

doi: 10.6046/zrzyg.2021179

引用格式: 孔爱玲, 张承明, 李峰, 等. 基于知识引导的遥感影像融合方法[J]. 自然资源遥感, 2022, 34(2): 47–55. (Kong A L, Zhang C M, Li F, et al. Knowledge-based remote sensing image fusion method[J]. Remote Sensing for Natural Resources, 2022, 34(2): 47–55.)

基于知识引导的遥感影像融合方法

孔爱玲¹, 张承明¹, 李峰², 韩颖娟³, 孙焕英⁴, 杜漫飞⁴

(1. 山东农业大学信息科学与工程学院, 泰安 271018; 2. 山东省气候中心, 济南 250031;
3. 中国气象局旱区特色农业气象灾害监测预警与风险管理重点实验室,
银川 750002; 4. 航天宏图信息技术股份有限公司, 北京 100195)

摘要: 遥感图像融合技术能够将含有互补信息的多源图像进行融合, 从而得到内容更丰富、光谱质量更高的图像, 是遥感应用的关键和基础。针对遥感影像融合过程中容易出现的光谱失真和空间结构失真问题, 该文以注意力机制为基础, 利用归一化植被指数(normalized difference vegetation index, NDVI)和归一化水指数(normalized difference water index, NDWI)作为先验知识, 构建了基于知识引导的遥感影像融合模型(remote sensing image FuseNet, RS-FuseNet)。首先, 基于高通滤波能够充分提取边缘纹理细节的优势, 构建了高通滤波模块提取全色图像高频细节信息; 其次, 提取多光谱图像的 NDVI 和 NDWI 信息; 再次, 构建自适应挤压激励(squeeze-and-excitation, SE)模块, 对输入特征进行重标定; 最后, 将自适应 SE 模块与卷积单元结合, 对输入特征进行融合处理。以高分六号遥感影像作为数据源, 选择施密特正交化(gram-schmidt, GS)、主成分分析(principal component analysis, PCA)、深度网络结构的图像融合网络(a deep network architecture for pan-sharpening, PanNet)、以卷积神经网络为基础的融合网络(pansharpening by convolutional neural networks, PNN)模型作为对比模型开展实验, 实验结果表明: 该文提出模型的峰值信噪比(peak signal to noise ratio, PSNR)指标(40.5)和结构相似性(structural similarity, SSIM)指标(0.98)均优于对比模型, 表明这一方法在遥感影像图像融合方面具有明显的优势。

关键词: 图像融合; 先验知识; 卷积神经网络; 注意力机制; 高分六号

中图分类号: TP 79 **文献标志码:** A **文章编号:** 2097-034X(2022)02-0047-09

0 引言

目前, 由于传感器技术的限制, 单一传感器还不能直接获取到同时具有高光谱分辨率、高空间分辨率的图像, 卫星遥感平台上一一般同时安装有两台传感器, 分别用于获取高空间分辨率的全色图像和多光谱图像^[1]。在应用时, 通常需要先利用图像融合技术对全色图像和多光谱图像进行融合, 以获取同时具有高光谱分辨率和高空间分辨率的图像^[2]。

传统的图像融合方法主要有成分替代(component substitution, CS)^[3]、多分辨率分析(multiresolution analysis, MRA)^[4]和稀疏表示(sparse representa-

tion, SR)^[5]。CS 融合方法首先将多光谱图像(multispectral, MS)转化到另一个空间, 将空间结构和光谱信息分离为不同的成分; 然后, 将转化后的 MS 图像中具有空间结构的分量替换为全色图像(panchromatic, PAN)。经典的基于 CS 的融合方法包括强度-色调-饱和度(intensity-hue-saturation, IHS)融合方法^[6]、主成分分析(principal component analysis, PCA)融合方法^[7]和施密特正交化方法(gram-schmidt, GS)^[8]。基于 CS 的方法可以获得丰富的细节, 但光谱失真往往比较严重。MRA 融合方法的核心是多尺度细节提取和注入。通常采用 MRA 方法从 PAN 图像中提取空间细节, 然后将其注入上采样的多光谱图像中。广泛使用的 MRA 方

收稿日期: 2021-06-03; 修订日期: 2021-09-16

基金项目: 山东省自然科学基金“基于多源遥感影像的冬小麦精细空间分布提取方法”(编号: ZR2021MD097)、山东省自然科学基金“基于知识融合的遥感影像分割方法研究”(编号: ZR2020MF130)、中国气象局旱区特色农业气象灾害监测预警与风险管理重点实验室开放基金项目“旱区作物种植信息智能化提取技术研究”(编号: CAMF-202001)、青海省基础 Research 计划“基于遥感图像超分辨率技术的油菜地土壤水分监测”(编号: 2021-ZJ-739)及中国气象局旱区特色农业气象灾害监测预警与风险管理重点实验室指令性项目“基于遥感技术的小麦、水稻不同发育期长势评估”(编号: CAMP-201916)共同资助。

第一作者: 孔爱玲(1996-), 女, 硕士研究生, 主要从事计算机视觉研究。Email: 2019110580@sdau.edu.cn。

通信作者: 张承明(1971-), 男, 博士, 教授, 主要研究方向为遥感信息提取。Email: chming@sdau.edu.cn。

法包括拉普拉斯金字塔^[9]、小波变换^[10,11]、curvelet 变换^[12]、非下采样 contourlet 变换^[13,14]、shearlet 变换^[15]、非下采样 shearlet 变换^[16]。与基于 CS 的方法相比,基于 MRA 的方法将提取到的全色图像的细节信息注入到多光谱图像中能够更好地保持光谱特性。稀疏表示的核心思想是将图像表示为一个过完整字典中最少原子的线性组合,但是该方法比较复杂和费时^[17]。

近年来,基于深度学习的遥感图像融合方法受到了广泛的关注,如以卷积神经网络为基础的融合网络(pansharpening by convolutional neural networks, PNN)^[18],基于卷积神经网络的目标自适应融合网络(target-adaptive CNN-based pansharpening, Target-PNN)^[19],多尺度多深度卷积神经网络(multi-scale and multi-depth convolutional neural network, MSDCNN)^[20],采用深度网络结构的图像融合(a deep network architecture for pan-sharpening, PanNet)^[21],遥感图像融合深度卷积神经网络(remote sensing image fusion with deep convolutional neural network, RSIFNN)^[22],基于卷积自编码器的 MS 融合(convolutional autoencoder-based multispectral image fusion, CAE)^[23]。其中,PNN 网络首次在图像融合任务中使用卷积神经网络,其将插值后的多光谱图像和全色图像拼接起来输入网络中进行端对端的训练,网络直接学习输入和高分辨率图像之间的关系;PanNet 网络将全色图像和多光谱图像的高频细节信息拼接起来输入到残差网络中进行特征提取与融合,然后将提取到的高频细节信息注入上采样后的低分辨率多光谱图像中。与传统的基于 CS 和基于 MRA 的算法相比,基于 CNN 的方法显著提高了图像融合的性能,但这些方法还存在一些问题,主要有:Target-PNN 和 RSIFNN 都缺乏具体的纹理细节处理,导致融合后的图像纹理细节不够清晰。PNN 只是单纯地将全色图像和上采样的多光谱图像输入卷积神经网络进行训练,没有根据全色图像和多光谱图像各自的特点有针对性的进行特征提取和融合,可能会导致一些光谱和空间结构的失真。PanNet 虽然增强了纹理细节,但没有考虑 MS 图像光谱通道之间的关系,可能会导致一些光谱失真。

注意力机制源于对人类视觉的研究,在认知学中,人类会有选择地关注所有信息的一部分,同时忽略其他可见信息。该思想最早应用在自然语言的处理中。后来,研究者提出了注意力机制模型(squeeze-and-excitation networks, SENet),该模型通过对输入的信息进行挤压、激励和特征重标定,对每个特征通道赋予不同的权重^[24]。该网络结构被

广泛应用于目标检测、图像分类等任务中,取得了很好的效果。在图像融合任务中,需要通过神经网络对全色和多光谱图像的所有特征进行融合操作,但是这样会产生大量的冗余信息,如何更准确高效地融合特征信息成为了一个亟待解决的问题。注意力机制则可以通过对特征赋予权重来解决这个问题。

先验知识是人们通过经验或者数理统计等方法得到的对于事物特性的描述,被广泛应用于自然图像去模糊、自然图像去噪中^[25]。常用的自然图像先验知识有自然图像的局部平滑性、非局部自相似性、稀疏性等特征。受自然图像先验知识的启发,考虑到归一化植被指数(normalized difference vegetation index, NDVI)能够反映土地植被的分布信息,归一化水指数(normalized difference water index, NDWI)反映了水体的分布信息,而植被和水体是能够影响光谱的两种主要信息。所以,我们利用这两种信息作为先验知识,对遥感影像的融合过程进行约束,可以进一步提高融合效果,改善光谱失真的问题。

基于上述分析,本文提出一种新的图像融合模型——基于知识引导的遥感影像融合模型(remote sensing image FuseNet, RSFuseNet)来改善融合过程中的光谱失真和空间结构失真问题。本文在 RSFuseNet 模型中引入了 SENet 网络中的挤压激励(squeeze-and-excitation, SE)结构,通过对图像的特征进行挤压、激励操作学习每个特征通道的权重,将学到的权重赋予图像在每个特征通道,以此来有针对性地抑制无用信息,增强有用信息,减少了冗余信息在网络中的传递和积累,提高了融合精度,缩短了模型的训练时间。根据全色图像与多光谱图像的特点引入先验知识——NDVI 和 NDWI 来对融合过程中的光谱信息进行约束。提高了融合过程中光谱信息的保真度,解决了图像融合过程中容易产生的光谱失真问题。

1 研究数据

1.1 遥感数据及其预处理

本文选择山东省泰安市的两幅高景一号影像,以及其对应的高分六号影像制作数据集。每幅高景一号影像包含了全色波段和多光谱波段,全色波段的分辨率为 0.5 m,多光谱波段的分辨率为 2 m。每景高分六(GF-6)影像包含多光谱波段(蓝、绿、红、近红外)和全色波段,空间分辨率分别为 8 m 和 2 m,表 1 给出了本文所使用的高分六号卫星的主要参数。

表 1 高分 6 号卫星的主要参数
Tab.1 Main parameters of GF-6 satellite

相机类型	谱段号	谱段范围/ μm	星下点地面 像元分辨率	覆盖宽度	静态转函	量化值	辐射定标精度
离轴 TMA 全反射式	P	0.45 ~ 0.90	优于 2 m	大于 90 km	优于 0.2	12 bit	绝对定标精度 优于 7%, 相对 定标精度优于 3%
	B1	0.45 ~ 0.52	优于 8 m				
	B2	0.52 ~ 0.60					
	B3	0.43 ~ 0.69					
	B4	0.76 ~ 0.90					

图像预处理主要包括几何纠正、辐射校正等步骤。本文利用航天宏图信息技术股份有限公司生产的新一代遥感与地理信息一体化软件(pixel information expert,PIE)对高分六号卫星的图像以批处理的方式进行预处理,包括:大气校正、几何纠正、正射校正。经过预处理后,得到空间分辨率为 8 m 的多光谱图像和空间分辨率为 2 m 的全色图像,其中,多光谱图像包含蓝、绿、红、近红外 4 个波段,全色图像包含一个波段。在遥感图像处理平台(the environment for visualizing images,ENVI)中对高景一号多光谱影像进行辐射定标、大气校正及正射校正,得到分辨率为 2 m 的多光谱影像。

1.2 数据集制作

通过 1.1 节中所述步骤,得到高分六号的全色与多光谱图像及高景一号的多光谱图像,将高分六号全色图像裁剪成 512 × 512 大小的 TIF 格式的图片,图片命名时采用在后缀的前面添加_pan。将对应的高分六号多光谱图像裁剪成 128 × 128 大小的

TIF 格式的图片,图片命名时采用在后缀的前面添加_lrms。将对应的高景一号的多光谱影像裁剪成 512 像素 × 512 像素大小的 TIF 格式的图片,图片命名时采用在后缀的前面添加_target。将裁剪得到的数据的 80% 作为训练样本,20% 的数据作为测试样本,分别放在 train 和 test 文件夹下,即得到实验所用的数据集。

2 研究方法

2.1 模型结构

RSFuseNet 模型的输入数据为训练样本中的高分六号全色、多光谱图像,高景一号的多光谱图像。其中,高分六号的全色图像和多光谱图像用来提取特征并进行特征融合;高景一号的多光谱图像是参考图像,用来计算模型的损失。模型结构如图 1 所示,主要包括特征提取模块、自适应 SE 模块以及结合自适应 SE 的卷积模块。

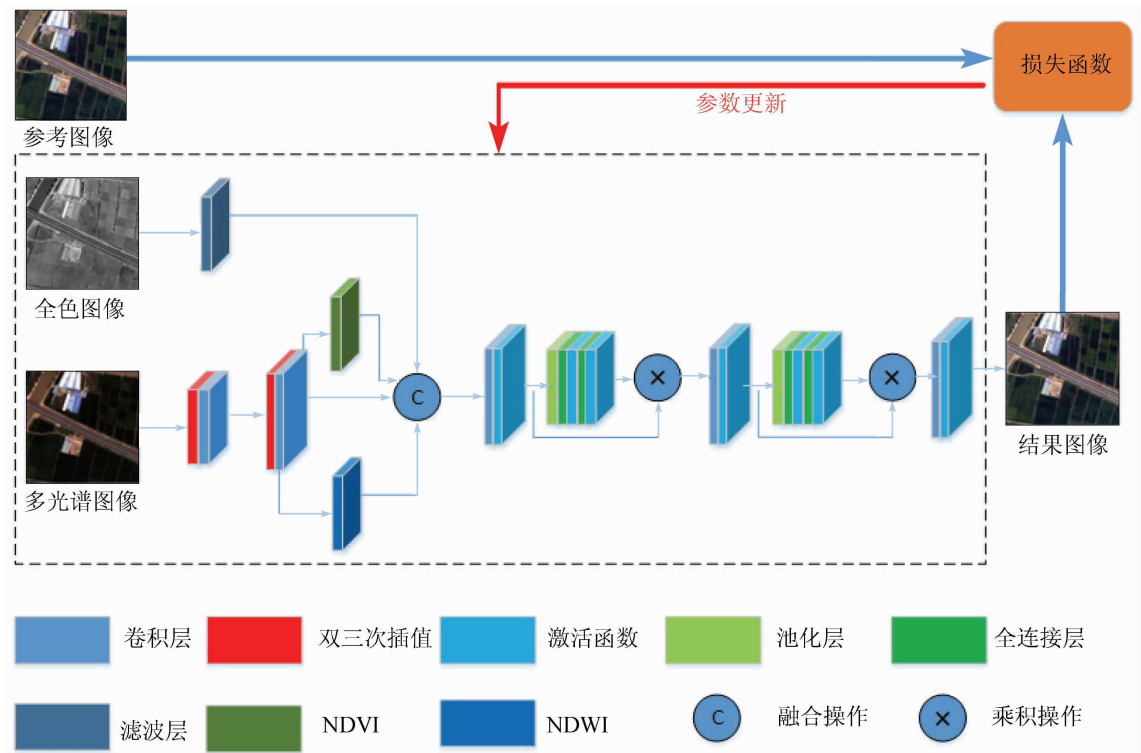


图 1 RSFuseNet 模型基本结构图

Fig.1 Basic structure of the remote sensing image FuseNet

2.1.1 特征提取模块

在进行图像融合前, 全色图像要经过高通滤波器, 提取高频细节信息, 过滤掉冗余的噪声信息。构建高通滤波器时, 首先用均值滤波器获取图像的低频信息, 然后用原始图像减去低频信息即得到高频信息。获取全色图像的高频信息可以减少全色图像中的噪声信息对多光谱图像中光谱信息的影响, 充分发挥全色图像中高频纹理细节的优势, 进而使融合图像的光谱信息和纹理信息达到最优的效果。

多光谱图像输入模型后首先要进行上采样以达到与全色图像相同的分辨率。鉴于双三次插值法可以充分保持多光谱图像的光谱信息, 故采用双三次插值进行上采样。考虑到全色图像的分辨率是多光谱图像分辨率的 4 倍, 遂采用分两次上采样来提升多光谱图像的分辨率, 每一次在原来的基础上提升两倍。上采样模块由一个双三次插值层和一个 1×1 卷积的调整层组成。在上采样后引入先验知识, 即通过 NDVI 和 NDWI 的计算公式分别计算出上采样后多光谱图像的每个像素点的 NDVI 和 NDWI 的值, 这些值分别组成 2 个矩阵, 与上采样后的多光谱图像一起作为下一层卷积的输入。NDVI 和 NDWI 的计算公式为:

$$NDVI = \frac{NIR - R}{NIR + R}, \quad (1)$$

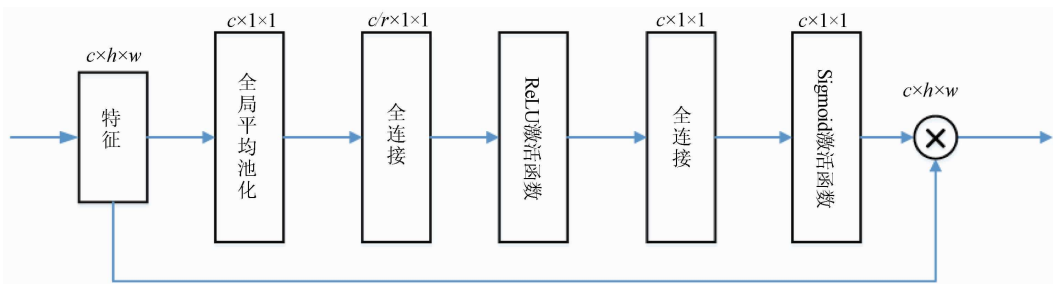


图 2 自适应 SE 模块结构图

Fig. 2 The structure diagram of adaptive squeeze-and-excitation module

2.1.3 结合自适应 SE 的卷积模块

RSFuseNet 模型的影像融合部分采用结合自适应 SE 的卷积模块。该模块主要包括 3 个卷积单元, 第一个和第二个卷积单元组成相同, 都是由一个卷积层、一个激活函数层和一个自适应 SE 模块组成; 第三个卷积单元由一个卷积层和一个激活函数层组成。该模块输入为经过高通滤波提取的全色图像的高频纹理信息、上采样后分辨率与全色图像相同的多光谱图像和经过逐像素计算得到的 NDVI 和 NDWI。这样, 第一个卷积单元一共输入 7 个波段, 故该单元卷积层的输入通道为 7, 卷积核大小为 9×9 , 步长大小为 1, 填充大小为 4, 输出特征通道大小为 48。第二个卷积单元卷积层的输入通道为 48, 卷积

$$NDWI = \frac{G - NIR}{G + NIR}, \quad (2)$$

式中, NIR , R 和 G 分别为多光谱图像近红外波段、红色波段和绿色波段的反射率。

2.1.2 自适应 SE 模块

该模块输入的特征图大小为 $c \times h \times w$, 其中 c 为通道数, h 为高度, w 为宽度。该模块主要由一个全局平均池化层, 两个全连接层, 一个 ReLU 激活函数层和一个 Sigmoid 激活函数层组成。全局平均池化层将输入的特征图压缩为 $c \times 1 \times 1$ 的大小, 也就是每一个通道得到一个值, 这个值具有全局的感受野, 它表征着在特征通道上响应的全局分布。第一个全连接层将全局平均池化层中得到的结果降维到 $c/r \times 1 \times 1$ 的大小, 这里的 r 采用上一层卷积的通道数也就是 c , 这也是该模块叫作“自适应”SE 模块的原因, 即降维所用的参数能够随着输入特征通道数的变化而变化, 通过这种“自适应”机制可以极大地减少参数量和计算量。ReLU 层对第一层全连接层的结果进行激活操作, 第二个全连接层将特征升回到原始的维度。通过全连接层的降维和升维以及 ReLU 的激活操作可以增加模型的非线性, 更好地拟合通道间复杂的相关性。Sigmoid 激活函数获得每个特征通道的权重。最后将获得的权重乘到每个通道的特征上, 也就是图中的 $\times$ 操作。具体见图 2。

核大小为 5×5 , 步长大小为 1, 填充大小为 2, 输出通道的大小为 32。第三个卷积单元卷积层的输入通道为 32, 卷积核大小为 5×5 , 步长大小为 1, 填充大小为 2, 输出通道大小为 4。由于数据输入模型时首先对其进行归一化操作, 故卷积层后面没有使用归一化层。

2.2 训练模型

2.2.1 模型损失

由于图像融合任务属于回归问题, 所以模型以均值平方差为基础定义损失函数, 均值平方差损失函数定义为:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i, j) - K(i, j)\|^2, \quad (3)$$

式中: $I(i, j)$ 为目标图像第 (i, j) 个像素的值, 即真值; $K(i, j)$ 为训练完成后结果图像第 (i, j) 个像素的值; m 为图像的宽度; n 为图像的高度。

根据 MSE 函数, 定义 RSFuseNet 模型的损失函数为 $loss$, 计算公式为:

$$loss = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (L_{\text{spectral}}(i, j) + L_{\text{spatial}}(i, j))^2, \quad (4)$$

$$L_{\text{spectral}}(i, j) = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|R(i, j) - \text{upLR}(i, j)\|, \quad (5)$$

$$L_{\text{spatial}}(i, j) = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|R(i, j) - \text{HRMS}(i, j)\|, \quad (6)$$

式中: $L_{\text{spectral}}(i, j)$ 为融合图像光谱部分的损失值; $L_{\text{spatial}}(i, j)$ 为融合图像结构部分的损失值; $R(i, j)$ 为参考图像的值; $\text{upLR}(i, j)$ 为上采样后的多光谱图像的值; $\text{HRMS}(i, j)$ 为融合后得到的高分辨多光谱图像的值。

2.2.2 训练方法

模型采用端对端的方式进行训练, 使用 $loss$ 作为损失函数, 使用随机梯度下降算法 (stochastic gradient descent, SGD) 作为优化算法, 具体训练步骤为:

- 1) 确定训练过程中的超参数, 并对模型的参数进行初始化。
- 2) 将制作出的训练图像 (包括全色, 多光谱以及参考图像) 作为训练数据输入到模型。
- 3) 使用模型对当前的训练数据进行一次前向计算。
- 4) 使用 $loss$ 损失函数进行计算损失。
- 5) 使用随机梯度下降算法, 对模型的参数进行更新, 完成一次训练过程。
- 6) 重复步骤 3)—5), 直到损失函数小于指定的期望值或者损失值不再减小。

3 实验

3.1 实验设计

本文以 pytorch 框架为实验平台, 以 python 语言对模型进行编码实现。使用一台服务器开展对比实验, 实验使用的操作系统为 Ubuntu16.04, 为了加速数据的运行, 装载 GPU TITAN X, 并安装 CUDA 和 cuDNN 加速。实验使用的数据集在 1.2 节已经进行了说明, 本实验采用 1 850 组图像作为训练样本、422 组图像作为测试样本, 训练样本和测试样本互

不交叉。每一组图像包括全色图像、多光谱图像、参考图像。由于全色和多光谱图像的空间分辨率之比为 1:4, 参考图像与全色图像的空间分辨率相同, 所以在这些样本中, 全色图像和参考图像的尺寸为 512 像素 × 512 像素, 多光谱图像的尺寸为 128 像素 × 128 像素。这些数据均由高分六号卫星的 PMS 传感器获得, 经过预处理得到。另外, 本文实验得到的结果图像与高分六号全色图像具有相同的空间分辨率, 均为 2 m, 图像尺寸为 512 像素 × 512 像素。

本文将 RSFuseNet 模型与 5 种常见的图像融合方法进行比较, 对比方法分别是 GS 变换、PCA、最邻近插值融合方法 (nearest neighbor diffusion, NNDiffuse)、PNN 和 PanNet。前 3 种是传统的图像融合算法, PNN 和 PanNet 是深度学习的图像融合算法 (表 2)。

表 2 对比实验使用的模型

Tab. 2 Models used in the comparison experiment

名称	模型介绍
GS 变换	该融合方法通过统计分析方法对参与融合的各波段进行最佳匹配
PCA	该融合方法将多光谱图像的多个波段看作多维数据集, 通过对该数据集进行 PCA 变换得到多个主成分分量
NNDiffuse	图像融合算法, 主要是采用最邻近插值
PNN	与 RSFuseNet 具有类似的结构, 使用 3 层卷积提取特征
PanNet	使用两路分支的卷积神经网络提取特征

3.2 客观评价指标

客观评价指标可以定量地评价融合图像的质量情况, 尤其是在图像的一些微小的细节, 仅仅依靠人眼很难发现图像的优劣, 而采用客观评价指标, 则可以通过数值的大小定量地评价图像的好坏。因此, 本文选择峰值信噪比 (peak signal to noise ratio, PSNR)^[26]、结构相似性指数 (structural similarity, SSIM)^[27] 和光谱映射角 (spectral angle mapper, SAM)^[28] 3 种评价指标对融合结果进行定量评价, 从不同的方面对图像的质量进行评估。

PSNR 主要用来评价图像质量的敏感性误差, 是衡量两幅图像差别的重要指标, 其计算公式为:

$$PSNR = 10 \lg \left(\frac{(2^n - 1)^2}{MSE} \right), \quad (7)$$

$$MSE = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W (F(i, j) - R(i, j))^2, \quad (8)$$

式中: H 和 W 分别是图像的高度和宽度; F 和 R 分别为融合结果图像和参考图像。PSNR 的值越大, 表明融合图像从原始图像获得的信息量越多, 图像和原始图像相似度越高, 融合效果越好。

SSIM 从亮度 (均值)、对比度 (方差) 和结构 3

个层次中比较图像的失真程度,取值范围为 $[0,1]$,值越大越好,其计算公式为:

$$SSIM(F, R) = \frac{2\mu_F\mu_R + C_1}{\mu_F^2 + \mu_R^2 + C_1} \cdot \frac{2\sigma_{FR} + C_2}{\sigma_F^2 + \sigma_R^2 + C_2}, \quad (9)$$

式中: μ_F 为 F 的均值; μ_R 为 R 的均值; σ_F^2 为 F 的方差; σ_R^2 为 R 的方差; σ_{FR} 为 F 和 R 的协方差; $C_1 = (k_1L)^2$, $C_2 = (k_2L)^2$ 为2个常数,避免除0; L 为像素值的范围; $k_1 = 0.01$, $k_2 = 0.03$ 为默认值。

SAM是可以计算两幅图像在每个像素处的光谱相似度的一种指标。其值较低则表示较低的光谱失真和较好的图像融合质量。其计算公式为:

$$SAM(F_{\{j\}}, R_{\{j\}}) = \arccos\left(\frac{(F_{\{j\}}, R_{\{j\}})}{\|F_{\{j\}}\| \|R_{\{j\}}\|}\right), \quad (10)$$

式中: $F_{\{j\}}$ 和 $R_{\{j\}}$ 分别为 F 和 R 的第 j 个像素向量, $(F_{\{j\}}, R_{\{j\}})$ 为两个向量的内积, $\|\cdot\|$ 为向量的 $L2$ 范数。

3.3 实验结果

图3—图6分别给出了使用NNDiffuse, GS变换、PCA, PanNet, PNN和RSFuseNet方法对测试样本进行融合得到的结果图。其中a列为用NNDiffuse融合的结果图,b列为用GS变换融合得到的结果图,c列为用PCA融合的结果图,d列为用PanNet模型融合得到的结果图,e列为PNN模型的融合结果图,f列为用RSFuseNet模型得到的融合结果图。这些图像中包含了建筑、农田、大棚、道路、河流等地物,具有一定的代表性。

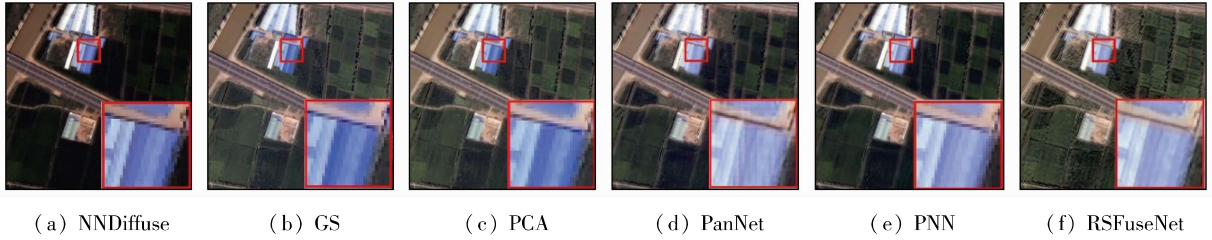


图3 区域一对比实验结果
Fig.3 Comparison of experimental model results

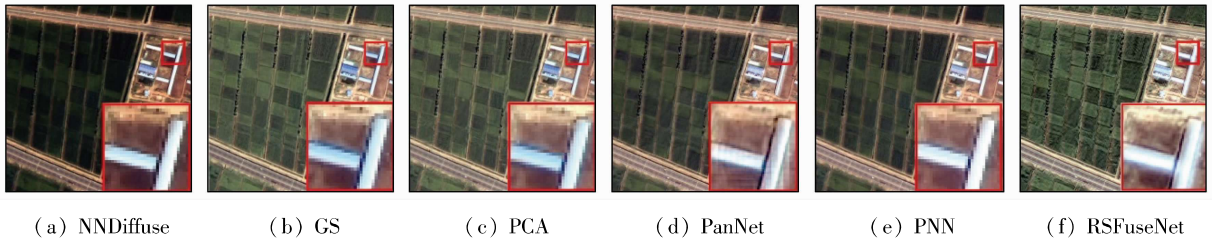


图4 区域二对比实验结果
Fig.4 Comparison of experimental model results

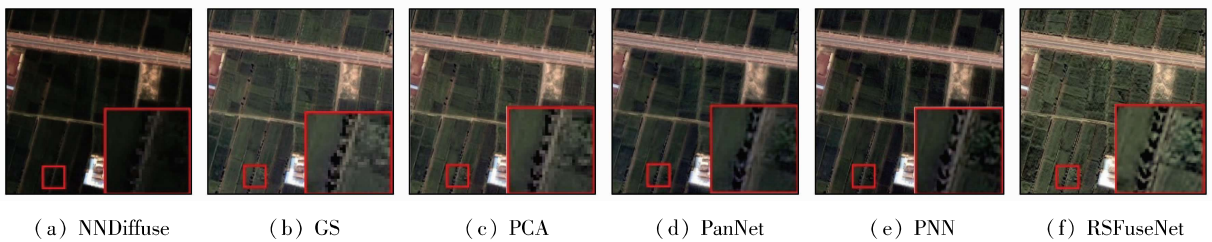


图5 区域三对比实验结果
Fig.5 Comparison of experimental model results

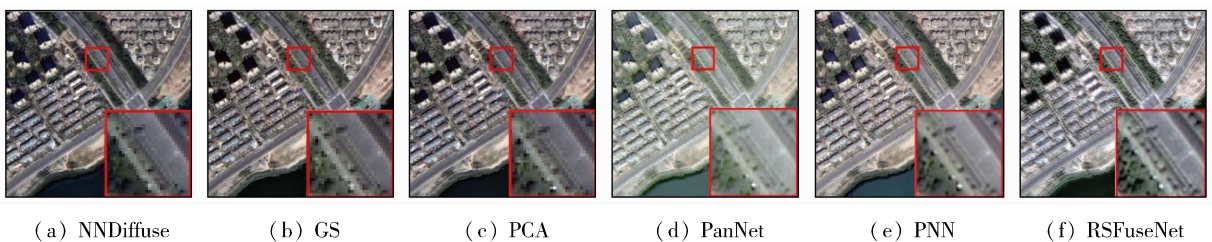


图6 区域四对比实验结果
Fig.6 Comparison of experimental model results

综合图3—图6可以看出,a列融合图像颜色偏深,尤其是在有植被的地方,植被的不同区域展现出

的不同颜色在融合图像中无法具体区分。b列和c列图像还是会存在部分光谱失真的现象,d列、e列、

f列图像在区分不同区域的植被颜色上都有了改进,可以区分出不同区域不同的颜色。在图像的纹理特征方面,d列和e列要比f列稍逊色一些。在图像的水体特征方面,通过图6可以看出,与f列图像相比,d列和e列图像在水体的颜色方面存在着明显的光谱失真,a列、b列和c列图像存在着轻微的光谱失真,这也进一步说明了引入NDWI信息可以改善水体的光谱失真现象。仔细观察上述图像可以发现,f列图像的结果在图像的边缘部分更清晰,比如在房子的边缘、大棚的边缘以及不同作物的边缘,f列图像得到的结果纹理更明显,d列和e列则有一些模糊。这说明了利用卷积神经网络进行图像融合的方法是优于传统的图像融合方法的,而且不同的卷积神经网络算法对融合结果的影响不同。通道注意力层在图像的融合过程中发挥了作用,能够通过学习每个通道的特征信息并根据其重要程度赋予其权重,这样,权重低的噪声信息在图像融合的过程中就会被相应的抑制,从而使融合结果更精确。

各个模型的定量评价结果如表3所示,表中向上的箭头表示该指标的值越高,融合效果越好,向下的箭头表示该指标的值越低,融合效果越好。可以看到RSFuseNet模型的3个评价指标的值均优于对比模型。

表3 对比模型性能指标表

Tab.3 Compare the models performance metrics

指标	NNDiffuse	GS	PCA	PNN	PanNet	RSFuseNet
PSNR ↑	21.13	24.26	26.16	35.50	28.02	40.50
SSIM ↑	0.92	0.92	0.90	0.97	0.91	0.98
SAM ↓	4.62	3.68	3.57	3.71	4.22	2.09

通过图3—图6和表3的分析,可以发现不论是主观视觉感受还是在客观评价指标,本文构建的RSFuseNet模型都比其他方法表现出更好的效果。

4 讨论

4.1 自适应SE对结果精度的影响

使用自适应SE模块的目的,是为了对特征重标定,从而达到根据特征的重要程度进行特征融合的目的。在本研究开展的对比实验中,RSFuseNet, PanNet和PNN都是通过对输入的特征进行卷积操作达到特征融合的效果。不同的是,RSFuseNet在融合的过程中采用了自适应SE模块对特征进行重标定,而PanNet和PNN在融合的过程中没有进行特征的重标定操作。

对比RSFuseNet模型和PanNet模型、PNN模型的融合结果,可以发现PanNet模型生成的图像还普

遍存在着棋盘化的现象;PNN模型的结果图像在光谱上还是存在着一定的失真。RSFuseNet模型融合的结果图像光谱信息较好,细节较丰富。尤其是在图像中建筑的边缘、道路、农田中不同作物的边缘以及大棚的边缘,RSFuseNet的融合结果比PanNet, PNN的融合结果具有更加清晰的纹理。

造成这种现象的原因,主要是因为RSFuseNet模型中使用了自适应SE模块。该模块通过特征重标定增强有用的特征,抑制无用的特征,从而能够有针对性地进行图像融合,减少了冗余信息对融合结果的影响。

4.2 先验知识对结果精度的影响

目前,在图像分类,图像分割任务中需要根据先验知识对训练图像做一个真值图,以此来告诉深度学习模型哪些类型的地物是农田,哪些地物类型是建筑、小麦、玉米、水体、道路等等。在图像融合中无法细致地对每一类地物做出标记,但是可以利用遥感影像便于计算的NDVI和NDWI属性充当先验信息来进行图像融合。因此,使用NDVI和NDWI这两种先验知识的目的,是为了在模型中添加一些限制条件,达到对图像融合约束的作用。在本研究开展的对比实验中,只有RSFuseNet模型使用了先验知识NDVI和NDWI。

对比RSFuseNet模型和GS,PCA,PanNet,PNN及NNDiffuse,可以发现NNDiffuse得到的结果图像光谱失真比较严重,图像颜色发暗,难以区分不同地物。GS和PCA得到的结果图像存在着少许的光谱失真。PanNet和PNN模型的结果图也存在着轻微的光谱失真。而RSFuseNet模型的结果图像则几乎没有出现光谱失真现象,光谱信息更丰富,图像质量更高。

5 结论

针对遥感影像融合过程中容易出现的光谱失真和空间结构失真问题,本文利用注意力机制在特征提取方面的优势和先验知识的特性,建立了一种适用于高分六号遥感影像的融合模型。该模型利用注意力机制对图像的特征通道赋予重要性,并根据重要程度对特征进行后续的处理;同时在融合过程中利用先验知识对特征进行逐像素约束,使得融合结果的光谱特征和纹理特征都取得了较高的保真度。

本文的不足之处在于本文构建的模型是有监督模型,但是当前图像融合的参考图像难以获得,融合结果相比理想结果可能还是存在一些差距。针对这个问题考虑在下一步的工作中根据遥感影像的特点

设计一个无监督的网络,解决参考图像难以获得的问题。

致谢:感谢山东农业大学高性能计算中心的支持。

参考文献 (References):

- [1] 王建宇,王跃明,李春来. 高光谱成像系统的噪声模型和对辐射灵敏度的影响[J]. 遥感学报,2010,14(4):607-620.
Wang J Y, Wang Y M, Li C L. Noise model of hyperspectral imaging system and influence on radiation sensitivity[J]. Journal of Remote Sensing, 2010, 14(4):607-620.
- [2] Vivone G, Alparone L, Chanussot J, et al. A critical comparison among pansharpening algorithms[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(5):2565-2586.
- [3] Xie B, Zhang H, Huang B. Revealing implicit assumptions of the component substitution pansharpening methods[J]. Remote Sensing, 2017, 9(5):443-458.
- [4] Yong Y, Wan W, Huang S, et al. A novel pan-sharpening framework based on matting model and multiscale transform[J]. Remote Sensing, 2017, 9(4):391.
- [5] 王晓艳,刘勇,蒋志勇. 一种基于结构相似度的 IHS 变换融合算法[J]. 遥感技术与应用, 2011, 26(5):670-676.
Wang X Y, Liu Y, Jiang Z Y. An IHS fusion method based on structural similarity[J]. Remote Sensing Technology and Application, 2011, 26(5):670-676.
- [6] Yang Y, Wan W, Huang S, et al. Remote sensing image fusion based on adaptive IHS and multiscale guided filter[J]. IEEE Access, 2017(4):4573-4582.
- [7] W Dong, S Xiao, X Xue, et al. An improved hyperspectral pansharpening algorithm based on optimized injection model[J]. IEEE Access, 2019(7):16718-16729.
- [8] Xu L, Zhang Y, Gao Y, et al. Using guided filtering to improve gram-schmidt based pansharpening method for GeoEye-1 satellite images[J]. Proceedings of the 4th International Conference on Information Systems and Computing Technology, 2016(64):33-37.
- [9] Aiazzi B, Alparone L, Baronti S, et al. MTF-tailored multiscale fusion of high-resolution MS and Pan imagery[J]. Photogrammetric Engineering & Remote Sensing, 2015, 72(5):591-596.
- [10] Vivone G, Restaino R, Mura M D, et al. Contrast and error-based fusion schemes for multispectral image pansharpening[J]. IEEE Geoscience and Remote Sensing Letters, 2013, 11(5):930-934.
- [11] Alparone L, Luciano, Baronti, et al. Spatial methods for multispectral pansharpening: Multiresolution analysis demystified[J]. IEEE Transactions on Geoscience & Remote Sensing, 2016(54):2563-2576.
- [12] Nencini F, Garzelli A, Baronti S, et al. Remote sensing image fusion using the curvelet transform[J]. Information Fusion, 2007, 8(2):143-156.
- [13] Lu X, Zhang J, Zhang Y. An improved non-subsampled contourlet transform-based hybrid pan-sharpening algorithm[C]//2017 IEEE International Geoscience and Remote Sensing Symposium. Fort Worth: IEEE, 2017:3393-3396.
- [14] James R, Vadivel M. Improvement of spectral and spatial information using modified WAMM and modified bi-cubic interpolation method in non-subsampled contourlet transform domain[C]//2015 International Conference on Robotics, Automation, Control and Embedded Systems (RACE). Chennai India: IEEE, 2015:1-5.
- [15] Zhang Z, Luo X, Wu X. A new pansharpening method using statistical model and shearlet transform[J]. IEEE Technical Review, 2014, 31(5):308-316.
- [16] Wang X, Bai S, Li Z, et al. The PAN and MS Image pansharpening algorithm based on adaptive neural network and sparse representation in the NSST domain[J]. IEEE Access, 2019(7):52508-52521.
- [17] Yin H T. Sparse representation based pan-sharpening with details injection model[J]. Signal Processing, 2015(113):218-227.
- [18] Giuseppe M, Davide C, Luisa V, et al. Pansharpening by convolutional neural networks[J]. Remote Sensing, 2016, 8(7):594.
- [19] Scarpa G, Vitale S, Cozzolino D. Target-adaptive CNN-based pansharpening[J]. IEEE Transactions on Geoscience & Remote Sensing, 2018, 56(9):5443-5457.
- [20] Yuan Q, Wei Y, Meng X, et al. A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(3):978-989.
- [21] Yang J, Fu X, Hu Y, et al. PanNet: A deep network architecture for pan-sharpening[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice Italy: IEEE, 2017:5449-5457.
- [22] Shao Z, Cai J. Remote sensing image fusion with deep convolutional neural network[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(5):1656-1669.
- [23] Azarang A, Manoochehri H E, Kehtarnavaz N. Convolutional autoencoder-based multispectral image fusion[J]. IEEE Access, 2019, 7:35673-35683.
- [24] Jie H, Li S, Gang S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8):2011-2023.
- [25] 聂文昌. 融合先验知识的深度学习目标识别与定位研究[D]. 哈尔滨:哈尔滨工程大学, 2019.
Nie Wenchang. Research on target recognition and location based on prior knowledge using deep learning[D]. Harbin: Harbin Engineering University, 2019.
- [26] Pushparaj J, Hegde A V. Evaluation of pan-sharpening methods for spatial and spectral quality[J]. Applied Geomatics, 2016, 9(1):1-12.
- [27] 杨丽萍,马孟,谢巍,等. 干旱区 Landsat8 全色与多光谱数据融合算法评价[J]. 国土资源遥感, 2019, 31(4):11-19. doi: 10.6046/gtzyyg. 2019. 04. 02.
Yang L P, Ma M, Xie W, et al. Fusion algorithm evaluation of Landsat 8 panchromatic and multispectral images in arid regions[J]. Remote Sensing for Land and Resources, 2019, 31(4):11-19. doi:10.6046/gtzyyg. 2019. 04. 02.
- [28] Vivone G, Alparone L, Chanussot J, et al. A critical comparison among pansharpening algorithms[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(5):2565-2586.

Knowledge – based remote sensing image fusion method

KONG Ailing¹, ZHANG Chengming¹, LI Feng², HAN Yingjuan³, SUN Huanying⁴, DU Manfei⁴

(1. *College of Information Science and Engineering, Shandong Agricultural University, Taian 271018, China*; 2. *Shandong Provincial Climate Center, Jinan 250031, China*; 3. *Key Laboratory for Meteorological Disaster Monitoring and Early Warning and Risk Management of Characteristic Agriculture in Arid Regions, CMA, Yinchuan 750002, China*; 4. *PIESAT Information Technology Co., Ltd., Beijing 100195, China*)

Abstract: The remote sensing image fusion technology can combine multi – source images containing complementary information to obtain images with richer content and higher spectral quality, thus it is the key and foundation of remote sensing applications. Aiming at the problems of spectral distortion and spatial structure distortion that are prone to occur in the process of remote sensing image fusion, the knowledge – based remote sensing image FuseNet (RSFuseNet) was constructed based on the attention mechanism and using normalized difference vegetation index (NDVI) and normalized difference water index (NDWI) as prior knowledge. Firstly, considering that the high – pass filtering can fully extract edge texture details, a high – pass filtering module was constructed to extract high – frequency details of panchromatic images. Secondly, NDVI and NDWI were extracted from multi – spectral images. Then, an adaptive squeeze – and – excitation (SE) module was constructed to recalibrate the input features. Finally, the adaptive SE module was combined with the convolution unit to perform fusion processing on the input features. The experiment was conducted using Gaofen 6 remote sensing image as the data source, and selecting Gram – Schmidt (GS) transformation, principal component analysis (PCA), a deep network architecture for pan – sharpening (PanNet), and pansharpening by convolutional neural networks (PNN) models as comparative models. The experimental results show that the peak signal to noise ratio (PSNR) index (40.5) and the structural similarity (SSIM) index (0.98) of the RSFuseNet model are better than those of comparative models, indicating that the method in this study has obvious advantages in remote sensing image fusion.

Keywords: image fusion; prior knowledge; convolutional neural network; attention mechanism; GF – 6

(责任编辑: 李 瑜)