



利用动态上采样滤波深度网络进行多角度遥感影像超分辨率重建

陈行, 罗斌

引用本文:

陈行, 罗斌. 利用动态上采样滤波深度网络进行多角度遥感影像超分辨率重建[J]. 武汉大学学报·信息科学版, 2021, 46(11): 1716–1726.

CHEN Hang, LUO Bin. Multi-angle Remote Sensing Images Super-Resolution Reconstruction Using Dynamic Upsampling Filter Deep Network[J]. *Geomatics and Information Science of Wuhan University*, 2021, 46(11): 1716–1726.

相似文章推荐 (请使用火狐或IE浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

高分辨率可见光遥感图像舰船目标检测综述

A Survey on Ship Detection Technology in High-Resolution Optical Remote Sensing Images

武汉大学学报·信息科学版. 2021, 46(11): 1703–1715 <https://doi.org/10.13203/j.whugis20200481>

基于深度学习的立体影像密集匹配方法综述

A Review of Dense Stereo Image Matching Methods Based on Deep Learning

武汉大学学报·信息科学版. 2021, 46(2): 193–202 <https://doi.org/10.13203/j.whugis20200620>

面向无人机倾斜影像的高效SfM重建方案

Solution for Efficient SfM Reconstruction of Oblique UAV Images

武汉大学学报·信息科学版. 2019, 44(8): 1153–1161 <https://doi.org/10.13203/j.whugis20180030>

联合全极化SAR和IRI估计三维电子密度分布

Estimation of Three-Dimensional Electron Density Distribution Using Polarimetric SAR and IRI Observations

武汉大学学报·信息科学版. 2021, 46(11): 1677–1685 <https://doi.org/10.13203/j.whugis20210061>

利用AB算法进行高分四号卫星数据反照率反演

Retrieval of GF-4 Satellite Image Data Surface Albedo Based on Angular Bin Algorithm

武汉大学学报·信息科学版. 2020, 45(4): 542–549 <https://doi.org/10.13203/j.whugis20180291>



利用动态上采样滤波深度网络进行多角度 遥感影像超分辨率重建

陈行¹ 罗斌¹

¹ 武汉大学测绘遥感信息工程国家重点实验室,湖北 武汉,430079

摘要:近年来,基于深度学习的超分辨率重建技术已经广泛应用于多时相高光谱影像、高分影像超分辨率重建等领域。多角度遥感影像之间具有丰富的互补信息,可用于超分辨率重建。针对高分辨率多角度遥感影像提出了一种基于动态上采样滤波网络的超分辨率重建方法。该方法的网络结构为端到端双路网络,其中一个分支网络通过动态上采样滤波模块来实现分辨率提升,另一个分支网络用来学习影像中的高频信息,将两个分支网络输出的结果相加即可得到最终的超分辨率重建影像。为了验证该方法的有效性,利用 WorldView-2 美国亚特兰大地区和巴西里约热内卢地区多角度遥感影像数据分别进行了2倍、3倍、4倍超分辨率重建模拟实验和真实实验,并进行了多组对比实验。实验结果表明,所提方法可以在顾及多角度影像角度维信息的同时有效提升目标影像空间分辨率,并且较好地保持了影像的细节信息。

关键词:超分辨率重建;多角度遥感影像;深度学习;动态上采样滤波

中图分类号:P237

文献标志码:A

高分辨率遥感影像可以提供地物丰富的细节信息,已经被广泛应用于城市监控、精准农业、灾害监测以及军事目标检测等领域。然而,诸多条件的限制导致高分影像难以获取,而影像超分辨率重建技术通过处理多景具有互补信息的低分辨率观测影像可以重建出一景或多景高分辨率影像,该技术可以充分利用已有的低分辨率成像系统经济高效地获取高分辨率影像。目前,超分辨率重建已经应用于多种不同种类的遥感影像上,比如美国陆地卫星计划(Landsat)专题制图仪(thematic mapper, TM)影像、Landsat时序分析(extracted timing model, ETM)影像、中分辨率成像光谱仪(moderate-resolution imaging spectroradiometer, MODIS)影像、机载可见光近红外成像光谱(airborne visible infra red imaging spectrometer, AVIRIS)影像以及 WorldView 影像^[1-9]。

多角度遥感影像成像系统可以在几乎同一时间内获取目标地物多个角度的观测影像,因此该系统可以有效避免由于场景变化和天气条件变化造成的影响,应用多角度遥感影像进行超分辨率重建具有天然优势。文献[10]首次提出超

分辨率重建的概念并应用于 Landsat 4 影像。随后,超分辨率重建技术得到了广泛关注,并衍生出了一系列经典的超分辨率重建算法,如最大后验估计法^[11]、凸集投影法^[12]、非均值内插法^[13]、最大似然法^[14-15]、混合最大后验估计/凸集投影法^[16]等。超分辨率重建技术在遥感领域也得到了发展,文献[9]提出了一种基于可变像素线性重建的超分辨率重建方法,并应用于 Landsat ETM+ 数据;文献[7]针对 MODIS 影像提出了一种超分辨率重建方法;文献[6]提出了一种针对遥感影像数据的基于隐马尔可夫树模型的超分辨率重建方法,并且使用多时相 Landsat 7 数据进行了测试。上述方法都是针对多时相遥感影像的超分辨率重建提出的,然而,多时相遥感影像成像周期长,在不同时相的遥感影像中,地面目标会发生变化,且受天气条件影响也比较大。

多角度遥感影像通过摆镜扫描可以在几乎同一时间获取目标地物多个角度影像,一定程度上克服多时相影像应用于超分辨率重建的难点。针对多角度遥感影像,学者们也提出了相应的超分辨率重建算法。文献[3]提出了一种基于

德劳内三角化非标准插值方法的超分辨率重建方法;文献[17]使用资源3号多角度遥感影像进行超分辨率重建,该方法首先利用傅里叶变换方法对多角度影像进行配准,再利用最大后验概率方法对影像进行重建;文献[5]针对多角度WorldView-2影像提出了一种超分辨率重建算法,该算法在影像配准阶段考虑了影像局部几何畸变以及影像之间存在的视差问题,并在超分辨率重建模型中加入了L1范数约束以及全变差正则项;文献[18]提出了一种基于核回归的超分辨率重建方法,该方法将影像分为局部影像块,并使用 n 阶泰勒展开近似表示。然而上述方法都没有考虑多角度影像之间存在影像空间分辨率变化的情况,基于此,文献[19]提出了一种自适应权重的超分辨率重建方法,通过对不同角度低分辨率影像进行加权来获取最终的超分辨率重建影像;文献[20]考虑了多角度遥感影像的低秩特性,提出了一种同时实现多角度遥感影像配准及影像超分辨率重建的框架,重建结果同时保留了影像光谱信息及空间细节信息。上述方法都是基于传统机器学习框架对多框幅遥感影像超分辨率重建进行研究。

近年来,利用深度学习技术对遥感影像进行超分辨率重建的研究也越来越多,如文献[21]将生成对抗网络作为超分辨率重建的基本框架,使用密集残差网络增强网络的特征提取能力,并在机载遥感影像上进行了实验;文献[22]提出了一种基于密集卷积神经网络的遥感影像超分辨率重建框架,该框架有效提升了重建影像的纹理细节。深度学习技术同样被广泛应用于多框幅遥感影像超分辨率重建,文献[23]针对多框幅遥感影像提出了一种能够同时提取多框幅影像空间信息及时相信息的卷积神经网络(convolutional neural networks, CNN)结构;文献[24]提出了一种端到端神经网络框架,该框架包含了一个编码器模块、融合模块以及一个解码模块,首先在编码阶段提取配准后的多景低分辨率影像特征表达,然后在融合模块中整合多景影像特征表达,最后通过解码器获取重建后影像。然而,上述方法都是针对多时相遥感影像,尚未开展有关多角度遥感影像的研究。相比多时相遥感影像,多角度遥感影像存在由于观测角度变化引起目标地物偏移的问题。多幅不同角度遥感影像之间存在空间分辨率差异,这对多角度影像信息的融合也构成挑战。

由此,本文提出了一种面向多角度遥感影像

超分辨率重建的端到端深度网络结构(multiview super-resolution net, MVSR-Net)。首先将多个角度遥感影像输入MVSR-Net网络中;然后利用输入的多角度遥感影像提取高维特征,并堆叠为一个高维特征立方体,将提取的高维特征分别输入残差学习网络和动态上采样滤波网络,并输出高频影像和动态上采样影像;最后将高频影像与动态上采样影像相加即可得到超分辨率重建的影像。为了验证该网络结构的有效性,利用WorldView-2美国亚特兰大地区和巴西里约热内卢地区多角度遥感影像数据集进行了实验。实验结果表明,本文方法的网络结构相比没有加入动态上采样滤波器的传统三维卷积神经网络(three-dimensional convolutional neural networks, 3D-CNN),在超分辨率重建结果上有了比较显著的提升。

1 方法原理

1.1 网络结构设计

多角度遥感影像超分辨率重建的目标是利用多个角度观测到的低分辨率遥感影像 $\{X_{t-N}, X_{t-N+1}, \dots, X_t, \dots, X_{t+N}\}$ 重建出高分辨率影像 $\{\hat{Y}_t\}$,其中, t 代表最接近正射观测角度的中心观测影像, X_{t-N}, X_{t+N} 表示距离正射影像最邻近角度影像, N 越大表示距离正射角度影像的角度越大。根据本文提出的多角度遥感影像超分辨率重建框架和网络参数 θ 进行影像重建,计算如下:

$$\hat{Y}_t = G_\theta(X_{t-N:t+N}) \quad (1)$$

式中, \hat{Y}_t 表示网络输出的重建影像; G_θ 表示超分辨率重建网络; $X_{t-N:t+N}$ 表示输入的多角度影像。

假设输入模型 G 的张量尺寸为 $T \times H \times W \times C$, $T = 2N + 1$, H 和 W 分别表示输入低分辨率影像的高和宽, C 表示遥感影像通道数,则该模型的输出张量大小为 $1 \times rH \times rW \times C$, r 表示上采样倍数。

本文方法通过生成动态上采样滤波器 F_t 和残差影像 R_t 来获取最终的高分辨率影像,输入的中心角度影像首先通过动态上采样滤波得到滤波后的影像,然后与通过网络得到的残差影像 R_t 相加得到最终的高分辨率重建影像 \hat{Y}_t 。

本文方法的网络结构如图1所示,首先,每个输入的低分辨率多角度遥感影像经过 $1 \times 3 \times 3$ 的三维卷积层得到特征图;其次,将提取的特征沿角度维进行堆叠,得到一个三维的空间维角度维特征立方体;然后,将提取到的特征输入到两

个共享网络参数的分支网络(残差影像生成网络和滤波器生成网络),该部分的设计参考密集网络^[25],并且根据多角度遥感影像的数据结构对网络结构进行了改进;最后,从密集网络处理得到的特征分别输入到动态上采样滤波模块以及残

差学习模块。本文方法的网络结构将原本密集网络中的二维卷积层替换成三维网络结构,可以更有效地提取多角度影像中的角度维特征。相比于二维卷积网络,三维卷积网络可以有效提取多维遥感影像的空间时相维特征^[26-27]。

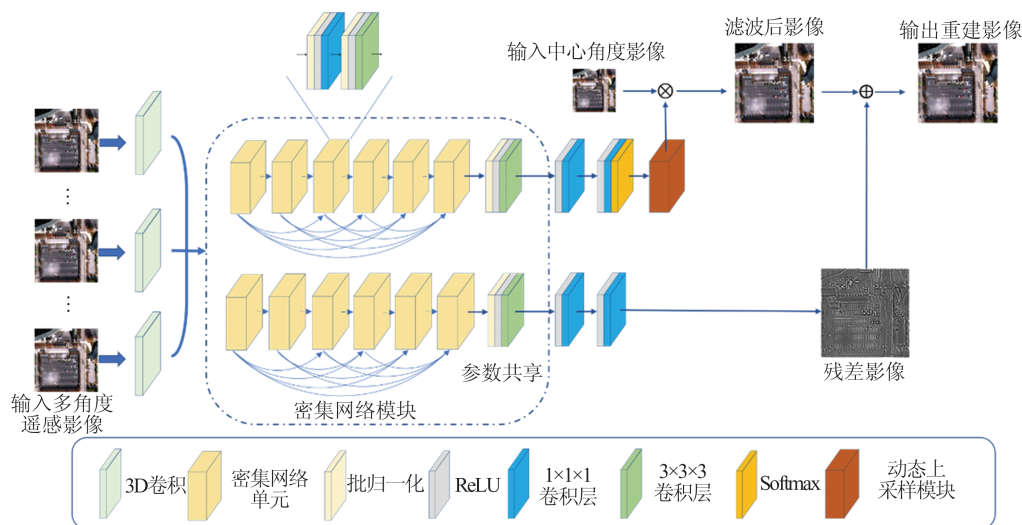


图1 MVSR-Net网络结构

Fig.1 Network Architecture of MVSR-Net

1.2 动态上采样滤波模块

本文将动态上采样滤波模块(dynamic upsampling filters, DUF)^[28]应用于多角度遥感影像,根据低分辨率多角度遥感影像中每个像素的空间维角度维邻域动态生成上采样滤波器,主要流程如图2所示。

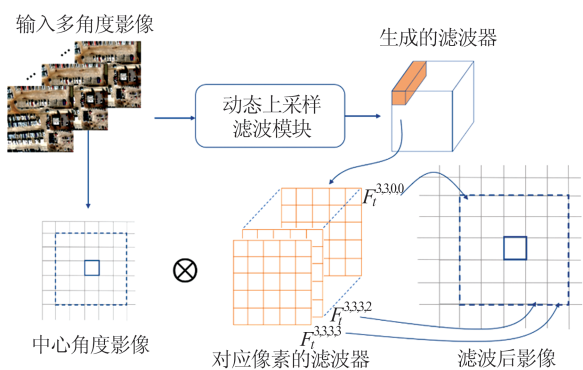


图2 动态上采样模块结构

Fig.2 Structure of the Dynamic Upsampling Filter Block

首先将多帧低分辨率多角度影像 $\{X_{t-N}, X_{t-N+1} \cdots X_t \cdots X_{t+N}\}$ 输入到动态上采样滤波网络结构,假设 $N=3$,共输入7个角度低分辨率影像,训练好的网络结构会生成总数为 r^2HW 、大小为 5×5 的上采样滤波器,生成的滤波器与原始输入的低分辨率影像相乘得到滤波后的高分辨率影像帧 \bar{Y}_t 。然后利用每个像素值对应的滤

波器 $F_t^{y,x,v,u}$ 对输入的原始影像 X_t 中的低分辨率像素进行局部滤波,即可输出对应位置的高分辨率像素,具体计算如下:

$$\bar{Y}_t(yr+vr, xr+ur) =$$

$$\sum_{j=-2}^2 \sum_{i=-2}^2 F_t^{y,x,v,u}(j+2, i+2) X_t(y+j, x+i) \quad (2)$$

式中, (x, y) 表示低分辨率影像中像素的位置; (v, u) 表示滤波器中的位置坐标。上述操作和反卷积操作类似,因此该网络框架可以实现后向传播进而实现端到端的训练。

在动态上采样滤波框架中,本文利用深层神经网络学习最佳上采样滤波器,基于给定的低分辨率遥感影像重建得到高分辨率遥感影像。根据像素偏移创建动态上采样滤波器,避免了多角度影像匹配过程。

1.3 残差学习

使用动态上采样滤波器后的重建影像实际是输入像素的加权,缺乏清晰度。为了改善这种现象,本文方法的网络结构中加入了一个残差学习分支来学习影像中的高频信息。经典方法深度网络超分辨率重建方法(very deep super-resolution, VDSR)^[29]利用网络学习到的残差信息与基于双三次上采样插值的影像相加来获得最后重建的影像。与VDSR方法不同,本文方法输入的是多角度遥感影像,而VDSR输入的是单幅影

像。其次,本文在网络结构中使用动态上采样滤波框架作为超分辨率重建的基准影像,相比双三次上采样插值重建的影像具有更高的精度。且本文网络结构利用残差学习获取的高频信息与动态上采样滤波模块得到的影像相加来获得最终的重建影像。因此,本文方法在顾及多角度影像角度维信息的同时得到空间清晰度较高的重建影像。

1.4 网络设置

在 MVSR-Net 网络框架中,对原始影像进行信息提取的卷积层包含 64 个滤波器,其卷积核大小都为 3×3 。密集网络框架总共包含 6 个密集网络单元,每个密集网络单元依次包含一个批归一化层、一个 ReLU 激活层、一个 $1 \times 1 \times 1$ 卷积层、一个批归一化层、一个 ReLU 激活层及一个大小为 $3 \times 3 \times 3$ 的 3D 卷积层。在第一个密集网络模块中,共有 64 个 $1 \times 1 \times 1$ 卷积层和 32 个 3D-CNN 卷积层,每个密集网络模块卷积层输出层数依次为 96、128、160、192、224 和 256。动态上采样滤波模块以及残差影像生成模块的网络设置见图 1,输入网络的低分辨率影像的大小分别为 96×96 、 128×128 和 192×192 ,分别对应 4 倍、3 倍和 2 倍影像重建网络结构。网络训练的批处理大小设置为 6,输入的影像都经过了归一化处理。

为了训练网络 G_θ ,本文方法利用 Huber 损失作为网络稳定收敛的代价函数 L ,计算如下:

$$L = H(\hat{Y}_t, Y_t) = \begin{cases} \frac{1}{2} \|\hat{Y}_t - Y_t\|_2^2, & \|\hat{Y}_t - Y_t\|_1 \leq \delta \\ \delta \|\hat{Y}_t - Y_t\|_1 - \frac{1}{2} \delta^2, & \text{其他} \end{cases} \quad (3)$$

式中, $H(\hat{Y}_t, Y_t)$ 表示 Huber 函数; Y_t 表示参考影像; 阈值 δ 设置为 0.01。本文选择 Adam 优化器进行网络优化,初始的学习率设置为 0.001。为了避免损失函数区域发散,每经过 10 次循环训练,学习率缩小 10 倍。

2 实验与结果分析

2.1 实验设置

本文实验是在 Linux 下 python 设置的深度学习框架进行的, GPU 采用 32 GB 显存 NVIDIA TitanX 显卡。实验使用的训练数据为 SpaceNet 提供的多角度遥感影像公开数据集 (<https://spacenet.ai/off-nadir-building-detection/>), 该数据集包含 27 个角度 4 通道 WorldView-2 影像, 拍摄

地点位于美国佐治亚州亚特兰大地区, 该数据集按照观测角度 $0^\circ \sim 25^\circ$ 、 $26^\circ \sim 40^\circ$ 和 $41^\circ \sim 55^\circ$ 分为 3 组影像。较大的观测角度会导致目标地物发生较大的几何位移, 而且对于目标中高程较大的地物, 几何变形的现象更加显著。因此, 为了保证网络的学习能力, 实验的训练数据及测试数据选择观测角度为 $0^\circ \sim 25^\circ$ 的影像, 图 3 展示了部分样例数据。实验选取了最接近正射角度的 7 个角度的影像进行训练及测试, 其基本信息如表 1 所示。数据集中包含 1 064 景 900×900 大小的全色影像及 1 064 景 225×225 包含 4 个通道的多光谱影像, 本文在 1 064 景影像随机选取 181 景影像, 使用 Gram-Schmidt 融合算法将全色影像和多光谱影像进行融合, 获得 181 景大小为 900×900 且包含 4 个通道的多光谱影像, 然后将每景影像随机裁剪出 50 块 384×384 的影像, 最后得到包含 9 050 组多角度遥感影像的数据集。

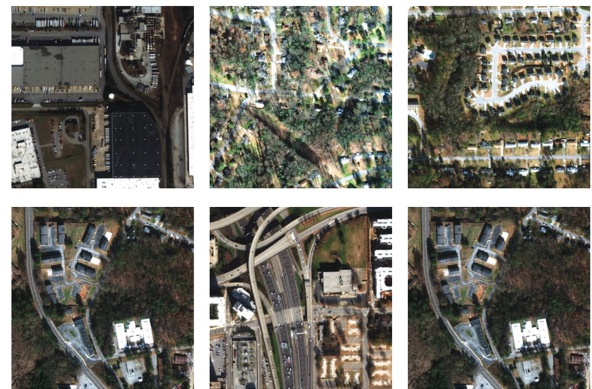


图 3 数据集中的样例数据

Fig.3 Samples of Dataset

表 1 实验数据基本信息

影像编号	观测角度/ $^\circ$	目标方位角/ $^\circ$	影像分辨率/m
角度 1	7.8	118.4	0.48
角度 2	8.3	78.4	0.49
角度 3	10.5	148.6	0.49
角度 4	10.6	57.6	0.48
角度 5	13.9	162.0	0.49
角度 6	14.8	43.0	0.49
角度 7	16.9	168.7	0.52

将准备好的数据集分别降采样 2 倍、3 倍、4 倍生成 3 组低分辨率多角度影像, 并进行了 3 组模拟实验验证了 MVSR-Net 的有效性。

为了验证本文方法的有效性, 实验中将其与双三次插值方法 (Bicubic)、基于 CNN 网络的超分辨率重建方法 (super-resolution CNN,

SRCNN)^[30]、基于 3D 卷积的视频超分辨率重建方法(3D super-resolution net,3DSRnet)^[31]方法进行比较。其中,SRCNN 方法首次将深度学习技术应用于影像超分辨率重建,SRCNN 网络的输入为单框幅影像;3DSRnet 方法的网络结构包含两个分支,一个分支用来学习影像的高频信息,另一个分支使用双三次插值方法获取分辨率增强后的影像,将两个分支结果相加即可获取最终超分辨率重建影像,该方法同样使用多框幅影像作为输入来获取超分辨率重建影像。此外,实验中加入了解融实验来探究动态上采样滤波模块对超分辨率重建结果的影响。本文采用峰值信噪比(peak signal-to-noise ratio ,PSNR)以及结构相似性测度(structural similarity ,SSIM)^[32-34]对实验结果进行定量评价。

2.2 模拟实验结果及分析

为了验证本文提出方法的有效性,设计了 3 组模拟实验,分别为 2 倍、3 倍、4 倍超分辨率重建实验,输入的影像分别通过原始高分辨率多角度影像降采样 2 倍、3 倍、4 倍获得。3 组模拟实验的定量评价结果如表 2 所示,最优结果加粗表示,次优结果标注下划线。由表 2 可知,MVSR-Net 方法得到了最高的 PSNR 值以及 SSIM 值;3 倍超分辨率重建实验结果中,MVSR-Net 方法的 PSNR 值相比 3DSRnet 方法的 PSNR 值提高了 0.731 0 dB;2 倍超分辨率重建实验中,输入数据通过将高分辨率原始影像降采样至原始大小的 1/2 而得到,由于下采样倍数较小,相比 3 倍、4 倍超分辨率重建实验,数据保留了最多的空间细节信息,MVSR-Net 方法也得到了更好的结果,相比 3DSR-net 方法,PSNR 值提高了 2.688 8 dB。

表 2 超分辨率重建定量结果
Tab.2 Quantitative Evaluation of Different Super-Resolution Methods

方法	2 倍重建实验		3 倍重建实验		4 倍重建实验	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
Bicubic	30.086 6	0.942 3	26.384 0	0.938 6	24.412 0	0.925 2
SRCNN	30.960 5	0.991 0	27.047 0	0.981 2	<u>24.903 9</u>	<u>0.935 0</u>
3DSRnet	<u>31.413 9</u>	<u>0.992 5</u>	<u>27.550 7</u>	<u>0.982 5</u>	19.086 6	0.925 2
MVSR-Net	34.102 7	0.993 5	28.281 7	0.989 0	26.794 2	0.985 3

图 4 展示了不同方法 2 倍超分辨率重建实验的目视结果,由图 4 可以看出,MVSR-Net 方法得到了最好的重建结果;Bicubic 方法和 SRCNN 方法没有考虑多角度影像角度维的信息,因此重建影像比较模糊,细节信息没有得到有效恢复;3DSRnet 方法在网络中考虑了影

像角度维信息,在重建的定量结果以及目视结果上都有一定的改善;MVSR-Net 方法加入了动态上采样滤波模块,可以更好地顾及多角度影像存在的分辨率差异以及角度变化引起的地物偏移等现象,因此得到了最好的定量评价结果以及目视结果。

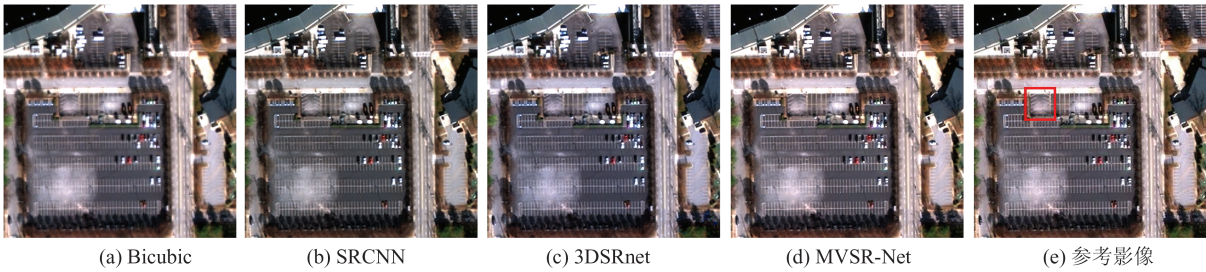


图 4 超分辨率重建结果比较(重建倍数=2)
Fig.4 Comparison of the Super-Resolution Reconstruction Results (Factor=2)

图 5 展示了图 4(e)中红框区域的重建结果。从图 5 中可以看出,MVSR-Net 的重建结果更清晰且相对比较平滑;Bicubic、SRCNN 的重建结果有较明显的马赛克效应;3DSRnet 方法的重建结果有一定改善,但仍然比较模糊。

图 6 展示了不同方法 3 倍超分辨率重建实验的目视结果,图 7 展示了图 6(e)中红框区域的重建结果,影像中的地物包含公路以及林地地区。由图 6、7 可知,Bicubic 方法存在严重的马赛克效应,影像的大量细节信息丢失;SRCNN 方法较

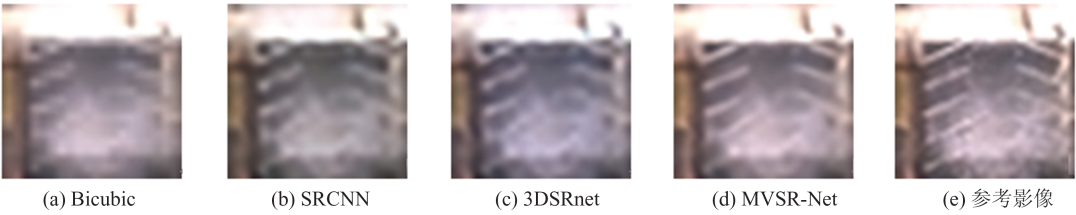


图5 图4(e)红框区域超分辨率重建结果局部放大图

Fig.5 Local Amplification of Super-Resolution Reconstruction Results of the Red Box Area in Fig.4(e)

Bicubic方法有一定的改善,但是在影像的具体细节恢复上存在人工伪痕的问题;3DSRnet和MVSR-Net方法在影像的细节上都有一定的改善,但是MVSR-Net可以恢复更多的细节纹理信息。可见,本文提出的MVSR-Net方法无论在空间细节的恢复上还是光谱的保持上,都得到了最佳的结果

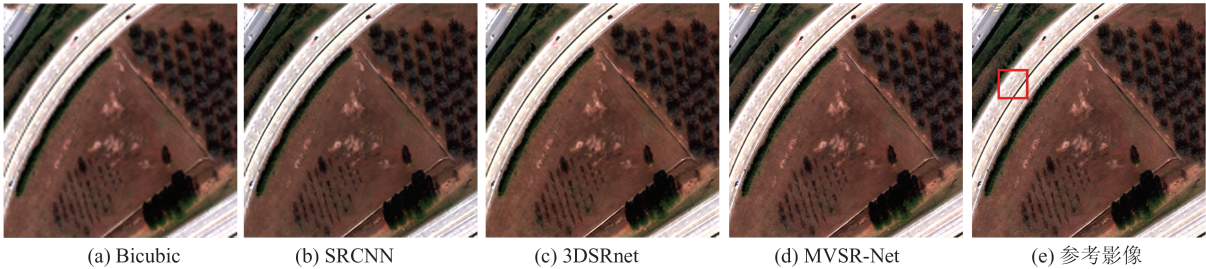


图6 超分辨率重建结果比较(重建倍数=3)

Fig.6 Comparison of the Super-Resolution Reconstruction Results (Factor=3)

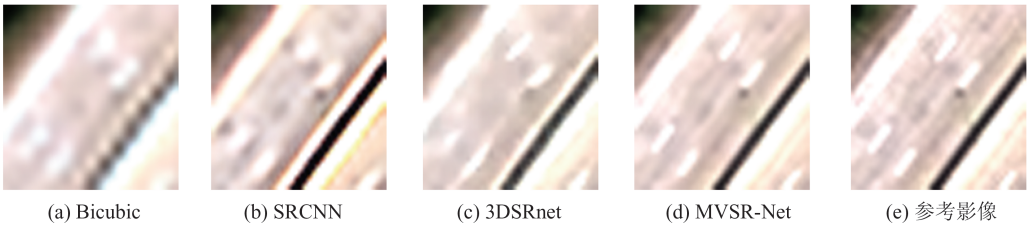


图7 图6(e)红框区域超分辨率重建结果局部放大图

Fig.7 Local Amplification of Super-Resolution Reconstruction Results of the Red Box Area in Fig.6(e)

图8展示了不同方法4倍超分辨率重建实验的目视结果,图9展示了图8(e)中红框区域的重建结果,实验数据主要是机场区域的影像。由图8、9以及表2可知,Bicubic方法、SRCNN方法以及3DSRnet方法都存在比较明显的细节纹理信息的模糊现象;MVSR-Net方法细节纹理信息也得到了修复,定量评价结果和目视结果最好。

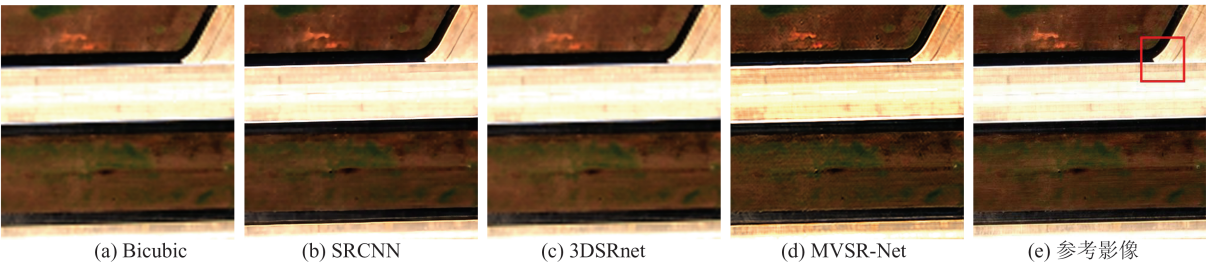


图8 超分辨率重建结果比较(重建倍数=4)

Fig.8 Comparison of the Super-Resolution Reconstruction Results (Factor=4)

综上所述,本文提出的方法相比其他对比方法在细节及纹理的修复上有更好的表现力。在Bicubic方法中,影像随着放大倍数的增大愈加模糊,影像的细节纹理信息重建都受到较大影响;SRCNN方法在目视结果上较Bicubic方法有提升,说明基于深度学习的SRCNN网络对于高分辨率遥感影像的超分辨率重建有一定的效果;3DSRnet方法由于加入了多角度影像角度维信

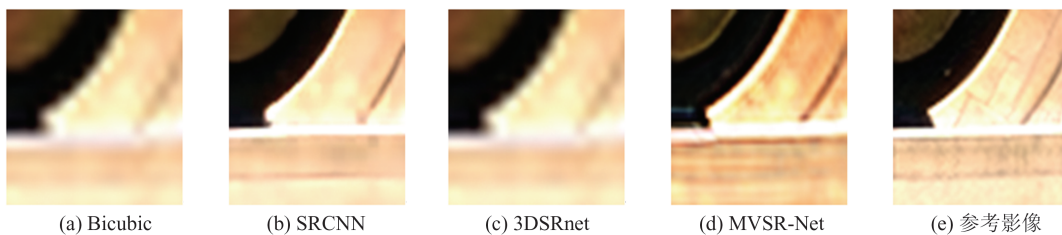


图9 图8(e)中超分辨率重建结果局部放大图

Fig.9 Local Amplification of Super-Resolution Reconstruction Results of the Red Box Area in Fig.8(e)

息,与SRCNN方法及Bicubic方法相比,重建结果在细节上得到了改善。虽然3DSRnet网络使用3D-CNN对多角度影像信息进行了融合,但3DSRnet没有顾及由于观测角度变化而引起的影像地物位移、变形等问题。而本文方法相比3DSRnet方法进一步改进了影像高频信息提取的网络结构,并且针对不同角度影像间存在的地物偏移等现象在网络中加入了动态上采样滤波模块,因此MVSR-Net可以有效利用多角度影像空间维信息以及角度维信息,进而在细节信息纹理信息的重建上有较大提升。实验通过3组不同重建倍数的超分辨率重建实验证明该网络可以有效提升多角度遥感影像的分辨率,并且得到较好的目视结果和定量评价结果。

2.3 真实实验结果及分析

为了验证MVSR-Net网络的泛化性,本文使用另一套多角度遥感影像数据对MVSR-Net网络的性能进行测试。真实实验中同样设计了3组放大倍数,分别为2倍、3倍和4倍超分辨率重建。真实实验的数据选自2012年数据融合大赛中提供的WorldView-2多角度遥感影像^[33],该数据包含巴西里约热内卢地区5个角度观测影像,分别为 45.3° 、 34° 、 8.6° 、 -30.2° 和 -45.4° 观测影像,其中观测角度绝对值越小,意味着越接近正射影像。本文从该数据分别裁取了大小为 192×192 、 128×128 和 96×96 的3幅子影像,分别使用Bicubic、SRCNN、3DSRnet和MVSR-Net网络对其进行超分辨率重建,并进行了定性及定量分析。实验中的网络设置与模拟实验中网络设置保持一致。定量评价指标选取无参考锐化因子(cumulative probability of blur detection, CPBD)^[34]来评价真实实验的重建结果,CPBD指标通常用来衡量重建影像的清晰度。不同倍数重建影像的结果如图10所示,为了增加影像对比度,本文选取影像的近红外波段、红波段与绿波段来显示重建结果。定量评价结果如表3所示,表中最优结果加粗表示,次优结果标注下划线。

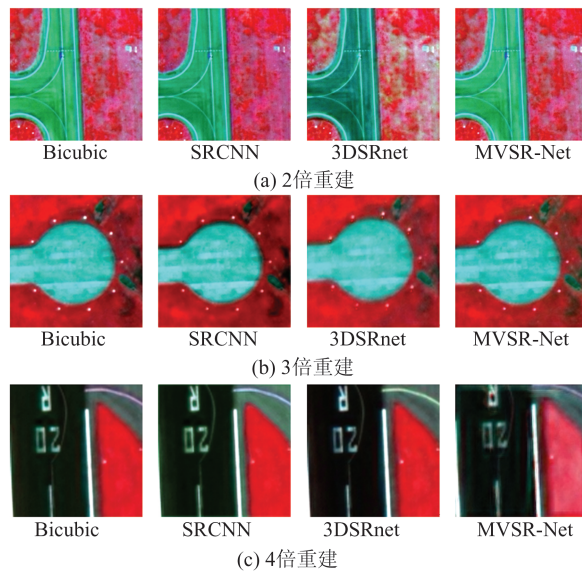


图10 真实实验超分辨率重建结果比较

Fig.10 Comparison of the Super-Resolution Reconstruction Results in Real Experiment

表3 不同超分辨率重建方法CPBD分析

Tab.3 CPBD of Different Super-Resolution Reconstruction Methods

方法	2倍重建	3倍重建	4倍重建
Bicubic	0.101 8	0.022 3	0.018 5
SRCNN	0.129 2	0.151 1	0.065 9
3DSRnet	<u>0.160 2</u>	<u>0.225 1</u>	0.261 4
MVSR-Net	0.183 1	0.837 2	<u>0.252 7</u>

图10(a)为4种方法的2倍超分辨率重建结果。从图10(a)中可以看出,影像的细节信息都得到了较好的保留;SRCNN方法和3DSRnet方法存在不同程度的光谱畸变效应;MVSR-Net方法保留了纹理细节的同时较好地重建了光谱信息。图10(b)为3倍超分辨率重建结果,覆盖区域为机场部分影像。由图10(b)可知,3DSRnet方法的重建结果出现了人工伪痕现象。图10(c)为4倍超分辨率重建结果,由图10(c)可知,SRCNN方法出现了一定程度的光谱畸变效应;MVSR-Net方法在地物标志细节的恢复上有一定程度的模糊。结合表3可知,MVSR-Net方法在2倍重

建和3倍重建上都表现出了相对较好的结果,而在4倍重建结果上3DSRnet方法表现更佳。综上所述,MVSR-Net方法总体上可以更有效地保留影像重建的细节,并较好地保留影像的光谱信息,而3DSRnet方法在光谱信息的保持上存在较大问题。

2.4 消融实验

为了进一步说明 MVSR-Net 方法的有效性,本文设计了消融实验对 MVSR-Net 网络结构中的动态上采样模块进行测试。将 MVSR-Net 的动态上采样模块去掉,使用 3DSRnet 中的双三次插值进行了替换,残差学习的部分仍保留在网络中,重建倍数设定为 2 倍。消融实验的定量评价结果如表 4 所示,由表 4 可知,加入了动态上采样模块的网络比未加入动态上采样模块的网络得到了更高的 PSNR 值以及 SSIM 值。图 11 展示了实验的目视结果,由图 11 可知,加入动态上采样滤波模块的结果有更清晰的细节信息。综上所述,加入动态上采样滤波模块的网络可以更好地融合多角度影像信息,从而得到更好的超分辨率重建结果。

2.5 多角度影像数据分析

为了更好地探究实验中多角度影像观测角度数目对于模型以及实验结果的影响,本文设置了模拟实验来比较多角度遥感影像观测角度数目对于实验的影响。在该模拟实验中,实验数据

集同样为 WorldView-2 亚特兰大地区多角度观测影像。实验数据中训练数据和测试数据的数量与 §2.1 中数据的设置相同,输入数据角度数量分别为 3 个角度、5 个角度和 7 个角度的多角度遥感影像,输入数据大小分别为 30×30 、 40×40 、 60×60 ,对应 4 倍、3 倍和 2 倍超分重建网络结构。MVSR-Net 网络的配置与 §1.4 中的网络配置相同,批处理数量设置为 12。3 组实验的定量评价结果如表 5 所示。

表 4 动态上采样模块消融实验分析

Tab.4 Analysis of Ablation Experiment on DUF Module

评价参数	MVSR-Net (去掉 DUF)	MVSR-Net
PSNR/dB	30.126 4	34.102 7
SSIM	0.945 6	0.993 5

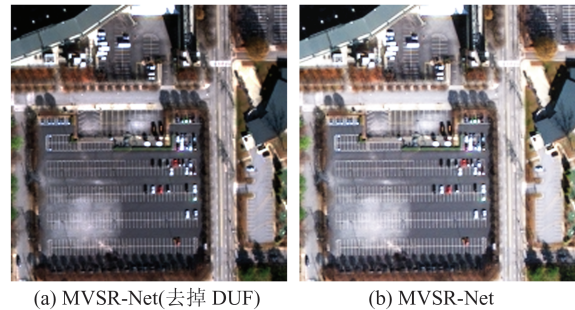


图 11 消融实验对比结果

Fig.11 Comparison of the Ablation Experiment Results

表 5 多角度影像数据定量分析

Tab.5 Quantitative Evaluation of the Different Multiview Datasets

输入数据角度数量	2 倍重建实验		3 倍重建实验		4 倍重建实验	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
3 个角度	30.193 7	0.993 7	26.594 3	0.988 8	24.103 8	0.963 4
5 个角度	31.121 9	0.993 9	26.656 0	0.989 1	24.247 3	0.974 5
7 个角度	31.248 9	0.994 7	26.764 3	0.991 9	24.354 4	0.975 7

由表 5 可知,随着参与训练的角度数目的增加,2 倍、3 倍以及 4 倍重建定量评价结果都呈现出增加的趋势。由此可知,参加训练的多角度影像观测角度越多,得到的重建结果越好。

2.6 改进三维卷积的 MVSR-Net

三维卷积可以有效提取数据角度维和时间维的特征,然而三维卷积大大增加了网络的参数量以及运算负担。文献[35]针对三维卷积计算效率低下的问题提出了快速时空三维卷积,并将该改进模块应用于视频超分辨率重建中。为了进一步提升网络效率,本文对动态上采样超分辨率重建网络中的三维卷积模块进行了改进,并将快速时

空三维卷积引入网络,将网络中的三维卷积替换为快速时空三维卷积。改进后的快速时空卷积如图 12 所示。由图 12 可知,原本的 $k \times k \times k$ 三维卷积由串联的 $1 \times k \times k$ 卷积和 $k \times 1 \times 1$ 卷积所替代。改进后的三维卷积计算量大幅下降,并且很大程度降低了 MVSR-Net 网络的参数量。改进三维卷积的 MVSR-Net 网络的参数量以及每秒所执行的浮点运算次数(floating-point operations per second, FLOPs)见表 6。由表 6 可知,网络参数量得到了 25.13% 的改进, FLOPs 得到了 30.48% 的改进。

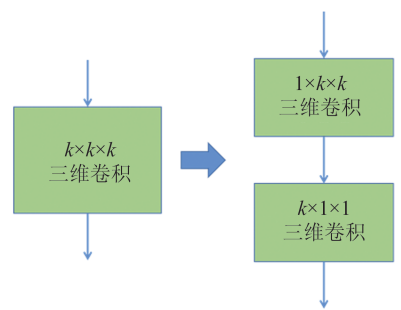


图 12 快速时空卷积

Fig.12 Fast Spatiotemporal Convolution

表 6 网络参数量及运算量比较

Tab.6 Comparison of Parameters and FLOPs of Different Nets

方法	参数量/ 10^6	FLOPs/ 10^9
MVSR-Net	1.91	5.02
改进三维卷积的 MVSR-Net	1.43	3.49

3 结 语

本文提出了一种面向多角度遥感影像的基于深度学习框架的端到端超分辨率重建方法。该方法的网络结构通过两个分支网络提取影像的高频信息并获取重建影像。本文提出的方法通过 WorldView-2 美国亚特兰大地区和巴西里约热内卢地区多角度遥感影像数据集得到了验证。

通过模拟实验和真实实验证明, MVSR-Net 可以得到精度较高的超分辨重建结果。通过消融实验、角度数目对比实验可以得到以下结论: (1) 在网络结构中加入动态上采样滤波模块可以进一步提高超分辨率重建精度; (2) 增加多角度遥感影像观测角度数目可以提高重建结果精度; (3) 将网络中 3D 卷积替换为快速时空卷积也可以降低网络参数量并提高运算效率。后续研究中, 将进一步提高算法的泛化能力和效率, 并且能够融合更多角度的信息。

参 考 文 献

[1] Kang Junmei, Sui Lichun, Li Li, et al. Super-Resolution Reconstruction of Remote Sensing Images Based on Double-Sparse K-SVD Dictionary Learning[J]. *Computer Engineering and Applications*, 2018, 54 (16): 187-191 (康军梅, 隋立春, 李丽, 等. 基于双稀疏度 K-SVD 字典学习的遥感影像超分辨率重建[J]. *计算机工程与应用*, 2018, 54 (16): 187-191)

[2] Zhou Ziyong. Super-Resolution Reconstruction of Remote Sensing Images by Using Empirical Mode

Decomposition and Compressed Sensing [J]. *Remote Sensing Technology and Application*, 2018, 33 (1): 96-102 (周子勇. 运用经验模态分解和压缩感知方法进行遥感影像超分辨率重建[J]. *遥感技术与应用*, 2018, 33(1): 96-102)

[3] Chan J C W, Ma J L, Kempeneers P, et al. Super Resolution Enhancement of Hyperspectral CHRIS/Proba Images with a Thin-Plate Spline Nonrigid Transform Model [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, 48 (6) : 2 569-2 579

[4] Galbraith A E, Theiler J, Thome K J, et al. Resolution Enhancement of Multilook Imagery for the Multispectral Thermal Imager [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, 43 (9): 1 964-1 977

[5] Ma J L, Chan J, Canters F. An Operational Super-resolution Approach for Multi-Temporal and Multi-angle Remotely Sensed Imagery [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2012, 5(1): 110-124

[6] Li F, Jia X P, Fraser D, et al. Super Resolution for Remote Sensing Images Based on a Universal Hidden Markov Tree Model [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, 48 (3) : 1 270-1 278

[7] Shen H, Ng M K, Li P, et al. Super-Resolution Reconstruction Algorithm to MODIS Remote Sensing Images [J]. *The Computer Journal*, 2008, 52(1) : 90-100

[8] Akgun T, Altunbasak Y, Mersereau R M. Super-Resolution Reconstruction of Hyperspectral Images [J]. *IEEE Transactions on Image Processing*, 2005, 14(11): 1 860-1 875

[9] Merino M T, Nunez J. Super-Resolution of Remotely Sensed Images with Variable-Pixel Linear Reconstruction [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, 45(5): 1 446-1 457

[10] Tsai R. Multiframe Image Restoration and Registration [J]. *Advance Computer Visual and Image Processing*, 1984, 1: 317-339

[11] Schultz R R, Stevenson R L. Extraction of High-Resolution Frames from Video Sequences [J]. *IEEE Transactions on Image Processing*, 1996, 5 (6): 996-1 011

[12] Stark H, Oskoui P. High-Resolution Image Recovery from Image-Plane Arrays, Using Convex Projections [J]. *Journal of the Optical Society of America A, Optics and Image Science*, 1989, 6(11): 1 715-1 726

- [13] Ur H, Gross D. Improved Resolution from Subpixel Shifted Pictures [J]. *CVGIP: Graphical Models and Image Processing*, 1992, 54(2): 181-186
- [14] Tom B C, Katsaggelos A K. Reconstruction of a High-Resolution Image by Simultaneous Registration, Restoration, and Interpolation of Low-Resolution Images[C]//International Conference on Image Processing, Washington DC, USA, 1995
- [15] Tom B C, Katsaggelos A K, Galatsanos N P. Reconstruction of a High Resolution Image from Registration and Restoration of Low Resolution Images [C]//The 1st International Conference on Image Processing, Austin, Texas, USA, 1994
- [16] Elad M, Feuer A. Restoration of a Single Super-resolution Image from Several Blurred, Noisy, and Undersampled Measured Images[J]. *IEEE Transactions on Image Processing*, 1997, 6(12): 1 646-1 658
- [17] Li L, Wang W, Luo H, et al. Super-Resolution Reconstruction of High-Resolution Satellite ZY-3 TLC Images[J]. *Sensors*, 2017, 17(5): 1 062
- [18] Ma J L, Chan J C W, Canters F. Robust Locally Weighted Regression for Superresolution Enhancement of Multi-angle Remote Sensing Imagery [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, 7(4): 1 357-1 371
- [19] Zhang H Y, Yang Z Y, Zhang L P, et al. Super-Resolution Reconstruction for Multi-angle Remote Sensing Images Considering Resolution Differences [J]. *Remote Sensing*, 2014, 6(1): 637-657
- [20] Chen H, Zhang H Y, Du J, et al. Unified Framework for the Joint Super-Resolution and Registration of Multiangle Multi/Hyperspectral Remote Sensing Images [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13: 2 369-2 384
- [21] Bi Xiaojun, Pan Mengdi. Super-Resolution Reconstruction of Airborne Remote Sensing Images Based on the Generative Adversarial Networks[J]. *CAAI Transactions on Intelligent Systems*, 2020, 15(1): 74-83 (毕晓君, 潘梦迪. 基于生成对抗网络的机载遥感图像超分辨率重建[J]. 智能系统学报, 2020, 15(1): 74-83)
- [22] Wang Zhi, Li Anyi, Fang Jinxiong. Super Resolution Reconstruction of Remote Sensing Images Based on Dense Convolution Neural Network [J]. *Geomatics & Spatial Information Technology*, 2020, 43(8): 4-8 (王植, 李安翼, 方锦雄. 基于密集卷积神经网络的遥感影像超分辨率重建[J]. 测绘与空间地理信息, 2020, 43(8): 4-8)
- [23] Molini M A, Valsesia D, Fracastoro G, et al. Deep-SUM: Deep Neural Network for Super-Resolution of Unregistered Multitemporal Images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(5): 3 644-3 656
- [24] Deudon M, Kalaitzis A, Goytom I, et al. HighRes-Net: Recursive Fusion for Multi-Frame Super-Resolution of Satellite Imagery[J]. *ArXiv Preprint*, 2020: 2002.06460
- [25] Huang G, Liu Z, van der Maaten L, et al. Densely Connected Convolutional Networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, Hawaii, USA, 2017
- [26] Ji S P, Zhang C, Xu A J, et al. 3D Convolutional Neural Networks for Crop Classification with Multi-temporal Remote Sensing Images [J]. *Remote Sensing*, 2018, 10(2): 75
- [27] Li Y, Zhang H K, Shen Q. Spectral - Spatial Classification of Hyperspectral Imagery with 3D Convolutional Neural Network[J]. *Remote Sensing*, 2017, 9(1): 67
- [28] Jia X, De Brabandere B, Tuytelaars T, et al. Multi-Dynamic Bayesian Networks [M]. Boston: The MIT Press, 2007
- [29] 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, Nevada, USA, 2016
- [30] 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): 295-307
- [31] Kim S Y, Lim J, Na T, et al. 3DSRnet: Video Super-Resolution Using 3D Convolutional Neural Networks[J]. *ArXiv Preprint*, 2018: 1812.09079
- [32] Xu Lina, Xiao Qi, He Luxiao. Fused Image Quality Assessment Based on Human Visual Characteristics [J]. *Geomatics and Information Science of Wuhan University*, 2019, 44(4): 546-554 (许丽娜, 肖奇, 何鲁晓. 考虑人类视觉特征的融合图像评价方法[J]. 武汉大学学报·信息科学版, 2019, 44(4): 546-554)
- [33] Pacifici F, Qian D. Foreword to the Special Issue on Optical Multiangular Data Exploitation and Outcome of the 2011 GRSS Data Fusion Contest [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2012, 5(1): 3-7
- [34] Narvekar N D, Karam L J. A No-Reference Perceptual Image Sharpness Metric Based on a Cumulative

- Probability of Blur Detection [C]//International Workshop on Quality of Multimedia Experience, San Diego, California, USA, 2009
- [35] Li S, He F X, Du B, et al. Fast Spatio-Temporal Residual Network for Video Super-Resolution[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, California, USA, 2019

Multi-angle Remote Sensing Images Super-Resolution Reconstruction Using Dynamic Upsampling Filter Deep Network

CHEN Hang¹ LUO Bin¹

¹ State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

Abstract: Objectives: Deep learning based on super-resolution reconstruction technology has been widely used in multi-temporal hyperspectral images, high-resolution image reconstruction. Multi-angle remote sensing images have rich complementary information, which is suitable for super-resolution reconstruction. **Methods:** An end-to-end super-resolution reconstruction method based on dynamic upsampling filter network is proposed for high-resolution multi-angle remote sensing images. The network of the method includes an end-to-end two-way network, in which one branch is used dynamic upsampling filter block to improve the image resolution. Another branch network is used to learn the high-frequency information in the image. In order to verify the effectiveness of the proposed method, 2, 3 and 4 times super-resolution reconstruction simulation experiments and real experiments are carried out with Worldview-2 multi-angle remote sensing images from Atlanta, America and Rio de Janeiro, Brazil, respectively. Several groups of comparative experiments are carried out. **Results and Conclusions:** Experimental results show that the proposed method can effectively improve the spatial resolution of the target image while taking into account the angle dimension information of multi angle images, and effectively maintain the image details.

Key words: super-resolution reconstruction; multi-angle remote sensing images; deep learning; dynamic upsampling filter

First author: CHEN Hang, PhD candidate, specializes in the theories and methods of the high resolution remote sensing images and multi-angle remote sensing images processing. E-mail: hangchen@whu.edu.cn

Foundation support: The National Natural Science Foundation of China (61871298); Foundation Frontier Project of Wuhan (2020010601012184); the Natural Science Foundation of Qinghai Province (2020-ZJ-927).

引文格式: CHEN Hang, LUO Bin. Multi-angle Remote Sensing Images Super-Resolution Reconstruction Using Dynamic Upsampling Filter Deep Network[J]. Geomatics and Information Science of Wuhan University, 2021, 46(11): 1716-1726. DOI: 10.13203/j.whugis20200651 (陈行, 罗斌. 利用动态上采样滤波深度网络进行多角度遥感影像超分辨率重建[J]. 武汉大学学报·信息科学版, 2021, 46(11): 1716-1726. DOI: 10.13203/j.whugis20200651)