

doi: 10.6046/gtzyyg.2020209
引用格式: 韩彦岭,崔鹏霞,杨树瑚,等. 基于残差网络特征融合的高光谱图像分类[J]. 国土资源遥感,2021,33(2):11-19.
(Han Y L,Cui P X,Yang S H,et al. Classification of hyperspectral images based on feature fusion of residual network [J]. Remote Sensing for Land and Resources ,2021,33(2):11-19.)

基于残差网络特征融合的高光谱图像分类

韩彦岭,崔鹏霞,杨树瑚,刘业锟,王 静,张 云
(上海海洋大学信息学院,上海 201306)

摘要:深度学习技术因其在深度挖掘地物特征方面的独特优势为高光谱图像分类提供了技术手段。但是在高光谱图像的像素级地物分类中,由于样本输入尺寸的影响导致深度学习的层数受限,不能充分挖掘高光谱图像中的深度特征,为此提出基于残差网络特征融合的高光谱图像分类方法。首先通过主成分分析(principal component analysis,PCA)方法提取原始高光谱图像中的第一主成分,利用残差网络有效提取地物空谱特征;再通过反卷积算法实现特征图的扩充,将反卷积后不同维度的特征进行多尺度特征融合,充分挖掘高光谱图像中的深度特征信息,进一步提升高光谱图像分类精度。对“珠海一号”卫星拍摄的江苏太湖和安徽巢湖两个区域进行地物分类实验,结果表明,与其他方法相比,该方法有效解决了高光谱图像分类中深度特征提取不足的问题,获得了更好的分类性能。
关键词:反卷积;特征融合;残差网络;高光谱图像分类
中图法分类号: TP 79 **文献标志码:** A **文章编号:** 1001-070X(2021)02-0011-09

0 引言

高光谱图像是若干个光谱波段组成的三维立体图,含有丰富的空间信息和光谱信息,并且具有较高的空谱分辨率,能提取到不同地物的细节特征,在农作物监测、海洋环境监测、地图绘制等方面发挥着重要作用^[1],因此利用高光谱图像进行地物分类识别已经成为国内外研究热点。近年来很多分类方法被应用于高光谱图像分类中,包括支持向量机(support vector machine,SVM)^[2]分类法、决策树分类法、最大似然分类法等,但是这些传统分类方法由于只能提取到高光谱图像的浅层特征信息而忽略了深层特征信息导致其分类精度的进一步提高受到限制,另外,高光谱图像数据量巨大,且存在较多冗余信息,也给传统分类方法的处理提出了挑战。

由于深度学习方法可以较好地提取图像的深层特征信息,近年来逐渐受到重视并广泛应用于高光谱图像分类领域,常用的网络模型有卷积神经网络(convolutional neural networks,CNN)、深度置信网络(deep belief nets,DBN)、循环神经网络(recurrent neural network,RNN)等,比如文献[3]将传统的深

度 CNN 进行改进,采用分层特征融合的方法,得到更加有效的深度特征;文献[4]基于反卷积提取高光谱图像的高层特征,提出了一种结合高层特征空间和迁移学习网络的遥感地物图像分类算法,有效提高了分类精度;文献[5]基于由光谱特征和纹理特征作为浅层特征,尺度不变特征变换(scale-invariant feature transform,SIFT)经编码后得到的中层特征,以及 CNN 提取的深层特征,基于不同层次特征选用不同的分类器进行分类。研究表明利用深度学习一定程度上可以有效获得深层特征信息,但是还不能充分融合深度学习网络中各个层之间的深层特征,需要结合其他算法使层与层之间的深层特征得以充分利用。另外,在像素级分类中,高光谱遥感图像的训练样本输入尺寸偏小,深度学习模型的层数受到限制,模型优化的可能性较少,一定程度上也限制了高光谱图像分类精度的提高。2015 年 He 等^[6]提出深度残差网络,在 ImageNet 大规模视觉识别竞赛中获得了图像分类和物体识别的优胜,解决了网络层数不足的问题;文献[7]提出一种基于深度残差 CNN 的高光谱遥感数据的方法,提高分类的准确度;文献[8]利用 CNN 提取到空间光谱特征的同时,结合了三层残差网络模块提高了高光谱图像

收稿日期: 2020-07-10; 修订日期: 2020-09-27
基金项目: 十三五“蓝色粮仓科技创新”国家重点研发计划项目(编号: 2019YFD0900805)和国家自然科学基金项目“基于激光测高和高分立体测绘卫星的大范围建(构)筑物灾害损失精细化评估方法研究”(编号: 41871325)共同资助。
第一作者: 韩彦岭(1975-),女,博士,副教授,主要研究方向为高光谱遥感图像分类。Email: ylh@shou.edu.cn。
通信作者: 杨树瑚(1983-),男,博士,讲师,主要研究方向为南极冰盖的演化。Email: shyang@shou.edu.cn。

的分类精度。

基于以上研究,针对高光谱图像分类中目前存在的样本输入小,网络层数有限,导致深层特征无法充分提取的问题,本文提出一种基于残差网络特征融合的高光谱图像分类方法。利用残差网络加深网络层数,解决了高光谱图像分类中由于训练样本输入尺寸偏小导致的网络层数受限问题;另外,通过反卷积算法实现特征图的扩充并进行不同尺度特征的深度融合,充分挖掘高光谱图像中的深度特征信息,进一步提升高光谱图像分类精度。在“珠海一号”卫星拍摄的江苏太湖和安徽巢湖两个区域数据上进行分类实验,以验证本文方法的有效性。

1 基于残差网络特征融合的高光谱图像分类方法

本文方法主要思想是利用残差网络加深网络层数,通过反卷积扩大特征图,并将不同尺度的特征进行特征融合,进一步挖掘高光谱图像中的深层特征,

以此获得较高的分类精度。本文首先对高光谱图像的原始数据进行主成分分析(principal component analysis,PCA),利用PCA将图像的重要特征集中在第一个波段中,处理成二维图像作为样本输入;考虑到深度学习网络层数加深会造成梯度消失,以及网络层数加深会造成参数过大的问题,利用残差网络中的恒等映射减少训练过程中的参数,并且缓解了梯度消失的问题;传统的残差网络只能够一层一层提取特征并输入到下一层残差单元,层与层之间的特征没有被充分地挖掘和利用,限制了高光谱图像分类精度的进一步提升,因此本文利用反卷积扩大特征图,并且通过多尺度融合算法使不同层的特征进行融合,充分挖掘高光谱图像中的深层特征信息以提高分类精度。本文总体架构如图1所示,主要包含3部分,第一部分为数据预处理,第二部分为算法实现部分,包含本文算法以及与其他方法的对比,第三部分为精度验证,通过计算混淆矩阵,得出总体分类精度(overall accuracy,OA)和Kappa系数验证精度的一致性。

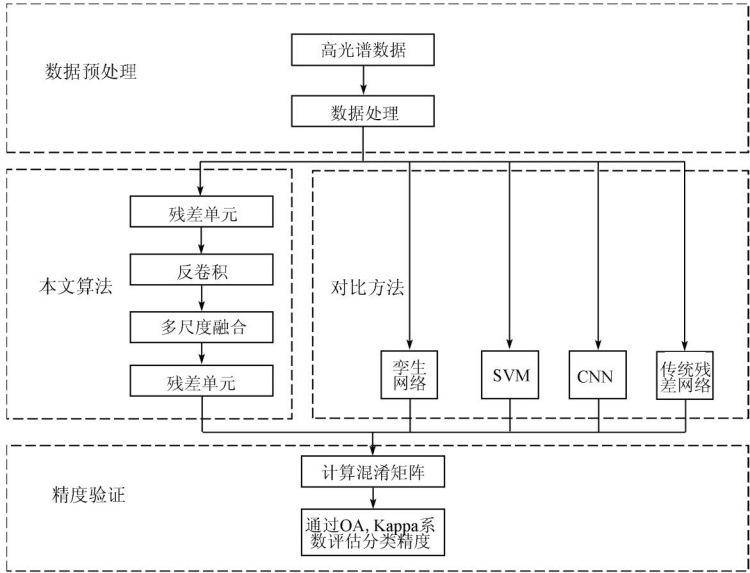


图1 本文总体框架
Fig. 1 Overall framework of this paper

1.1 改进的残差网络

残差网络的恒等映射的设计,使误差不会继续增加,因为可以从 x 恒等映射到 $F(x)$,这就是残差网络解决深度学习中因网络层数增加而导致梯度消失问题的关键所在。如图2所示,把输入 x 传到输出作为初始结果,输出结果为 $Y = F(x) + x$,当 $F(x) = 0$ 时, $Y = x$,也就是恒等映射。于是,残差网络相当于将学习目标改变了,不再是学习一个完整的输出,而是目标值 Y 和 x 的差值,也就是所谓的残差 $F(x) = Y - x$,因此,后面的训练目标就是要将残差结果逼近于0,随着网络加深,准确率不下降。

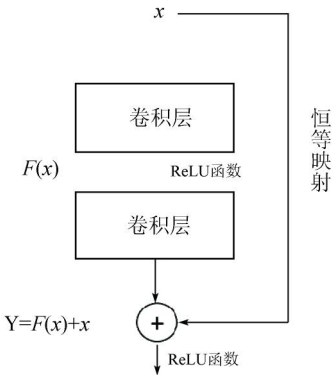


图2 残差网络构思图
Fig. 2 Conceptual diagram of the residual network

图 3 为本文改进的残差网络结构,conv ij 称为残差层,其中 i 表示第 i 个残差层, j 表示对应残差层中的残差单元,如 conv1 j 为第一残差层,conv2 j 为第二残差层,虚线框中称为残差块,包含 6 个残差单元和 2 个池化层,每个残差单元的卷积核大小均为 3×3 ,32 和 64 为卷积核个数。网络中卷积核大小一般为 3×3 , 5×5 , 7×7 ,考虑本文中样本输入大小一般为 27×27 ,若选用 5×5 和 7×7 大小的卷积核,会导致网络模型层数较少,优化网络模型的可能情况更少,因此本文中选取卷积核大小为 3×3 。该网络结构先经过 3 个残差单元和 1 个池化层,将池化后的输出作为反卷积 Dconv1 的输入,反卷积中卷积核大小为 5×5 ,卷积核个数为 64,步长为 1,无填充层,

激活函数为 ReLU。随后将第一个池化层的输出作为第二个残差层的输入,经过三个残差单元和一个池化后的输出作为反卷积 Dconv2 的输入,Dconv2 中除了卷积核大小为 3×3 与 Dconv1 中不同外,其余参数设置均一样。Dconv1 和 Dconv2 输出维度不同的特征进行多尺度融合 (M1),将低层特征扩大融合为同维度的特征后输入到残差块中进行特征提取,其中残差块与第一残差层和第二残差层参数设置完全一样,分别命名为第三残差层和第四残差层,最后通过 conv5 残差层,增加了特征的数量,最后将提取到的特征输入到全连接层 FC 和 softmax 分类器进行分类。

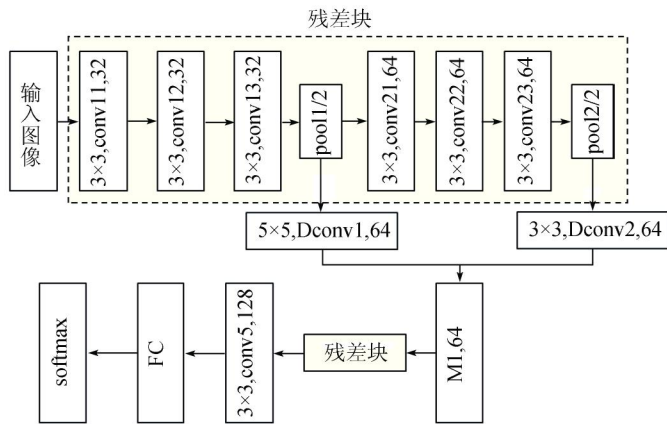


图 3 改进的残差网络结构

Fig. 3 Improved residual network structure

1.2 卷积和反卷积

CNN^[9] 包括一维 CNN、二维 CNN 和三维 CNN。一维 CNN 主要基于光谱特征应用于序列类的数据处理;二维 CNN 主要基于空间信息应用于图像的分类与识别;三维 CNN 基于空谱特征主要应用于医学图像识别等领域。由于二维 CNN 在充分利用空间信息方面具有独特优势和良好的效果,本文采用二维 CNN 进行高光谱图像的特征提取及分类,主要包括输入层、卷积层、池化层和输出层。二维卷积中,第 i 层第 j 个特征图的位置 (x,y) 处的神经元 v_{ij}^{xy} 的计算公式为:

$$v_{ij}^{xy} = f(\sum_m \sum_{p=0}^{p_i-1} \sum_{q=0}^{q_i-1} w_{jm}^{pq} v_{(i-1)m}^{(x+p)(y+q)} + b_{ij}) \quad , \quad (1)$$

式中: m 为索引连接到当前第 j 个特征映射的第 $i-1$ 层中的特征映射; w_{jm}^{pq} 为连接到第 m 个特征映射的位置 (p,q) 的权重; p_i 和 q_i 为第 i 层中的卷积核的长度和宽度; b_{ij} 为第 i 层中第 j 个特征映射的偏差; $f()$ 为激活函数。

本文采用 ReLU 函数作为激活函数,该函数不

存在梯度消失问题,使得模型的收敛速度维持在一个稳定的状态,其公式为:

$$ReLU(x) = \begin{cases} x & x > 0 \\ 0 & x \leq 0 \end{cases} \quad . \quad (2)$$

卷积层的主要作用是提取特征信息,反卷积为卷积的逆运算,反卷积的主要作用是通过信息填充使特征图扩大。在本文算法中,由于不同层的特征图维度不同,不能直接进行融合,需要将高层特征图扩大。卷积和反卷积示意图如图 4 所示。图 4 中蓝色为卷积输入的特征图,绿色为卷积(反卷积)之后的特征图,图 4(a)为卷积过程,卷积核大小为 3×3 ,步长为 1×1 ,无填充层,将图中大小为 4×4 的特征图卷积为 2×2 。图 4(b)为反卷积过程,卷积核大小为 3×3 ,步长为 1,填充 2 层(虚线部分为填充层),将图中大小为 2×2 的特征图反卷积为 4×4 。在步长为 1 时,反卷积输出特征的计算公式为:

$$out = n + 2s - d + 1 \quad , \quad (3)$$

式中: out 为输出特征图大小; n 为输入特征图大

小; s 为填充层数; d 为卷积核个数。

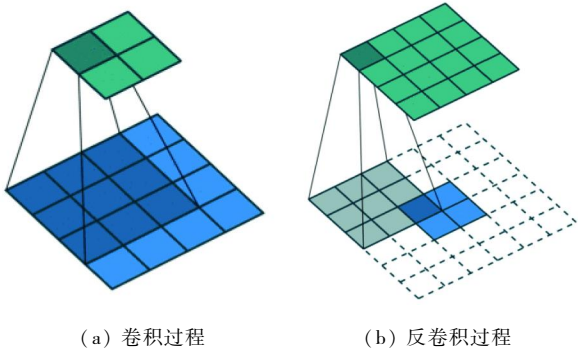


图 4 卷积和反卷积

Fig. 4 Convolution and deconvolution

1.3 多尺度融合

在卷积网络进行特征提取的过程中,随着网络层数的加深,每层获得的特征信息会有所不同,低层网络侧重提取地物颜色、边缘等纹理特征信息,高层网络侧重提取高层抽象特征,同时获得的特征图维

度也会逐渐缩小,由于不同层的特征信息具有不同的特点,只利用最后的高层信息进行地物分类会在一定程度上丢失部分信息。因此通过多尺度特征融合可以将低层网络的特征图和高层网络的特征图进行融合,通过特征融合获取更多来自不同卷积层的空谱特征,从而进一步提高光谱图像的分类精度。本文多尺度融合思想^[10]源于 RefineNet 网络中的多分辨率融合 (multi-resolution fusion, MRF) 模块,经过残差块后将提取到的不同尺度的特征图实现融合,不同尺度的特征融合过程示意如图 5 所示,对 18×18 , 9×9 和 3×3 的特征图分别进行二维卷积 (其中, padding 层为 same、卷积核大小为 3×3), 再进行不同 Scale 的上采样过程 (Scale 为扩大整数倍), 其中 9×9 特征图扩大 2 倍变为 18×18 , 3×3 的特征图扩大 6 倍变为 18×18 , 最后将所有的 18×18 的特征图加和后输入到下一层。

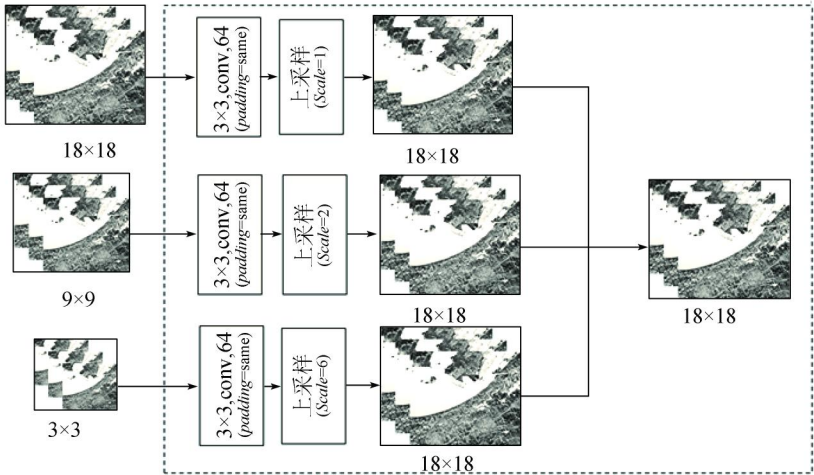


图 5 多尺度融合结构图

Fig. 5 Multi-scale fusion structure diagram

1.4 算法描述

基于以上算法分析, 本文的算法实现过程描述如下:

开始:

输入: 原始高光谱图像

1) 样本标签。

- ①对原始高光谱图像做样本标签, 形成样本库;
- ②按一定比例将样本库分为训练样本和测试样本;

2) 提取空间特征。

③通过 PCA 算法求得原始图像的第一主成分作为输入;

3) 改进残差网络。

a. 训练阶段:

④从训练样本中按照迭代批次 (batch) 随机输入到第一残差层中, 进行卷积核大小为 $l \times l$ 的最大池化;

⑤将步骤④中的输出作为第二残差层的输入, 再进行卷积核大小为 $l \times l$ 的最大池化;

⑥将步骤④, ⑤中的输出作为反卷积的输入, 卷积核大小为 $c_i \times c_i$ (表示第 i 层卷积核大小为 c), 步长为 1;

⑦将两个反卷积之后图像进行多尺度特征融合;

⑧将⑦中的输出作为第三残差层的输入, 然后进行卷积核大小为 $l \times l$ 的最大池化;

⑨将步骤⑧中的输出作为第四残差层的输入, 再进行卷积核大小为 $l \times l$ 的最大池化;

- ⑩将步骤⑨中的输出作为 conv5 残差层的输入,获得更多的特征;
- ⑪将提取到的特征输入到全连接层;
- ⑫迭代模型直至模型收敛;
- ⑬模型训练完毕;
- b. 测试阶段:
- ⑭将测试样本输入到训练好的模型中,计算混淆矩阵并得到分类精度。
- ⑮测试完成。
- 输出: 混淆矩阵,OA,Kappa 系数。
- 结束

2 实验结果与分析

2.1 数据描述

本文采用的数据集为珠海一号卫星高光谱数据,该数据共有 32 个波段,空间分辨率是 10 m,原始图像大小为 5 056 × 5 056。分别在江苏省无锡市太湖地区和安徽省合肥市巢湖地区裁剪 2 个 3 000 × 3 000 的实验区域,其中太湖地区影像成像时间为 2018 年 10 月 1 日,左上角经度为 E120°11'34",纬度为 N31°06'52",文中称为太湖,如图 6(a);巢湖

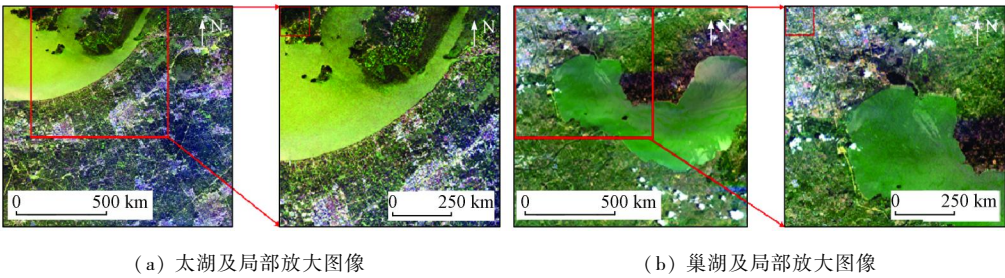


图 6 珠海一号 B14(R),B7(G),B2(B)假彩色合成图像
Fig. 6 B14(R), B7(G), B2(B) false color composite image of Zhuhai - 1

区域影像成像时间为 2019 年 4 月 17 日,左上角经度为 E117°14'41",纬度为 N31°50'25",文中称为巢湖,如图 6(b)。

实验中将数据分为湖水、房屋、农田 3 类,依据光谱曲线和谷歌地图两者结合一起以像素点为单位做样本标签,太湖和巢湖的平均光谱曲线如图 7 所

示,红色表示湖水,绿色表示房屋,蓝色表示农田,从这 2 个图中可以看出数据中 3 种类别地物可以较好地区分。太湖实验数据中带标签的像素共有 29 346 个,巢湖带标签的像素共有 29 512 个,实验中训练样本与测试样本都是按照 1:1 的比例,具体数量见表 1。

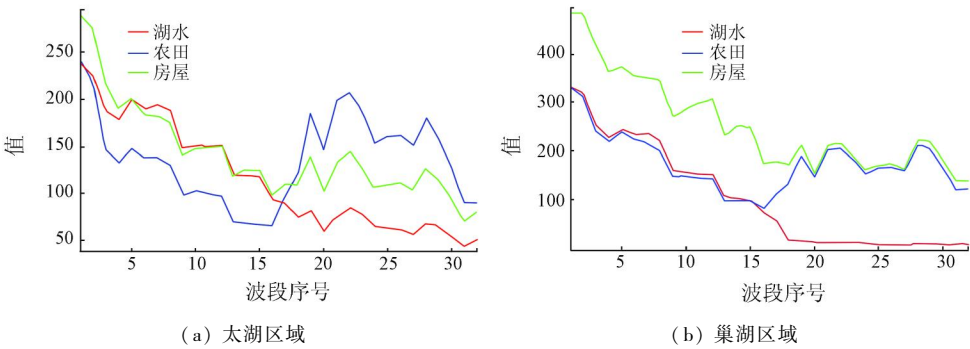


图 7 实验区光谱曲线
Fig. 7 Spectral curves of study areas

表 1 数据集样本个数

Tab. 1 Number of data set samples (个)

类别	太湖		巢湖	
	训练数据	测试数据	训练数据	测试数据
湖水	4 983	5 069	4 767	5 308
房屋	4 809	4 803	4 984	4 451
农田	4 881	4 801	5 001	4 998
总计	14 673	14 673	14 755	14 757

2.2 预处理与实验设置

预处理方式采用 PCA 算法对高光谱图像降维,

将 32 个波段的数据降维为 1 个波段作为输入,其中对于训练样本是随机输入,测试样本为总样本数减训练样本,训练样本的随机性使模型精度不完全一致,因此每个实验都是训练 5 次确保实验结果的稳定性,使实验结果更加具有说服力。

深度学习中包含大量待训练参数,网络结构直接影响训练参数的多少,结构越复杂训练参数越多,训练难度也会增加,固定部分参数是非常有必要的。本文实验中设置学习率为 0.000 5,丢弃比例(drop_

prob) 为 0.5, 批量处理 (batch_size) 大小为 20, 迭代次数为 20 000 次。改进残差网络参数及其网络设置见表 2, 同样的两套数据进行孪生网络^[11], SVM^[12], CNN^[13], GLCMCNN^[14] 和传统残差网络的对比实验。孪生网络 (Siamese) 是将两个输入映射到统一特征空间中, 并共享权值, 判别两个样本是否为正样本或者负样本, 只需一次前向传播操作就可以计算出精度; SVM 采用径向基函数作为核函数, 参数 g 和惩罚因子 c 都是经过 5 次交叉验证寻优得到; CNN 参数设置见表 3; GLCMCNN 中选取了平均值、方差、熵、偏度 4 个值, 再结合 CNN 网络进行分类, 其中 CNN 网络参数与表 3 相同; 传统残差网络参数及其网络设置见表 4。

表 2 改进残差网络
Tab. 2 Improved residual network

层数	输出尺寸大小	卷积核大小及个数	步长	激活层
conv1i	27 × 27	$\begin{bmatrix} 3 \times 3 & 32 \\ 3 \times 3 & 32 \\ 3 \times 3 & 32 \end{bmatrix}$	1 × 1	ReLU
pool1	14 × 14	—	2 × 2	—
conv2i	14 × 14	$\begin{bmatrix} 3 \times 3 & 64 \\ 3 \times 3 & 64 \\ 3 \times 3 & 64 \end{bmatrix}$	1 × 1	ReLU
pool2	7 × 7	—	2 × 2	—
Dconv1	18 × 18	5 × 5 64	1 × 1	ReLU
Dconv2	9 × 9	3 × 3 64	1 × 1	ReLU
M1	18 × 18	32	—	—
conv3i	18 × 18	$\begin{bmatrix} 3 \times 3 & 32 \\ 3 \times 3 & 32 \\ 3 \times 3 & 32 \end{bmatrix}$	1 × 1	ReLU
pool1	9 × 9	—	2 × 2	ReLU
conv4i	9 × 9	$\begin{bmatrix} 3 \times 3 & 64 \\ 3 \times 3 & 64 \\ 3 \times 3 & 64 \end{bmatrix}$	1 × 1	ReLU
pool2	5 × 5	—	2 × 2	ReLU
conv5	4 × 4	3 × 3 128	1 × 1	ReLU
pool5	2 × 2	—	2 × 2	—

表 5 不同方法的对比结果

Tab. 5 Comparison results of different methods

模型	太湖		巢湖	
	OA/%	Kappa × 100	OA/%	Kappa × 100
孪生网络	61.27 ± 0.86	52.12 ± 0.15	64.57 ± 0.59	53.48 ± 0.25
SVM	76.73	69.02	80.40	70.52
CNN	86.23 ± 0.92	79.44 ± 1.47	83.25 ± 1.22	75.08 ± 2.47
GLCMCNN	88.94 ± 0.26	80.23 ± 0.33	85.32 ± 0.46	78.23 ± 0.7
传统残差网络	90.07 ± 1.88	85.41 ± 2.83	89.26 ± 1.18	84.18 ± 1.56
本文方法	92.27 ± 2.45	86.60 ± 4.00	91.31 ± 3.07	87.14 ± 4.64

9.68%, 7.02% 和 2.30%, Kappa 值分别高出了 62.94%, 23.57%, 16.06%, 11.39% 和 3.52%。孪生网络是将两个输入映射到统一特征空间中, 提取相似特征, 由于没能有效提取不同类别的地物特征, 导致其精度偏低; SVM 属于传统机器学习, 提取的特征属于浅层特征, 高光谱图像中的深层特征没有

表 3 CNN 网络结构

Tab. 3 CNN network structure

层数	卷积核大小	卷积核个数	步长	激活函数
conv1	4 × 4	32	1 × 1	ReLU
pool1	2 × 2	—	2 × 2	—
conv2	5 × 5	64	1 × 1	ReLU
pool2	2 × 2	—	2 × 2	—
conv3	4 × 4	128	1 × 1	ReLU

表 4 传统残差网络

Tab. 4 Traditional residual network

层数	卷积核大小	卷积核个数	步长	激活函数
conv11	—	16	1 × 1	ReLU
conv12	—	16	—	ReLU
conv13	—	16	—	ReLU
pool1	2 × 2	—	2 × 2	—
conv21	—	32	—	ReLU
conv22	—	32	—	ReLU
conv23	—	32	—	ReLU
pool2	2 × 2	—	2 × 2	—
conv3	—	64	—	ReLU

2.3 实验结果与分析

表 5 为本文方法与孪生网络, SVM, CNN, GLCMCNN 和传统残差网络在太湖和巢湖数据集进行对比的分类结果, 从总体分类精度和 Kappa 系数看, 本文方法获得了最高的分类精度, 证明该算法分类性能较好, 总体分类精度和 Kappa 系数的值均为“平均值 ± 标准差”格式。表 5 中, 相对传统方法, 本文方法取得了最好的效果, 在太湖数据中, 相比于孪生网络, SVM, CNN, GLCMCNN 和传统残差网络, 本文方法的 OA 分别高出了 50.60%, 20.25%, 7.00%, 3.74% 和 2.44%, Kappa 值分别高出了 66.16%, 25.47%, 9.01%, 7.94% 和 1.39%; 在巢湖数据中, 相比于孪生网络, SVM, CNN, GLCMCNN 和传统残差网络, 本文方法的 OA 分别高出了 41.41%, 13.57%,

得到充分利用, 影响了其分类精度; CNN 模型和 GLCMCNN 模型精度低是因为高光谱图像输入样本尺寸偏小, 导致 CNN 模型的网络层数被限制, 不能充分提取到高光谱图像的深层特征; 传统的残差网络由于加深了网络层次, 可以进一步提取高光谱图像中的深层特征, 获得了较高的精度, 但是由于它只

是将每层提取的特征输入到下一层残差单元,不同层之间的特征没有进行深度融合和挖掘,限制了高光谱图像分类精度的进一步提升;本文方法通过采用残差网络提高网络深度,并利用其捷径连接解决由于加深网络导致的梯度消失问题,提高了图像分类精度;同时利用反卷积扩大特征图,并且通过多尺度融合算法使不同层的特征进行深度融合,不仅

解决了网络层数被限制的问题,还充分挖掘高光谱图像中的深度特征信息进一步提高了分类精度,因此获得了最高的分类精度。

图 8 所示为可视化结果,图中(a)和(c)分别为太湖和巢湖经裁剪后的原始图像,(b)和(d)分别为本文方法在太湖和巢湖数据集进行分类后的可视化结果,从图中可以看出,本文方法取得了较好的分类效果。

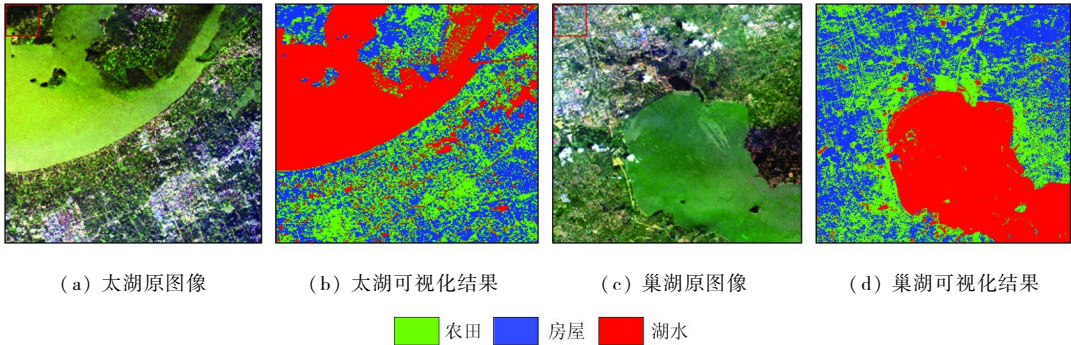


图 8 可视化结果

Fig. 8 Visualization results

2.4 参数对实验影响

2.4.1 卷积核个数对实验的影响

在深度学习中,卷积核个数与最终的分类精度并不是成正比的,该参数的设置尚未有明确的理论支持,本文采用根据经验值设置候选值,并通过实验分析确定最优的卷积核个数。一般而言,残差网络中第一层网络中卷积核个数为 16,本节做了第一层卷积核个数分别设置为 8,16,32 和 40 分别进行 5 次实验,结果如表 6 和表 7 所示。

表 6 不同卷积核个数下太湖数据的分类结果				
Tab. 6 Classification results of Taihu Lake data under different numbers of convolution kernels				
结果	8	16	32	40
OA/%	86.20 ± 1.2	89.11 ± 0.71	92.27 ± 2.45	88.24 ± 2.89
Kappa × 100	79.58 ± 1.81	83.96 ± 1.18	86.60 ± 4.00	82.65 ± 4.26

表 7 不同卷积核个数下巢湖数据的分类结果				
Tab. 7 Classification results of Chaohu Lake data with different numbers of convolution kernels				
结果	8	16	32	40
OA/%	87.00 ± 1.09	90.20 ± 1.54	91.31 ± 3.07	89.48 ± 4.18
Kappa × 100	80.78 ± 1.49	85.49 ± 0.74	87.14 ± 4.64	84.45 ± 7.38

可以看出:

1)卷积核个数为 16 时比卷积核个数为 8 时的精度高,验证了候选值的可靠性。

2)卷积核个数为 32 时精度最高,太湖和巢湖数据集卷积核为 32 时比卷积核个数为 16 时的总体分类精度上分别提高了 3.16% 和 1.11%,在 Kappa 系数上也分别提高了 2.64% 和 1.65%。

3) 在卷积核个数为 40 时精度比卷积核个数为 32 时在太湖和巢湖数据上分别降低了 4.03% 和 1.83%。综合考虑本文卷积核个数采用 32。

2.4.2 输入大小对实验的影响

本文中输入样本是以某一像素点为中心、以 $N \times N$ (N 为奇数)的邻域块作为二维卷积神经网络的输入,图像分类中邻域块的分类结果作为中心像素点的类别。样本输入大小不同会影响高光谱图像分类精度,本节主要讨论 15×15 , 25×25 , 27×27 和 29×29 这 4 个不同输入对太湖和巢湖的影响,表 8 和表 9 分别为太湖和巢湖在不同输入的分类结果,从表 8 中可以看出在输入为 27×27 时精度最高,而表 9 中是输入为 29×29 时精度最高,比 27×27 时提高了 0.9%,但是随着输入尺寸增加计算成本也会随之增加,综合考虑分类精度和计算代价,本文算法选取了输入大小为 27×27 。

表 8 不同输入大小下太湖的分类结果				
Tab. 8 Classification results of Taihu Lake under different input sizes				
结果	15 × 15	25 × 25	27 × 27	29 × 29
OA/%	77.88 ± 2.37	90.04 ± 1.08	92.27 ± 2.45	91.03 ± 4.31
Kappa × 100	67.08 ± 4.31	85.27 ± 1.79	86.60 ± 4.00	87.37 ± 5.80

表 9 不同输入大小下巢湖的分类结果				
Tab. 9 Classification results of Chaohu Lake with different input sizes				
结果	15 × 15	25 × 25	27 × 27	29 × 29
OA/%	78.90 ± 1.64	89.12 ± 2.20	91.31 ± 3.07	92.20 ± 1.70
Kappa × 100	68.62 ± 1.78	83.94 ± 3.26	87.14 ± 4.64	88.99 ± 1.87

3 结 论

为进一步充分挖掘高光谱图像中的深层特征,本文提出基于残差网络特征融合的高光谱图像分类方法,利用残差网络中残差单元进行特征提取,通过其恒等映射的特点加深网络层数,利用反卷积扩充特征图,并通过融合不同尺度的特征实现高光谱图像的深度特征提取和融合。实验结果表明,与其他学习方法相比,本文方法获得了最好的分类性能。具体总结如下:

1) 高光谱图像蕴含丰富的地物图谱信息,传统机器学习只能提取浅层特征,不能充分利用高光谱图像中的深层特征影响了其分类精度的提高;而深度学习技术由于其良好的深度特征提取能力获得了较好的分类效果。

2) 传统残差网络可以利用其恒等映射特点加深网络层数,解决了高光谱图像分类中由于样本输入尺寸偏小导致的网络层数有限的问题,同时缓解了由于网络层数加深容易出现的梯度消失问题,可以进一步深度提取地物特征,提高了高光谱图像的分类精度。

3) 本文方法充分利用了残差网络在加深网络层数、缓解梯度消失方面的优势,并结合反卷积通过填充信息扩大特征图、进行不同维度的特征图多尺度有效融合,实现层与层之间的深层特征互补,深度挖掘并充分利用高光谱图像中的深层特征信息,进一步提高了高光谱图像分类精度。

本文方法主要是针对高光谱图像进行分类,但是目前高光谱传感器由于时空分辨率的限制,有些区域的高光谱图像难以获取,为满足更广泛的地物分类识别,充分利用高空间分辨率、高光谱分辨率及高时间分辨率的遥感数据进行多源数据融合是我们接下来的研究方向。

参考文献 (References):

[1] 李秉璇,周 冰,贺 宣,等. 针对高光谱图像的目标分类方法现状与展望[J]. 激光与红外,2020,50(3):259-265.
Li B X,Zhou B,He X,et al. Status and prospects of target classification methods for hyperspectral images[J]. Laser and Infrared, 2020,50(3):259-265.

[2] 张 亮. 基于 PCA 和 SVM 的高光谱遥感图像分类研究[J]. 光学技术,2008,34(s1):184-187.
Zhang L. Research on classification of hyperspectral remote sensing images based on PCA and SVM[J]. Optical Technology,2008,34(s1):184-187.

[3] 芦国军,陈丽芳. 基于深度卷积神经网络的遥感图像场景分类[J]. 太原师范学院学报(自然科学版),2019,18(1):57-62.
Lu G J,Chen L F. Remote sensing image scene classification based on deep convolutional neural network[J]. Journal of Taiyuan Teachers College (Natural Science Edition),2019,18(1):57-62.

[4] 王云艳,罗冷坤,周志刚. 基于反卷积高层特征的遥感地物图像分类[J]. 计算机工程与应用,2020,56(11):200-206.
Wang Y Y,Luo L K,Zhou Z G. Remote sensing object image classification based on deconvolution high-level features[J]. Computer Engineering and Applications,2020,56(11):200-206.

[5] 于佩鑫,周 询,刘素红,等. 东北黑土区侵蚀沟遥感影像特征提取与识别[J]. 遥感学报,2018,22(4):611-620.
Yu P X,Zhou X,Liu S H,et al. Feature extraction and recognition of remote sensing image of eroded gully in black soil region of northeast China[J]. Journal of Remote Sensing,2018,22(4):611-620.

[6] He K M,Zhang X Y,Ren S Q,et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2016,6:770-778.

[7] 孟佳佳,王 毅. 基于深度残差网络的遥感数据分类[J]. 数字技术与应用,2019,37(1):99-101.
Meng J J,Wang T. Remote sensing data classification based on deep residual network[J]. Digital Technology and Application, 2019,37(1):99-101.

[8] 张怡卓,徐苗苗,王小虎,等. 残差网络分层融合的高光谱地物分类[J]. 光谱学与光谱分析,2019,39(11):3501-3507.
Zhang Y Z,Xu M M,Wang X H,et al. Classification of hyperspectral features by hierarchical fusion of residual networks[J]. Spectroscopy and Spectral Analysis,2019,39(11):3501-3507.

[9] 周俊宇,赵艳明. 卷积神经网络在图像分类和目标检测应用综述[J]. 计算机工程与应用,2017,53(13):34-41.
Zhou J Y,Zhao Y M. Overview of the application of convolutional neural networks in image classification and target detection[J]. Computer Engineering and Applications,2017,53(13):34-41.

[10] 裴君岩,刘义海. 基于多尺度感受野扩增融合的遥感目标检测算法[J]. 指挥控制与仿真,2020,42(1):34-39
Pei J Y,Liu Y H. Remote sensing target detection algorithm based on multi-scale receptive field amplification fusion[J]. Command Control and Simulation,2020,42(1):34-39.

[11] Nair V,Hinton G E. Rectified linear units improve restricted boltzmann machines[Z]. In ICML. 2010.

[12] Chen P H,Lin C J,Schölkopf,B. A tutorial on v-support vector machines[J]. Applied Stochastic Models in Business and Industry,2005,21(2):111-136.

[13] Krizhevsky A,Sutskever I,Hinton G. Imagenet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems,2012,25(2):1-9.

[14] Su H,Wang Y,Xiao J,et al. Improving MODIS sea ice detectability using gray level co-occurrence matrix texture analysis method: A case study in the Bohai Sea[J]. ISPRS Journal of Photogrammetry & Remote Sensing,2013,85(85):13-20.

Classification of hyperspectral image based on feature fusion of residual network

HAN Yanling, CUI Pengxia, YANG Shuhu, LIU Yekun, WANG Jing, ZHANG Yun
(School of Information, Shanghai Ocean University, Shanghai 201306, China)

Abstract: Deep learning technology provides technical means for hyperspectral image classification due to its unique advantages in deep mining of features. However, in the pixel – level feature classification of hyperspectral images, the number of deep learning layers is limited due to the influence of the sample input size, and the depth features in the hyperspectral images cannot be fully mined. The classification of hyperspectral image based on feature fusion of residual network is proposed in this paper. First, the principal component analysis (PCA) method is used to extract the first principal component in the original hyperspectral image, and the residual network is used to effectively extract the spatial spectrum features of the ground objects; then the feature map is expanded by the deconvolution algorithm, and after deconvolution, features of different dimensions are fused with multi – scale features to fully mine the depth feature information in the hyperspectral image, thus further improving the classification accuracy of the hyperspectral image. The ground feature classification experiment was conducted on the two areas of Taihu Lake in Jiangsu and Chaohu Lake in Anhui captured by the “Zhuhai – 1” satellite. The results show that, compared with other methods, this method can effectively solve the problem of insufficient depth feature extraction in hyperspectral image classification, thus showing better classification performance.

Keywords: deconvolution; feature fusion; residual network; hyperspectral image classification

(责任编辑: 张 仙)