去噪结果优于其他方法, 不仅提高图像的峰值信噪比, 得到最高的平均梯度值, 还在视觉上取得了最清晰的结果。实验结果表明, 本文提出的基于边缘增强的深层编解码卷积网络在去噪的同时可以改善边缘细节被模糊的问题, 保留更多遥感地物信息, 提高图像视觉效果。

关键词: 遥感图像去噪, 卷积神经网络, 边缘增强, 感知损失, 高分二号

引用格式: 吴从中, 陈曦, 詹曙. 2020. 结合残差编解码网络和边缘增强的遥感图像去噪. 遥感学报, 24(1): 27-36

Wu C Z, Chen X and Zhan S. 2020. Remote sensing image denoising using residual encoder-decoder networks with edge enhancement. Journal of Remote Sensing (Chinese), 24(1): 27-36 [DOI:10.11834/jrs.20208169]

1 引言

遥感图像分析已广泛应用于土地利用、城市规划及环境监测事等各个领域(Ghassemian, 2016)。其中高分辨率对地观测系统重大专项(简称高分专项)是中国2010年启动实施的重大专项之一, 高分二号卫星实现了亚米级高空间分辨率与高时间分辨率的有机结合, 使得卫星影像包含了更丰富的地物纹理信息, 在区域经济建设和社会发展发挥了重要的信息支撑和服务作用。遥感的根本目标是为了从图像上提取信息获取知识, 但在实际应用中, 遥感成像过程中通常会因为光信息转换传输而受到噪声干扰降低图像质量, 所以去噪作为遥感图像的预处理一直受到遥感应用领域的关注, 它能为图像带来更好的视觉效果, 也使其保留更高质量的特征。

目前应用于遥感图像的去噪方法主要有小波变换模型(徐冬等, 2015), 全变分(TV)模型(He W

等, 2016), 非局部自相似(NSS)模型(Li等, 2015)和稀疏分解模型(Liu等, 2016)等, 其中最具代表的是基于稀疏表示的理论。Chang等(2014)结合全变分模型和稀疏表示去除遥感图像中的条带和高斯噪声。夏琴等(2016)按照是否能够从过完备字典中选择较少原子进行K-SVD稀疏表示的原则去除遥感图像高频信息中的噪声。张从梅等(2016)将稀疏性和非局部自相似性结合起来应用于遥感图像去噪。尽管这些传统方法可以达到较好的去噪效果, 但由于复杂的优化问题, 往往需要大量的计算成本。

近些年来, 深度学习快速发展并在图像复原任务中表现优异, 越来越多的深度学习模型被用来解决自然图像去噪问题。这些深度学习算法可以分为两类: 卷积神经网络模型(CNN)和多层感知模型(MLP)。Jain和Seung(2008)首次将CNN用于去噪, 证明卷积神经网络可以通过学习端到端的非

收稿日期: 2018-04-10; 预印本: 2018-09-24

基金项目: 国家自然科学基金(编号:61371156)

第一作者简介: 吴从中, 1964年生, 男, 副教授, 研究方向为信号与信息处理。E-mail: 329161005@qq.com

通信作者简介: 詹曙, 1968年生, 男, 教授, 研究方向为模式识别, 计算机视觉和医学图像处理。E-mail: shu_zhan@hfut.edu.cn

线性映射实现去噪；基于MLP的模型有Vincent等(2008)和Xie等(2012)提出的降噪自编码器。后来随着网络学习能力不断增强，算法性能随着层数的增加而大大提升，Mao等(2016)提出了一个比较深的对称卷积编解码网络用于自然图像复原并取得显著成果。目前基于深度学习的去噪方法主要应用于传统自然图像中并取得理想效果，所以将深度学习应用于高分遥感图像去噪是十分有前景的。

不同于传统自然图像去噪，遥感图像中包含更多的地物纹理边缘等细节信息，这些细节对于遥感任务十分重要，所以遥感图像去噪的难点在于解决去噪结果模糊的问题。虽然传统的逐像素损失函数使重建结果有较高的信噪比，但是丢失了高频边缘信息，出现过度平滑的纹理边缘，因此直接将现有的去噪网络直接应用于高分遥感影像还存在一些不足。为了解决以上问题，本文提出了一个基于联合损失的残差编解码网络REDJ (Residual Encoder-Decoder Network with Joint Loss)，用来实现边缘增强的高分辨率遥感图像去噪。此网络通过一系列采样运算在多尺度空间上学习端到端非线性映射来直接重建去噪图像，同时还使用高频层分解和学习残差映射两个技巧压缩映射范围简化训练难度。为了解决生成的去噪结果过模糊的缺点，该算法使用联合损失函数，包括MSE损失和感知损失，使得网络训练时不仅遵循逐个像素比较匹配的方式学习，还通过比较学习图像高级语义特征的差异训练网络保留更多边缘和细节信息，生成视觉更清晰的去噪结果。本文采用高分二号卫星拍摄的高分辨率图像作为主要研究数据，实验表明该算法得到的结果在去噪效果和边缘细节保留方面都有很大改善。

2 基于边缘增强的残差网络

本节主要介绍所提出的去噪模型REDJ的网络结构和联合损失函数。网络模型由两部分组成：前端是带有残差模块的编解码网络，作为去噪图像的生成器；后面连接一个参数固定的损失网络，用来定义联合损失中的感知损失函数。

2.1 编解码网络

图1是生成器的结构框图，包括高频层分解，残差编解码网络和全局跳跃连接。其中卷积编解码器中又包括下、上采样和编码、解码模块。

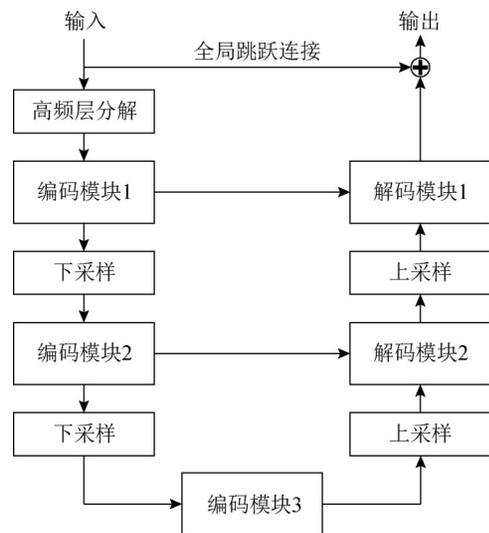


图1 编解码去噪网络的网络结构

Fig.1 Flow chart of the proposed network structure

去噪模型核心部分是左右对称带由残差模块的全卷积神经网络，借鉴U-Net的结构(Ronneberger等，2015)，包含一个编码网络和对应的解码网络，两路网络分别通过下采样和上采样进行特征维度的缩小和扩展。每对采样操作将特征表示划分到不同的空间尺度，实现多尺度信息的学习处理。具体而言，就是在原始空间维度下的特征经过下采样后缩小特征维度，使得编码模块在更小的尺度下进行卷积编码，并且每一次下采样后将卷积层的特征通道数量加倍。而上采样则是将缩小的特征维度再一步步扩大回原始尺寸，通过解码模块的学习填充之前丢失的细节信息，类似的每扩大一次尺度特征通道的数量减半。网络中每个模块的具体结构会在下文阐述。

下采样和上采样：通过下采样减小特征图大小来降低计算成本，之后再由对称的上采样将特征映射扩展为与先前比例相同的空间大小，实现多尺度处理信息学习。本文中降采样是通过大小为 2×2 的卷积层采用步长为2的方式在特征图上滑动实现的，卷积过后图的尺寸会缩减为原来的 $1/4$ 。上采样操作使用的是实现插值操作的反卷积运算。

编码模块：在每个尺度上设计一个特征编码模块。不同于传统编码器简单串联卷积层，本文引入了残差学习，结构如图2(a)所示，由一个卷积层和一个残差单元组成，每个残差单元又包括3个卷积层和一个恒等映射的捷径连接，这种三叠层的残差函数被证明能比传统两层的函数实现更

深的网络结构, 具有更好的特征学习能力(He K 等, 2016)。从上到下4个卷积层大小分别 $3\times 3, 1\times 1, 3\times 3$ 和 1×1 , 每个卷积层紧接着是批量标准化和ReLU, 第一个编码模块的卷积层特征通道数量为32,32,32,128, 之后的每经过一个下采样, 特征通道数量加一倍。

解码模块: 使用与特征编码模块类似的结构, 每层卷积层的特征通道数量也与相对应的编码卷积相同。不同的是, 解码模块还要实现上采样后的信息与之前相同尺度编码信息的融合, 因为池化时丢失部分信息, 使得直接插值上采样后还是会存在较粗糙的结果, 因此还要从高分辨率特征图中引入跳跃连接方式, 将前面编码特征图与反卷积后到同一尺度的特征图相加, 其框架如图2(b)。

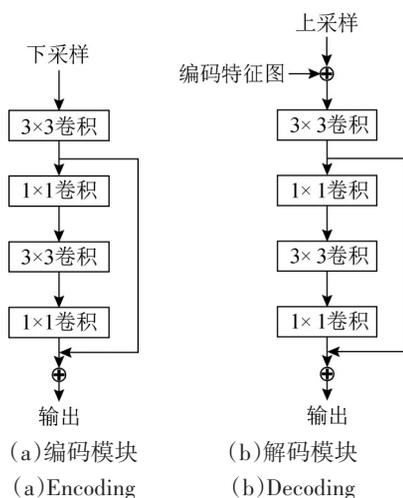


图2 编码模块和解码模块的结构
Fig.2 Flow chart of encoding and decoding

采用这样的编解码结构可以构造多尺度特征表示的深层网络, 这通常会得到更好的复原结果, 但网络的无限加深会为训练收敛带来很多困难。本文借鉴以往网络结构的经验(Kim等, 2016; Mao等, 2016)并通过权衡计算成本和恢复精度之间的关系, 最后选择3个尺度20层的卷积网络用于实验。

上文提到深层卷积网络因为计算量大、梯度消失等原因存在训练难度大不易收敛的问题。针对这些问题, 本文提出了高频层分解和学习残差映射的方法, 通过减小网络输入和输出的稀疏度来压缩映射范围, 实现简化训练的目的。原始含

噪图像可以根据灰度变化的强度分解为低频信息和高频信息两部分。其中高频信息包含大量边界信息和几乎全部的噪声, 又因为去噪任务就是要针对噪声部分进行处理, 因此可将高频层替代原始图像作为输入, 剩下包含背景内容的低频层无需被送进网络训练处理。类似的, 网络的输出也可由更稀疏的残差图片代替, 即先学习残差映射得到噪声图像和干净图像的差别, 再由全局跳跃连接计算得到最终去噪结果。为了更直观的观察改变后的输入和输出的稀疏范围缩小情况, 图3给出了一张样本图的原始噪声图 x 及其高频层 x_{detail} 、去噪图 y 和残差图 $x-y$ 和各自对应的直方图分布。通过比较可以明显看出, 不同于原始输入和原始输出直方图分布均匀的情况, 高频层和残差图的直方图分布更为集中, 多数像素的值都集中于0附近, 像素值范围明显减小。这表明使用高频分解和学习残差映射改变网络的输入输出确实能帮助减小映射范围降低计算维度简化训练。

因此本文最终的网络结构是先将噪声图像 x 经过低通滤波器(He等, 2013)分离出高频层 x_{detail} 输入残差编解码网络, 训练学习残差映射 $F(x_{detail})$, 最后由全局跳跃连接使 x 与其相减得到完整的去噪图像 y , 整个过程为

$$y = x - F(x_{detail}) \quad (1)$$

2.2 联合损失函数

以往基于学习的图像复原任务都是使用均方根误差(MSE)作为目标函数, 以取得更高的信噪比。然而这种逐像素比较的方法被发现容易丢失细节信息造成结果模糊(Ledig等, 2017)。最近很多研究(Johnson等, 2016)证明使用特征比较方式的感知损失更符合真实视觉感知, 可复原出更清晰的图像, 视觉效果远好于只使用MSE损失的模糊结果, 但有时也会因像素空间不均匀覆盖产生细微的视觉伪影。因此结合两者的优点本文提出了新的联合损失

$$L_{\text{Joint}} = L_{\text{MSE}} + L_{\text{Per}} \quad (2)$$

式中, L_{MSE} 和 L_{Per} 分别代表逐像素比较和语义特征比较的损失函数。

(1)MSE损失函数。本文的逐像素损失函数使用传统的MSE方法, 计算生成图像与干净图像的均方根误差, 使两者逐个像素点比较匹配, 公式为

$$L_{\text{MSE}} = \frac{1}{W \cdot H} \|F(x_{\text{detail}}) - (x - y)\|^2 \quad (3)$$

式中, x 是噪声图像, y 是干净的标签图像, W 和 H 分别代表输入图像对 $\{x, y\}$ 的宽度和高度。

(2) 感知损失函数。为了解决现有去噪结果模糊的缺点, 本文引入了感知损失函数。它能实现边缘增强的本质是因为: 传统 MSE 损失逐像素比

较的方式会丢失高频信息, 而感知损失通过比较图像特征差异的方法重建更多细节得到清晰的结果。又因为实验证明用于图像分类分割的神经网络本身就可以很好的学习图像的纹理边缘等语义特征, 因此可以串联预训练好的卷积网络用于提取所需的特征图。

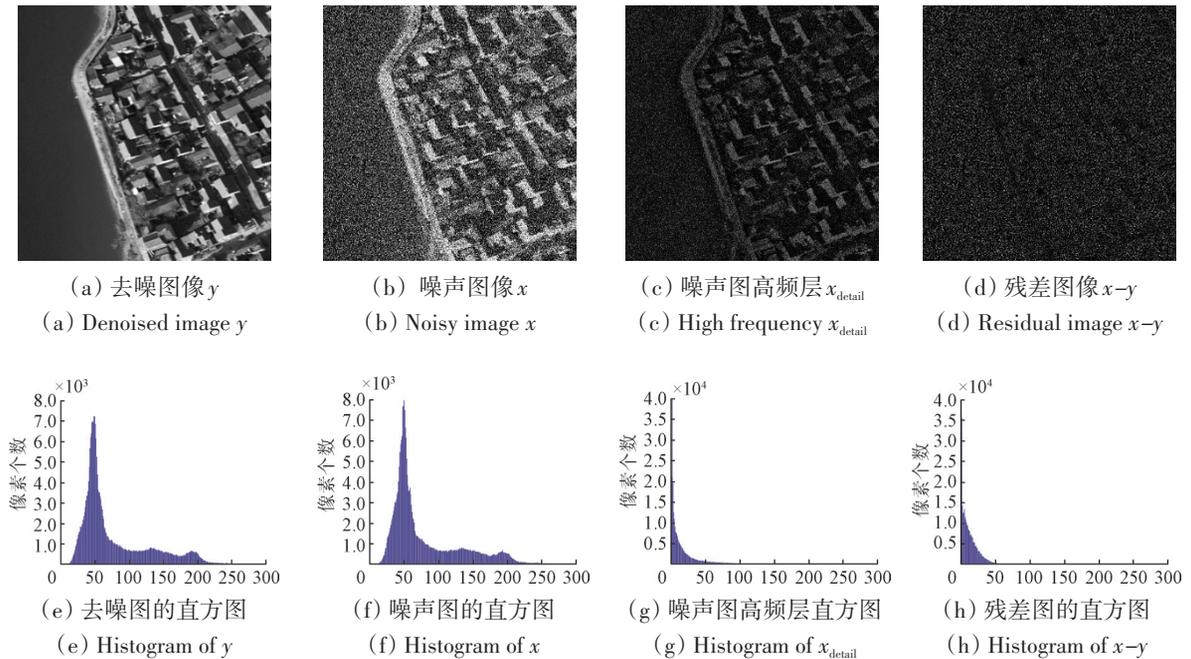


图3 高频层与残差图像的稀疏范围缩小示意图

Fig.3 Range reduction and sparsity of the residual image and high-frequency layer

本文选择能达到较好语义分割效果的 SegNet 模型作为损失网络。SegNet 包含一组卷积编码层和镜像的反卷积解码层, 其中编码部分使用的是泛化能力极强的 VGG 模型, 解码部分使用对称的结构恢复池化丢失的信息。为了保证损失网络提取特征的能力, 本文直接使用文献(Badrinarayanan 等, 2017)中预训练好分割效果优秀的 Caffe 模型。

确定好损失网络后, 便可以将感知损失定义在语义特征层面上: 将前端网络初始生成的模糊结果 $x - F(x_{\text{detail}})$ 和原始干净图像 y 分别输入 SegNet, 从其中固定的一层卷积层里提取两者的特征图, 再计算这两者特征表示的欧氏距离, 公式如下

$$L_{\text{Per}} = \frac{1}{W_i H_i} \|\varphi_i(x - F(x_{\text{detail}})) - \varphi_i(y)\|^2 \quad (4)$$

式中, W_i 和 H_i 代表所选特征图的宽度和高度, φ_i

是特征图提取器, 本文实验使用第 22 个卷积层选择图像中较丰富复杂又不至于过于抽象的边缘纹理特征和语义信息用来比较学习。

算法最终使用的联合感知损失如图 4 所示由 MSE 损失和感知损失共同组成, 具体实现方法是先将噪声图像 x 与其对应的干净图像 y 输入到去噪网络中通过 MSE 损失函数逐个像素点地比较学习两者的不同, 得到与 y 像素匹配的初始去噪结果 $x - F(x_{\text{detail}})$, 此时输出的图像比较模糊。然后接着将 y 和生成的 $x - F(x_{\text{detail}})$ 分别输入到损失网络 SegNet 中, 然后从其中一层卷积层里提取两者的特征图 $\varphi(x - F(x_{\text{detail}}))$ 和 $\varphi(y)$ 用来定义感知损失函数, 通过最小化感知损失继续训练网络来学习两者图像语义特征的差异, 使两张图像在特征感知上更相似, 原本模糊的区域因此重建了更多的边缘和细节信息, 最终生成较清晰的去噪结果。

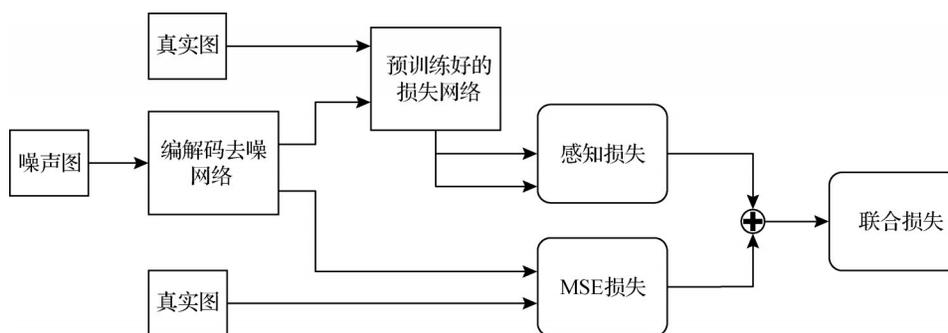


图4 级联网络框架的联合损失

Fig.4 Proposed joint loss of cascaded network

3 实验设置与结果分析

3.1 数据源

高分二号卫星作为中国首颗分辨率达到亚米级的宽幅民用遥感卫星, 搭载有两台高分辨率 1 m 全色、4 m 多光谱相机, 其中多光谱 4 个波段, 星下点空间分辨率可达全色 0.8 m、多光谱 3.2 m, 是我国目前分辨率最高的民用陆地观测卫星。本文数据采用高分二号卫星于 2017-03-21 获取的全色与多光谱影像, 数据质量符合实验要求。

影像区域为山东省日照市五莲县松柏镇及其周边地区, 范围为 119.3021° E—119.3545° E、35.7620° N—35.7323° N。五莲县地处鲁东南, 地貌以山地丘陵为主, 境内山岭起伏, 河川纵横。松柏镇位于县城东 12 km, 有丰富的土地和水资源。总体来说, 该研究区地表覆盖类型多样, 主要包括建设用地、山脉、水体、农田和林地等, 如图 5 所示。



图5 研究区图像

Fig.5 The image of research area

3.2 实验设置

(1)训练集。实验中使用全色影像, 首先将完整

的影像分割成尺寸为 512 的子块, 得到 1200 个不重叠的图像块, 然后随机选取其中 70% 作为训练样本, 用来训练去噪模型。为了确保噪声样本的质量, 本实验都是人为地对干净样本添加高斯噪声生成相对应的噪声图像。添加的噪声有两种: 一种是只添加均值为 0 标准差为 30 的高斯噪声训练单一噪声级别去噪模型 REDJ, 这种正态分布的噪声容易出现在真实卫星图像中; 另一种是随机添加 $\sigma = [1, 50]$ 范围内的高斯噪声, 用来训练盲去噪模型 REDJ(blind)。由于训练中网络的输入需要尺寸更小的子块, 将每张大小为 512×512 的图像块再次分割成 32×32 的子块, 最终可以得到 215040 个子块。

(2)测试集。为了验证本文算法的有效性, 一共选择 3 种测试数据进行实验。第 1 种是分辨率为 0.8 m 的高分二号全色影像, 使用除去训练集后剩下的 30%, 即 360 张尺寸为 512 的图像方块; 第 2 种是分辨率为 3.2 m 的高分二号多光谱影像, 使用的是与前面全色测试集对应区域的多光谱图像块; 最后从 UC Merced_LandUse 公开数据集中选择两张含真实噪声的图像作为测试样本, 用来验证本文算法对于实际含噪声的遥感卫星影像去噪的效果。

(3)参数设置。训练网络的输入是大小为 32×32 的图像。训练是通过使用带动量参数的小批量处理随机梯度下降法对学习目标进行优化的, 并将批大小和动量参数分别设置为 10 和 0.9。初始学习率被设置为 0.001, 每完成 12000 次迭代后学习率减少 5 倍, 并在 60000 次迭代后停止训练过程。所有网络的训练与测试都在 Caffe 工具上进行, 硬件参数: Intel Core i7-4790K 4.0GHz, 1X NVIDIA Tian X GPU。

3.3 结果分析

为了客观验证算法的有效性, 分别从去噪结果定量、定性分析和运行时间几个方面进行对比实

验, 并与目前各类去噪模型中最具代表性的算法作对比: 小波阈值法, TV模型法, K-SVD法。量化评价采用峰值信噪比(PSNR)和平均梯度值作为衡量标准, 其中PSNR作为去噪效果的评价指标, 平均梯度值作为图像边缘信息保留多少的评价标准。

(1) 全色影像。全色影像中只含有一个波段。类似于训练样本, 首先对所有测试图像进行加噪处理(标准差为30的高斯噪声)。然后分别采用对比算法和本文训练好的单一噪声级别去噪网络REDJ、盲去噪网络REDJ(blind)对分辨率0.8 m的含噪影像进行去噪处理, 表1给出了几种方法对测试集去噪后的峰值信噪比和平均梯度值的统计情况。通过对比可以看出本文算法无论REDJ还是REDJ(blind)得到的PSNR均高于传统算法, 说明改进的算法去除遥感图像噪声的效果明显优于其他。除此之外, 该算法能得到最高的平均梯度值, 这说明相较于其他结果, 本文的去噪结果能保留更多的边缘信息和更清晰的纹理特征, 实现了边缘增强的效果。

除了以上定量评估, 还从测试集中选取比较有代表的样本图片的去噪结果图进行视觉对比,

样本图像由建设用地、耕地、水体、山脉等地物混合构成。图6是一张完整尺寸的去噪结果实例, 由图6可知本文的算法在有效去除噪声的同时, 保持了更好的图像清晰度, 学习到更多的细节特征。图7给出了3处局部放大图, 由此可以更清晰的比较, 小波阈值和TV模型处理还残留了部分残余噪声点, 使得图像纹理和边缘比较模糊; K-SVD较前两种方法去噪效果更彻底, 但依然存在边缘细节丢失, 纹理区域模糊的问题; 相比而言, REDJ不仅在多尺度上进行训练, 还考虑感知损失的方式, 使得去噪结果能够得到最好的视觉效果, 含有更丰富的地物纹理信息, 尤其是局部边缘, 如山脊、房屋、农田的边界轮廓等。

表1 全色遥感图像去噪结果质量评价指标统计表

Table 1 The quality parameters of the panchromatic denoised image

	原图	小波 阈值	TV 模型	K- SVD	REDJ	REDJ (blind)
PSNR	—	28.68	29.35	30.41	31.10	30.92
平均梯度值	2.91	2.62	2.57	2.48	2.79	2.69

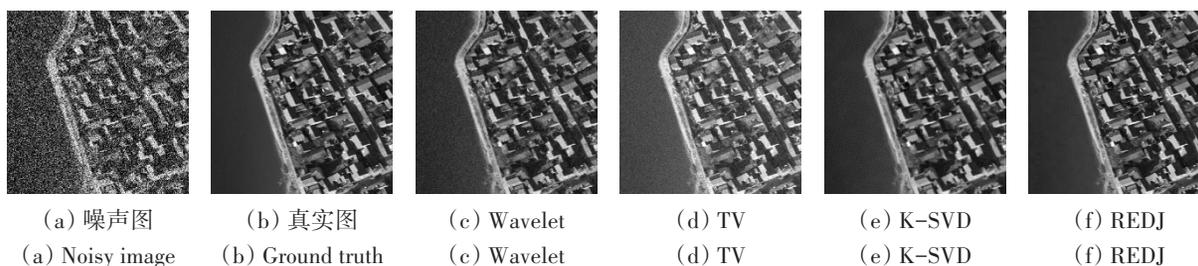


图6 全色影像去噪结果

Fig.6 Denoised results of panchromatic image

为了更直观地说明网络增强边缘的能力, 本文还比较了两个PSNR较高算法的边缘检测效果: K-SVD和REDJ。采用Sobel算子进行检测, 图8为检测结果。与干净图像的边缘比较, K-SVD结果图的边缘线条不连续, 部分地物轮廓丢失。相比而言REDJ的结果更为完整, 与原图检测出的边缘信息十分接近。由此说明, 本文算法处理后的影像残留噪声较少, 遥感的边缘轮廓信息保留更全, 有利于后期的检测、分类等遥感应用。

此外为了验证算法处理噪声级别不一致等复杂情况的效果, 本文增加了一组盲去噪实验, 将一张包括多种噪声级别的遥感图像用训练好的盲去噪模型REDJ(blind)进行去噪处理。图9(a)的混合

噪声图像由3种噪声级别组成, 即左边部分为标准差10的噪声, 中间为30, 右边为50。通过比较图9(b)残差图和图9(c)结果图可以看到, 对于噪声级别不一致的情况, 盲去噪模型也可以一次性处理并产生满意的去噪结果。为了便于观察, 其中图9(b)残差图是适当增加对比度后更清晰的结果。

(2) 多光谱影像。为验证算法的波段适用性, 还对分辨率低一些的多光谱图像进行实验。高分二号的多光谱影像有4个波段, 本实验使用第2波段为例。与全色影像相同, 先对测试图像加标准差为30的高斯噪声。表2和图10展示了不同算法去噪图像的定量指标结果和视觉结果, 比较可知本文算法依

旧获得的最高的指标值和最清晰的视觉效果, 去噪最为彻底, 保留的纹理边缘信息也最丰富。证明了

该方法适用于不同波段图像去噪, 泛化能力较强。

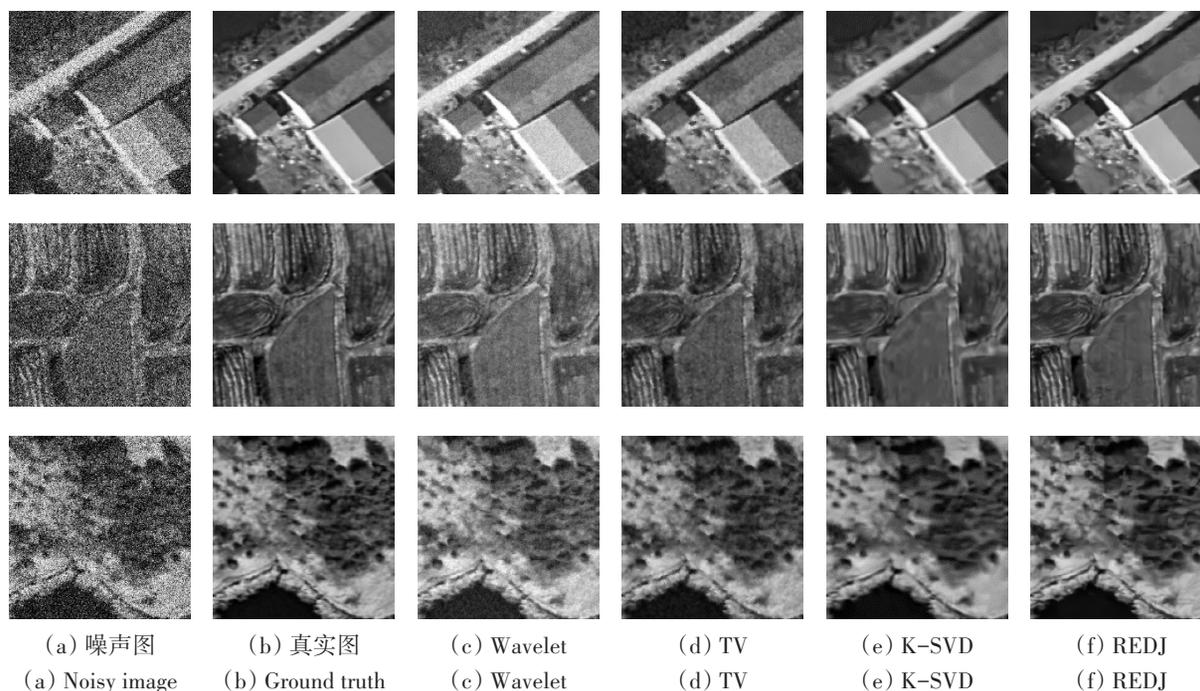


图7 局部图放大显示
Fig.7 Zoom in partial view

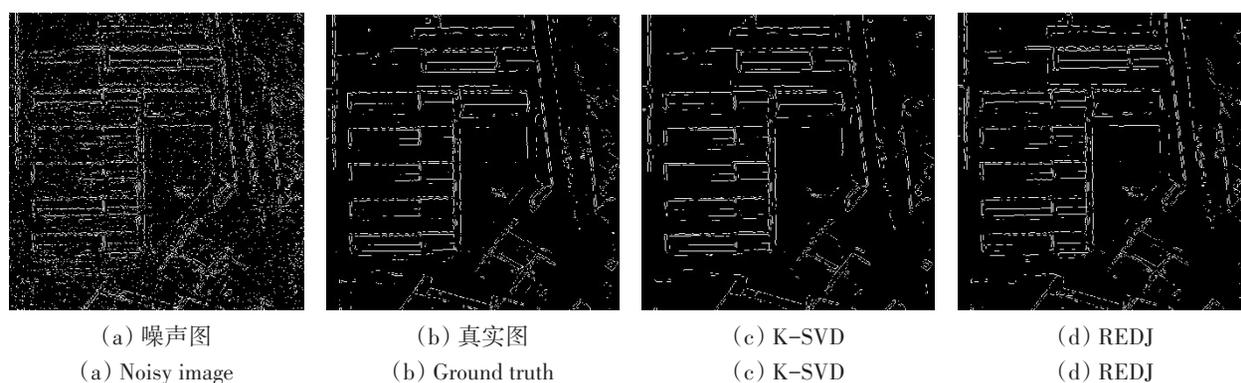


图8 Sobel算子进行边缘检测的结果
Fig.8 Edge detection results of Sobel

表2 多光谱遥感图像去噪结果质量评价指标统计表
Table 2 The quality parameters of the hyperspectral de-noised image

	原图	小波 阈值	TV 模型	K- SVD	REDJ	REDJ (blind)
PSNR	—	29.20	31.87	33.98	34.60	34.42
平均梯度值	1.85	1.44	1.42	1.05	1.81	1.68

(3)真实噪声图像。为验证本文算法的普适性, 进行了真实噪声去噪实验。由于真实噪声的分布是

随机未知的, 无法获取其对应的干净图像, 因此实验只选择主观性的对比方法。图11给出两张真实遥感影像使用3种算法的去噪结果, 其中小波阈值法和K-SVD都是使用的去噪模型, 本文算法使用的是盲去噪模型 $\sigma = 10$ 。从实验结果图中可以看出小波阈值法去噪并不彻底, K-SVD虽然去噪效果很好, 但丢失了大量的边缘信息, 视觉效果较为模糊。而本文算法不仅去除了大量真实噪声, 还保留了更多的细节, 得到更清晰真实的结果, 其地物纹理丰

富且轮廓特征明显突出。证明对于分布未知的真实噪声本文算法也能较好的处理，具有一定普适性。

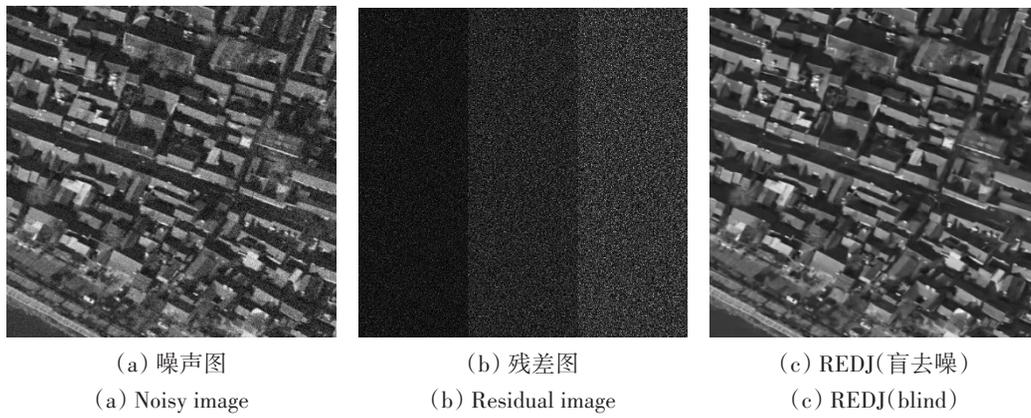


图9 盲去噪模型处理多种噪声级别的去噪结果

Fig.9 Denoised result of blind denoising

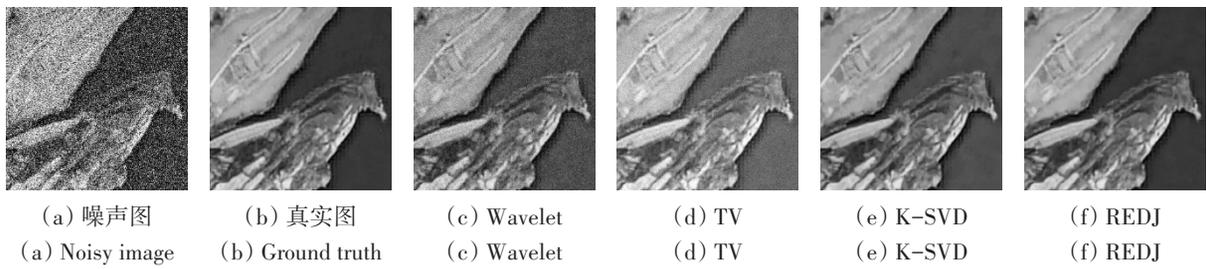


图10 多光谱影像去噪结果(第2波段)

Fig.10 Denoised results of hyperspectral image(The second band)

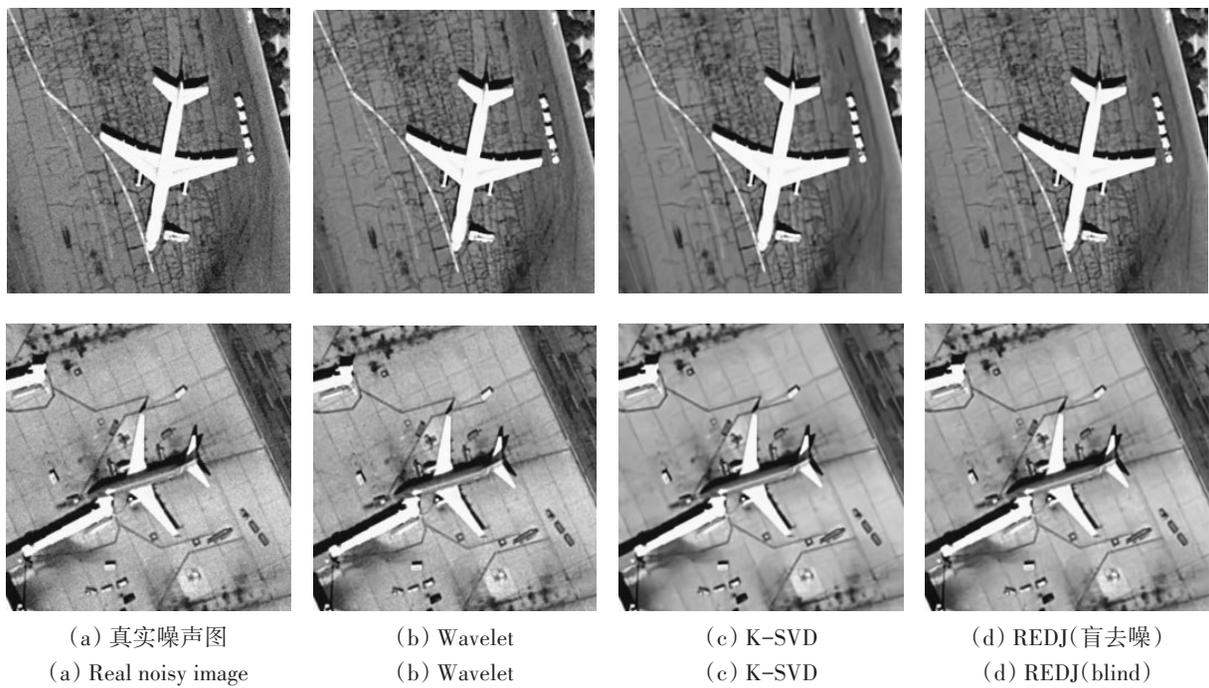


图11 真实噪声图像去噪结果

Fig.11 Denoised results of real noisy image

(4)运行时间。除了分析去噪结果的优劣,本文还对比了时间成本。表3列出各算法测试一张全色噪声影像的运行时间。以上所有的测试都是基于同样的机器配置,其中REDJ同时给出CPU时间(左)和GPU时间(右)。比较可知,CPU版本的运行时间相对TV和K-SVD已经有显著提升,GPU时间更是比其他算法快速更多。

表3 运行时间
Table 3 Running time

	小波	TV	K-SVD	REDJ
时间/s	0.19	140.69	197.27	3.41/0.06

注:REDJ给出了CPU时间(左)和GPU时间(右)。

4 结论

深度学习近几年来在图像处理方面得到了广泛的应用,本文将深度学习理论应用于遥感图像去噪,提出了一个20层的残差编解码去噪网络用于高分二号影像。相对之前的去噪方法,本文模型结合了残差学习和感知损失不仅能有效的去除噪声,还能指导生成更真实清晰的结果,改善了边缘模糊的问题,对于纹理信息丰富的遥感图像可以保留更多的地物轮廓和特征,有利于后期的检测、分类等遥感技术应用。对于未来的工作,还需要探索如何处理其他类型的噪声,特别是存在于现实遥感影像复杂真实的噪声种类。

参考文献(References)

- Badrinarayanan V, Kendall A and Cipolla R. 2017. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12): 2481-2495 [DOI: 10.1109/TPAMI.2016.2644615]
- Chang Y, Yan L X, Fang H Z and Liu H. 2014. Simultaneous destriping and denoising for remote sensing images with unidirectional total variation and sparse representation. *IEEE Geoscience and Remote Sensing Letters*, 11(6): 1051-1055 [DOI: 10.1109/LGRS.2013.2285124]
- Ghassemian H. 2016. A review of remote sensing image fusion methods. *Information Fusion*, 32: 75-89 [DOI: 10.1016/j.inffus.2016.03.003]
- He K M, Sun J and Tang X O. 2013. Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(6): 1397-1409 [DOI: 10.1109/TPAMI.2012.213]
- 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, NV, USA: IEEE: 770-

- 778 [DOI: 10.1109/CVPR.2016.90]
- He W, Zhang H Y, Zhang L P and Shen H F. 2016. Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration. *IEEE Transactions on Geoscience and Remote Sensing*, 54(1): 178-188 [DOI: 10.1109/TGRS.2015.2452812]
- Jain V and Seung H S. 2008. Natural image denoising with convolutional networks//Proceedings of the 21st International Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada: Curran Associates Inc.: 769-776
- Johnson J, Alahi A and Li F F. 2016. Perceptual losses for real-time style transfer and super-resolution//European Conference on Computer Vision. Cham: Springer: 694-711 [DOI: 10.1007/978-3-319-46475-6_43]
- 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, NV, USA: IEEE: 1646-1654 [DOI: 10.1109/CVPR.2016.182]
- Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z H and Shi W Z. 2017. Photo-realistic single image super-resolution using a generative adversarial network//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE: 105-114 [DOI: 10.1109/CVPR.2017.19]
- Li J, Yuan Q Q, Shen H F and Zhang L P. 2015. Hyperspectral image recovery employing a multidimensional nonlocal total variation model. *Signal Processing*, 111: 230-248 [DOI: 10.1016/j.sigpro.2014.12.023]
- Liu X X, Lu X L, Shen H F, Yuan Q Q, Jiao Y L and Zhang L P. 2016. Stripe noise separation and removal in remote sensing images by consideration of the global sparsity and local variational properties. *IEEE Transactions on Geoscience and Remote Sensing*, 54(5): 3049-3060 [DOI: 10.1109/TGRS.2015.2510418]
- Mao X J, Shen C H and Yang Y B. 2016. Image restoration using convolutional auto-encoders with symmetric skip connections. arXiv: 1606.08921
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//Navab N, Hornegger J, Wells W and Frangi A, eds. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Vincent P, Larochelle H, Bengio Y and Manzagol P A. 2008. Extracting and composing robust features with denoising autoencoders//Proceedings of the 25th International Conference on Machine Learning. Helsinki, Finland: ACM: 1096-1103 [DOI: 10.1145/1390156.1390294]
- Xia Q, Xing S, Ma D Y, Mo D L, Li P C and Ge Z X. 2016. An improved K-SVD-based denoising method for remote sensing satellite images. *Journal of Remote Sensing*, 20(3): 441-449(夏琴, 邢帅, 马东洋, 莫德林, 李鹏程, 葛忠孝. 2016. 遥感卫星影像K-SVD稀疏表示去噪. *遥感学报*, 20(3): 441-449) [DOI: 10.11834/jrs.20165149]
- Xie J Y, Xu L L and Chen E H. 2012. Image denoising and inpainting

- with deep neural networks//International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada: Curran Associates Inc.: 341-349
- Xu D, Sun L and Luo J S. 2015. Denoising of hyperspectral remote sensing imagery using NAPCA and complex wavelet transform. *Infrared and Laser Engineering*, 44(1): 327-334(徐冬, 孙蕾, 罗建树. 2015. 结合 NAPCA 和复小波变换的高光谱遥感图像去噪. *红外与激光工程*, 44(1): 327-334) [DOI: 10.3969/j.issn.1007-2276.2015.01.056]
- Zhang C M, Sun Q S, Wang C, Feng L and Gu Y H. 2016. Remote sensing image de-noising via sparse representation based on non-local self-similarity. *Remote Sensing Technology and Application*, 31(4): 739-747(张从梅, 孙权森, 王超, 封磊, 顾一禾. 2016. 基于非局部自相似性的遥感图像稀疏去噪算法. *遥感技术与应用*, 31(4): 739-747)[DOI: 10.11873/j.issn.1004-0323.2016.4.0739]

Remote sensing image denoising using residual encoder-decoder networks with edge enhancement

WU Congzhong, CHEN Xi, ZHAN Shu

School of Computer and Information, Hefei University of Technology, Hefei 230601, China

Abstract: Remote sensing images are often affected by noise in the process of digitization and transmission processes. Denoising is an indispensable way of improving image quality. Despite showing an excellent noise removal performance, the existing denoising algorithms however typically suffer from a common drawback. Specifically, in the learning process, some edge information is lost, thereby over-smoothing the denoising result. Given the importance of details—including sharp edges and texture information—in remote sensing images, we propose a residual encoder - decoder denoising network with joint loss (REDJ) for GF-2 satellite data.

Inspired by U-net, we use a deep convolutional framework is used to learn the end-to-end mapping from noisy images to the original ones. The encoder acts as a feature extractor that captures semantic information of image contents while eliminating noise, whereas the decoder recovers the image details. The high-resolution features from the encoder are combined with the up-sampled output by skip connection. We also introduce high-frequency decomposition and residual mapping to simplify the training process by reducing the solution space. As for the loss function, we modify the traditional denoising per-pixel loss. Given a well-trained convolutional neural network for defining perceptual loss, we instead to learn the perceptual differences of the extracted features instead of merely matching the low-level pixel information. Unlike the loss of detail resulting from normal per-pixel MSE loss, we recommend a new joint loss that combines the advantages of both per-pixel reconstruction and feature reconstruction, preserves additional edge and texture information, and generates clear denoised results. We employ the GF-2 satellite images in the experiments. To obtain enough training and testing data, we divide the entire high-resolution image is divided into 1200 pictures of size 512 and then allocate 70% of these images for training and the other 30% for testing. We generate the noisy images by adding Gaussian noise.

To verify the effectiveness of our proposed network, we compare our quantitative and qualitative results with those of other state-of-the-art methods, including wavelet threshold, total variation, and K-SVD. Our proposed method REDJ can obtain the best index values both of PSNR and average gradient. In the qualitative visual sense, REDJ obtains clear denoising results because of the joint of perceptual loss. Compared with other methods that produce blurred regions generated by other methods, REDJ preserves more edge information and texture details. We also compare the run times of different methods for denoising images and find that REDJ has a relatively high CPU speed and achieves an excellent computational efficiency on GPU time.

This paper successfully applies deep learning theory for denoising remote sensing images. We use the proposed network is used to remove noise from high-resolution GF-2 remote sensing images and to preserve the edge contours and fine details, which is conducive to facilitate later detection, classification, and other remote sensing applications. In our future work, we will explore to handle other types of noises, especially the complex real-world noises, and consider a single comprehensive network for more image restoration tasks.

Key words: remote sensing image denoising, convolutional neural network, edge enhancement, perceptual loss, GF-2 remote sensing image

Supported by National Natural Science Foundation of China (No.61371156)