

基于改进 3D-CNN 的多源遥感数据树种识别

栗旭升¹, 陈冬花^{2,3**}, 刘赛赛³, 张乃明⁴, 李虎^{2*}

¹新疆农业大学草业与环境科学学院, 新疆 乌鲁木齐 830052;

²安徽师范大学地理与旅游学院, 安徽 芜湖 241000;

³滁州学院计算机与信息工程学院, 安徽 滁州 239000;

⁴新疆师范大学地理科学与旅游学院, 新疆 乌鲁木齐 830001

摘要 针对森林复杂冠层结构和林分高密度下遥感树种识别精度不高的问题, 将能够提取高维数据立体特征的三维卷积神经网络(3D-CNN)引入到遥感影像树种识别中, 并利用残差网络(ResNet)对其进行改进, 提出三维残差卷积神经网络(3D-RCNN), 以减小网络深度带来的误差, 降低退化现象的影响。联合高分五号高光谱数据(GF-5 AHIS)和高分六号高空间分辨率数据(GF-6 PMS), 辅以森林资源数据和外业调查数据构建样本集。结合 3D-RCNN 思想构建树种识别模型。实验结果表明: 相较于传统 3D-CNN, 3D-RCNN 将模型网络从 12 层增加到 18 层, 能够深化网络结构, 缓解网络退化; 联合 GF-5 AHIS 和 GF-6 PMS, 3D-RCNN 能够有效地识别北亚热带森林树种, 且识别精度(91.72%)要优于传统 3D-CNN(85.65%)和支持向量机算法(85.22%)。

关键词 图像处理; 卷积神经网络; 残差网络; 树种识别; 高光谱影像

中图分类号 TP751.1

文献标志码 A

doi: 10.3788/LOP57.242804

Tree-Species Identification of Multisource Remote-Sensing Data using Improved 3D-CNN

Li Xusheng¹, Chen Donghua^{2,3**}, Liu Saisai³, Zhang Naiming⁴, Li Hu^{2*}

¹College of Grassland and Environment Sciences, Xinjiang Agricultural University, Urumqi, Xinjiang 830052, China;

²School of Geography and Tourism, Anhui Normal University, Wuhu, Anhui 241000, China;

³College of Computer and Information Engineering, Chuzhou University, Chuzhou, Anhui 239000, China;

⁴College of Geography and Tourism, Xinjiang Normal University, Urumqi, Xinjiang 830001, China

Abstract Aiming to address the low identification accuracy of remote-sensing tree species of forests with a complex canopy and high density, a three-dimensional convolution neural network (3D-CNN) that can extract the stereoscopic features of hyper-dimensional data is introduced herein, and it can identify remote-sensing images. Furthermore, it is improved through residual network (ResNet) to build a 3D residual convolution neural network (3D-RCNN) to reduce the influence of degradation phenomenon and the inaccuracy caused by network depth. The sample set is constructed by combining GF-5 hyperspectral data (GF-5 AHIS) and GF-6 high spatial resolution data (GF-6 PMS), supplemented by forest resource data and field survey data. Then, a tree species recognition model is constructed based on the concept of 3D-RCNN. The experimental results show that compared with traditional 3D-CNN, the proposed 3D-RCNN increases the model network's density from 12 layers to 18 layers, which can deepen the network structure and alleviate network degradation. By combining GF-5 AHIS and GF-6 PMS, 3D-RCNN can effectively identify northern subtropical forest species, providing better recognition accuracy (91.72%) than traditional 3D-CNN (85.65%) and support vector machine algorithm (85.22%).

Key words image processing; convolutional neural network; residual network; tree species identification;

收稿日期: 2020-04-23; 修回日期: 2020-05-15; 录用日期: 2020-05-29

基金项目: 安徽省高校协同创新项目(GXXT-2019-047)、安徽省高校学科优秀拔尖人才学术培育项目(gxbjZD44)、高分专项省(自治区)域产业化应用项目(76-Y40G05-9001-15/18)

* E-mail: lihu2881@aliyun.com; ** E-mail: chendonghua1981@sina.com

hyperspectral image

OCIS codes 100.2000; 100.4996; 120.0280

1 引言

树种空间分布信息的获取是林业科学生产经营的前提,可为林业部门的森林资源管理提供科学的决策依据,对森林经管有着重要的实际意义。遥感技术逐渐成为森林树种识别的主要技术手段,但常规光学数据的光谱分辨率低,“同物异谱”和“同谱异物”现象严重,其树种识别精度往往达不到应用要求,而高光谱数据能够很好地解决这个问题^[1-2]。以往,星载高光谱数据的获取途径较少,相关研究基础较为薄弱,GF-5号的成功发射改变了这一困境。关于GF-5 AHIS和GF-6 PMS联合的树种识别研究鲜有报道,开展国产多源遥感数据树种识别研究不仅可以推动国产高分数据的应用,同时也为进一步开展森林动态监测、森林生态多样性保护等研究奠定基础。

随着分类算法研究深入,模型驱动识别算法因参数精度的限制,识别精度在到达反演参数精度上限后,难以再有突破^[3]。基于深度学习的数据驱动型识别算法可以通过大量数据自行挖掘隐含特征并进行分类反演,成为遥感图像识别研究的新方向^[4]。卷积神经网络(CNN)算法经权值共享和局部连接能够很好地挖掘高光谱数据丰富的空间信息和光谱信息,在高光谱影像识别领域大放异彩。2015年,Hu等^[5]首次将CNN应用于高光谱数据分类,但仅利用光谱信息进行分类;2016年,Chen等^[6]提出一种能够同时提取空间信息和光谱信息的3D-CNN模型,并利用随机丢弃(dropout)和正则化(regularization)等优化策略提高模型的泛化能力,有效地提高了高光谱分类精度;2019年,闫苗等^[7]构建3类样本集对不同CNN模型高光谱数据分类的可行性进行探究,证明CNN模型高光谱数据识别具有较高的准确率和可行性;2019年,刘玉珍等^[8]依据高光谱图像中像素之间的光谱关系和空间关系构建超图,通过超图构建具有空谱联合信息的样本,再结合CNN算法进行分类识别,结果证明空谱联合信息能够更好地表征地物特征,提高分类精度。大量研究说明,高光谱数据中空间和光谱信息的联合能够有效提高影像分类精度。

综上所述,开展联合GF-5 AHIS和GF-6 PMS数据的树种识别研究具有现实应用意义,CNN对高维数据的识别分类具有可行性。3D-CNN通过增加

光谱维度卷积计算能够对高光谱数据立体特征进行提取,从而提高识别精度,但维度的增加不可避免地增大了算法的时间复杂度,引起退化现象。为降低参数计算的复杂度,减小网络深度造成的误差,缓解退化现象,本文借鉴残差网络思想,在3D-CNN中引入“短路”结构,构建3D-RCNN。结合GF-5 AHIS、GF-6 PMS数据制作的样本集,构建并训练3D-RCNN树种识别模型。

2 基本原理

2.1 三维卷积神经网络

3D-CNN最初应用于视频分类和动作识别等领域,和传统二维卷积神经网络(2D-CNN)相比,3D-CNN除了能在二维进行卷积特征提取,还能够提取第三维度的特征信息。对应到视频分类和动作识别领域,就是在空间维度上增加时间维度,逐帧提取动态特征;而在高光谱识别领域,就是在空间维度的基础上增加光谱维度,同时提取光谱影像的空间和光谱特征^[9]。2D-CNN和3D-CNN公式分别为

$$v_{ij}^{xy} = g \left[b_{ij} + \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \omega_{ijm}^{pq} v_{(i-1)m}^{(x+p)(y+q)} + b_{ij} \right], \quad (1)$$

$$v_{ij}^{xyz} = g \left[b_{ij} + \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} \omega_{ijm}^{pqr} v_{(i-1)m}^{(x+p)(y+q)(z+r)} + b_{ij} \right], \quad (2)$$

式中: i 为神经网络层序数; j 为特征样本序数; P_i 和 Q_i 分别为二维空间上卷积核的长和宽; R_i 为卷积核在第三维度上的高; m 为与上一次网络连接的特征数; $g[\cdot]$ 为神经元的激活函数,这里为ReLU函数; v_{ij}^{xyz} 为神经元在 (x,y,z) 处的值; ω_{ijm}^{pqr} 为第 m 个特征的第 (p,q,r) 个神经元的传递权重; b_{ij} 为第 i 层第 j 个特征偏移量。

将两种CNN算法对应到高光谱影像识别中,如图1所示,2D-CNN只是一个提取空间特征的二维滤波器,同一时刻只能以 P 和 Q 为卷积窗口提取一个波段的空谱联合信息。3D-CNN则是将卷积核从二维增至三维,能够同时以 P 和 Q 为卷积窗口对 R 个波段进行卷积特征提取(图1中3D-CNN的 R 就是3,卷积核为 $P \times Q \times 3$),可以很好地解决2D-CNN无法解决的空谱结合的问题。

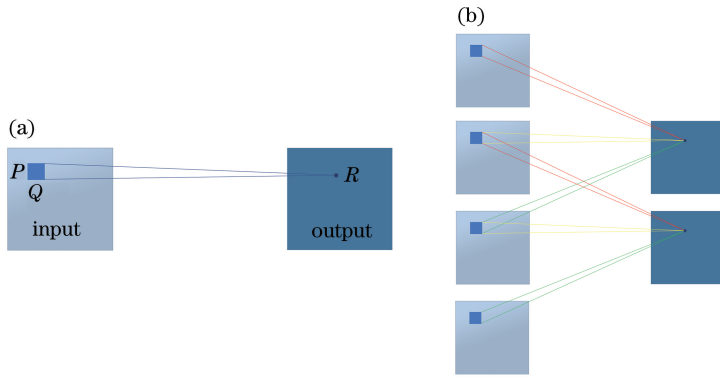


图 1 3D-CNN 和 2D-CNN 原理对比。(a) 2D-CNN; (b) 3D-CNN

Fig. 1 Principle comparison between 3D-CNN and 2D-CNN. (a) 2D-CNN; (b) 3D-CNN

2.2 残差网络

2016 年, He 等^[10] 借鉴 Highway networks 思想, 提出了残差神经网络 (ResNet), 该网络在第六届 ImageNet 年度图像识别大赛的多个类别测试中获得第一名。ResNet 的设计思路是通过在网络中增加恒等映射层将模型转换为一个浅层网络, 这在保证足够参数拟合的同时降低了网络深度, 以规避退化现象^[10], 在实际操作中就是部分数据不经过网络变换直接输入到下一层。ResNet 和常规网络模型结构如图 2 所示, 常规网络模型的数据输入均是通过上一层的输出计算得到的, 称之为平网络; 而残

差网络会有一个“短路 (shortcut)”的结构, 将经过多层网络处理后的数据和通过“短路”结构跳过多层网络处理的数据共同作为输入, 并将其传入到下一层网络内。在一个结构单元内, 要学习的期望映射均为 $H(x)$, 即图 2 中目标函数 $H(x)$ 和 $F(x)+x$ 是等价的。推导出 $F(x) = H(x) - x$, 称之为残差。这样 ResNet 的学习目标就转换为求解 $H(x)$ 和 x 间的差值, 从而消去映射前后相同的部分, 突出了微小的变化, 使深层网络结构在保证结构性能的同时不至于发生退化现象。

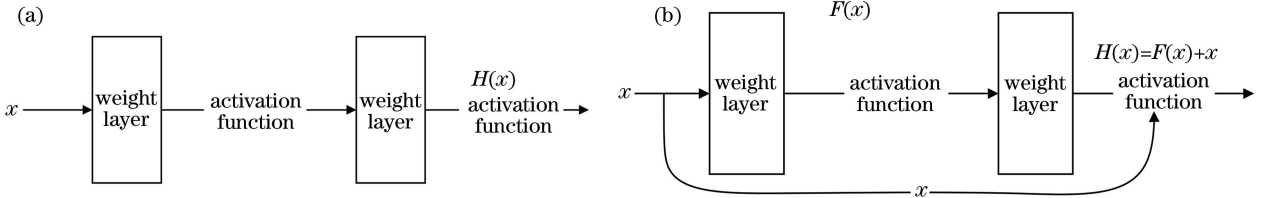


图 2 网络结构图。(a) 平网络; (b) 残差网络

Fig. 2 Diagram of network structure. (a) Flat network; (b) ResNet

2.3 三维残差卷积神经网络

CNN 算法本质上是通过不断修正多层次非线性参数来拟合真实属性的过程, 理论上, 网络越深, 层数越多, 提取特征越丰富, 拟合精度越高。但随着网络深度的增加, 可能会出现梯度消失和梯度爆炸现象^[11]。为了解决这一问题, dropout 和批归一化 (BN) 技术被用于网络结构的改进, 通过这些技术, 可以对数十层的网络结构进行训练。但随着网络深度的增加, 出现了训练精度下降的现象, 这就是退化问题, 这种退化不是过拟合, 而是额外网络层数带来的训练误差。残差神经网络的出现就是为了解决此类退化问题。

征提取, 但卷积维度的增加不可避免地使整个神经网络的参数量大幅增加, 致使整个算法的时间复杂度激增。如著名的 C3D 网络只有 11 层, 模型运行所需 GPU 内存却达到 321 MB, 甚至大于 152 层的 ResNet 模型, 同时网络深度的增加也会导致退化现象。ResNet 能够通过 shortcut 的独特设计, 降低参数计算的复杂度, 减少网络深度造成的误差, 有效地缓解退化现象。根据 3D-CNN 和 ResNet 的性能特点, 将二者有机结合, 提出 3D 残差神经网络 (3D-RCNN) 模型。该模型的基本结构如图 3 所示, 为降低网络层数, 减少非线性转换运算, 同时提高网络特征提取能力, 设计双 3D 卷积层, 在两卷积层之间不进行池化计算。在每次卷积后进行批正则

3D-CNN 虽然能够有效地对高光谱数据进行特

化计算,激活函数使用 ReLU 函数,降低了梯度消失或梯度爆炸的风险。在双卷积结构上添加 ResNet 结构,以降低参数计算的复杂度和模型退

化的风险。在双卷积结构后设置 $2 \times 2 \times 2$ 的 3D 最大池化层,以减少参数量和卷积层参数误差造成的估计均值偏移。

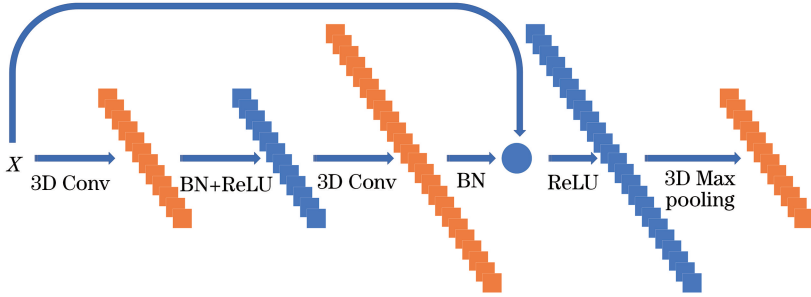


图 3 3D-RCNN 结构

Fig. 3 Structure of 3D-RCNN

3 实验与分析

3.1 研究区与数据源

皇甫山国家森林公园 ($32^{\circ}17' \sim 32^{\circ}22' N, 117^{\circ}58' \sim 118^{\circ}01' E$) 南北长约 10.5 km, 东西宽约 8 km, 总面积为 35.864 km², 是江淮地区保存最完整、面积最大的原始次生森林景观带, 国家级北亚热带阔叶林及鸟类保护区。从优势树种看, 阔叶林以朴树 (*Celtis sinensis*)、麻栎 (*Quercus acutissima*) 及黄檀 (*Dalbergia hupeana*) 为主, 还有榉树 (*Zelkova serrata Makino*)、榆树 (*Ulmus pumila*)、乌桕 (*Sapium sebiferum*) 等。针叶林以马尾松 (*Pinus massoniana*)、湿地松 (*Pinus elliotii*) 及杉木 (*Cunninghamia lanceolata*) 为主, 还有黑松 (*Pinus thunbergii*)、云杉 (*Picea asperata*) 等^[12]。研究区与样本分布如图 4 所示。

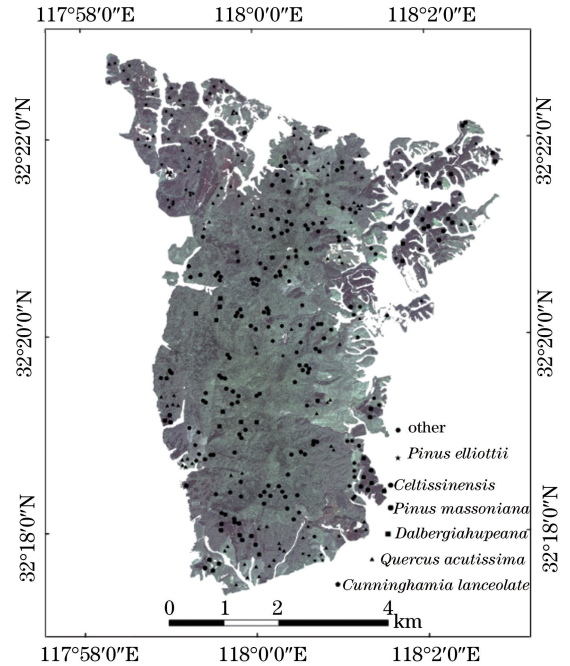


图 4 研究区及样本分布

Fig. 4 Distribution of study area and samples

主要数据为 GF-5 AHIS 数据, 辅助数据有 GF-6 PMS、森林资源调查数据及实地采集数据。GF-5 AHIS 数据, 幅宽 60 km, 空间分辨率为 30 m, 在可见光至短波红外 (400 ~ 2500 nm) 光谱范围里, 有 330 个波段, 其中可见光和近红外短波有 150 个波段, 近红外长波有 130 个波段。可见光及近红外短

波波段光谱分辨率为 5 nm, 近红外长波波段光谱分辨率为 10 nm^[13]。GF-6 PMS 数据的全色空间分辨率和多光谱空间分辨率分别是 2 m 和 8 m, 幅宽 90 km。数据情况如表 1 所示。

表 1 遥感影像数据基本信息

Table 1 Basic information of remote-sensing image data

Sensor	Product number	Latitude and longitude of center	Imaging time	Cloud cover / %
GF-5 AHIS	45157	118.02°E, 32.30°N	2019-05-22	<5
GF-6 PMS	1119873930	117.90°E, 32.10°N	2019-05-01	

森林资源调查数据为皇甫山林场管护人员在 2016 年第 8 次清查数据的基础上, 根据每年的森林更新清查情况制作的 2019 年森林自查数据。结合森林资源调查数据, 进行树种分布情况实地调查, 重

点对朴树、麻栎、黄檀、马尾松、湿地松及杉木 6 类树种样本进行测定。选取面积大于 900 m × 900 m 的纯林进行定位和调查记录, 形成实地采集样本数据。其中朴树样本 72 个, 麻栎样本 64 个、黄檀样本

66个、马尾松样本70个、湿地松样本77个、杉木样本78个、其他树种样本52个。

3.2 样本集构建

对GF-5 AHIS和GF-6 PMS数据进行预处理和特征提取。基于GF-5 AHIS数据,将得到的304个波段作为光谱特征,将26种植被指数作为植被指数特征。基于GF-6 PMS数据,利用灰度共生矩阵提取4个波段、4个方向上的8种纹理特征,共得到128个纹理特征。如图5所示,通过对特征因子进行相关矩阵分析得出表征各类特征因子间相关性

大小的相关系数,光谱特征、纹理特征及植被指数特征中相关系数绝对值大于0.5的强相关波段分别占84.1%、42.7%及32.0%,特征因子内相关性较高,存在冗余。为降低特征集的维度和相关性,增加树种类间可分性,利用Jeffreys-Matusita(JM)距离对特征因子进行排序,然后按类间可分性大小将特征因子逐一投入到线性判别分析模型内,对特征因子进行选择。最终选择出26个光谱特征、12个纹理特征及10个植被指数特征作为特征集进行树种识别。

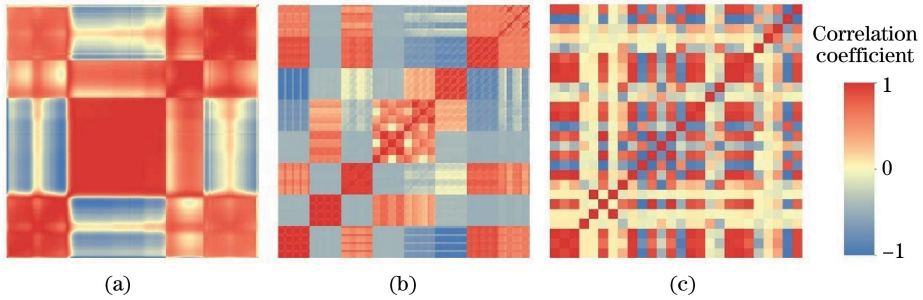


图5 特征因子相关矩阵。(a)光谱特征;(b)纹理特征;(c)植被指数特征

Fig. 5 Correlation matrix of characteristic factors. (a) Spectral feature; (b) texture feature; (c) vegetation index feature

3D-RCNN随着网络层数的增加,参数量也会急剧增加。在训练时如样本不足、反馈网络对参数修正不充分时,会引起过拟合现象。为提高分类精度,对样本进行三级扩充。首先利用已有的朴树、麻栎、黄檀、马尾松、湿地松、杉木及其他树种实测样本提取特征曲线并与未确定树种的区域进行相似性匹配。采用光谱特征拟合、二进制编码及光谱角分类3种方式进行匹配精度评价,将三种方法评分均大于0.8的对象拓展为样点。将6类树种的样本都拓展至100个,其他树种样本扩展至60个,这些样本称为扩展样本。然后利用大小为 100×100 的窗口,以100为步长对特征集进行规则化裁切,结合扩展样本和森林资源调查数据对裁切的规则瓦片进行逐

像元统计,如规则样本内单一树木占比超过80%,则认为该规则瓦片合格,将其标记为该树种的规则样本。最后将样本旋转 90° 、 180° 及 270° ,内旋转 45° 、 135° 、 225° 及 315° ,如图6所示,结合样本镜像,将规则样本扩充16倍,构建出最终的样本集,详细结果如表2所示。按比例将样本集随机分为训练集(60%)、测试集(20%)及验证集(20%)三部分。训练集用于分类模型的训练学习;测试集用来测试训练过程中临时模型的训练精度,使用总体精度(OA)和Kappa系数进行模型精度评价和稳定性评价,从而获取最优的网络结构和参数设置;验证集用以评估最终网络模型的性能,评价系统为混淆矩阵。

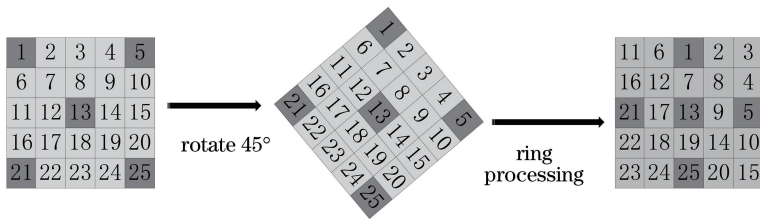


图6 内环旋转扩充样本

Fig. 6 Sample expansion by inner ring rotation

表2 样本数量

Table 2 Number of samples

Tree species	Measured samples	Expanded samples	Regular samples	Sample set
<i>Quercus acutissima</i>	64	100	231	3696
<i>Celtis sinensis</i>	72	100	197	3152
<i>Dalbergia hupeana</i>	66	100	147	2352
<i>Pinus massoniana</i>	70	100	292	4672
<i>Pinus elliotii</i>	77	100	206	3296
<i>Cunninghamia lanceolata</i>	78	100	175	2800
Others	52	60	155	2480

3.3 模型参数设置

参考3D双卷积神经网络模型,将3D-RCNN树种识别模型初始化为1个输入像元块大小为 $P \times P \times L$ 的输入层、 N 个3D残差卷积单元(包含2个3D卷积层、1个最大池化层及1个残差网络结构)、1个全连接层及1个Softmax分类器的组合^[14]。其中,卷积核大小为 $Q \times Q \times Q$,标准步长为 M ,卷积核数量随3D残差单元数量变化依次为8、16、32、 \dots 、 $8 \times N$,卷积核使用 L_2 正则,最大池化的尺寸为 $2 \times 2 \times 2$ 。在全连接层后使用BN和dropout策略,其中dropout系数设置为0.4。使用反向传播算法,batch-size设置为128,利用RMSprop优化器对损失函数进行优化,学习率为 S 。

设置初始参数 $P=5$ 、 $L=48$ 、 $M=1$ 、 $N=1$ 、 $Q=3$ 、 $S=0.0001$,按照控制变量的思想对影响模型精度的卷积步长 M 、3D-RCNN结构数量 N 、卷积核大小 Q 、输入像元大小 P 和学习率 S 5个因素进行实验分析,逐步寻找最优参数,构建最优识别模型。

1) 输入像元大小

在进行卷积运算时,如果输入像元 P 过大,会导致局部特征在全局上被平均,占用内存上升,运算速度降低;如果输入像元 P 过小,会导致特征提取不充分,在局部的效果差。将输入像元大小分别设置为13、15、17、19、21、23,在其他参数默认的情况下进行训练和测试,结果如表3所示。可以看出:随着输入像元尺寸变大,测试和训练所需的时间也随

表3 输入像元大小对运算时间和总体精度的影响

Table 3 Influence of input pixel size on operation time and overall accuracy

P	Training time /s	Test time /s	OA /%
3	879.22	9.12	83.22
15	1020.55	11.22	83.95
17	1159.53	14.71	84.87
19	1362.99	33.53	84.90
21	1505.88	52.12	84.73
23	1817.58	88.60	84.85

之变长,同时测试的总体精度增加速度也越来越慢,甚至出现下降的现象;当 P 超过17后,测试时间迅速增加,且精度基本稳定在85%左右。所以模型输入像元大小 P 设定为17。

2) 残差卷积单元数

3D残差卷积单元的数量 N 控制着整个网络隐藏层的数量和卷积层卷积核的数量。 N 过小,会导致模型参数量不足,最终拟合函数无法准确描述树种间的差异; N 过大,会导致提取参数过多,样本量无法达到训练要求,出现过拟合现象,且模型训练和测试耗时过长。模型输入像元大小 P 设置为17,其他参数保持默认设置,将 N 依次设置为1、2、3、4、5、6、7,进行模型训练和测试,结果如图7所示。可以看出:随着卷积结构 N 的增加,测试时间呈近乎指数式的增加,而总体精度则在 $N=5$ 时达到最大值,90.27%,随后因过拟合,总体精度下降。所以将3D残差卷积单元数 N 设置为5。

3) 卷积核大小

在神经网络模型中,卷积核的大小对模型参数量和局部特征提取精度有着重要影响。基于上文研究进行参数设置,将卷积核大小 Q 依次设置为 3×3 、 5×5 、 7×7 、 9×9 ,进行模型训练和测试,结果如表4所示。随着卷积核的变大,测试时间在逐渐增加,总体精度在卷积核增加到 5×5 时达到最大值90.79%,随后总体精度出现下降。卷积核尺寸 Q 由3增加到5时,总体精度增加0.52个百分点,测试时间增加3.90s,这对模型整体优化而言可以接受,所提模型卷积核尺寸设置为 5×5 。

4) 步长大小

步长是卷积核两次相邻卷积计算间相隔的距离。选取较小的步长能够充分提取特征,但会极大地增加参数量和计算量;选取较大的步长可以对数据进行稀疏降维,减少模型参数量和计算量,但是会在一定程度上造成信息的丢失。基于上文

研究进行参数设置,将步长 M 依次设置为 1, 2, 3, 4, 5, 并对模型进行验证分析,结果如图 8 所示。可以看出,步长设置为 2 时,虽然总体精度稍有下

降,但是模型参数量迅速减少,训练时间显著降低,这在实际生产应用中具有很大的意义。故模型的步长设置为 2。

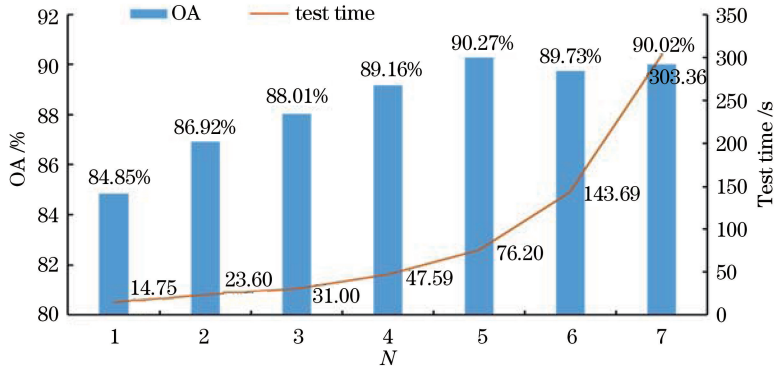


图 7 卷积单元数对测试时间和总体精度的影响

Fig. 7 Influence of number of convolution units on test time and overall accuracy

表 4 卷积核尺寸对运算时间和总体精度的影响

Table 4 Influence of convolution kernel size on operation time and overall accuracy

Q	Test time / s	OA / %
$3 \times 3 \times 3$	76.20	90.27
$5 \times 5 \times 5$	80.10	90.79
$7 \times 7 \times 7$	88.01	90.28
$9 \times 9 \times 9$	105.61	90.52

5) 学习率

学习率用以限制模型权重参数的修正幅度,能

够控制模型梯度下降的速率。在基于高光谱数据的 CNN 模型中,学习率基本上控制在 0.0001~0.001。设置输入像元大小 $P=17$ 、3D 残差卷积单元数 $N=5$ 、卷积核 $Q=5$ 、步长 $M=2$,其他参数保持默认,学习率取 0.0001~0.0010,进行模型训练和测试,结果如表 5 所示。可知:当学习率由 0.0001 增加到 0.0004 时,训练迭代次数显著减少,收敛速度和总体精度都有提高;当学习率超过 0.0004 以后,收敛速度开始降低且出现震荡现象,总体精度也开始出现降低,故将学习率 S 设置为 0.0004。

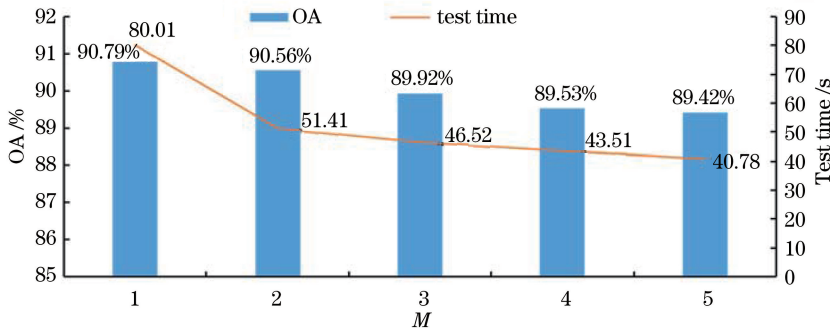


图 8 步长对测试时间和总体精度的影响

Fig. 8 Influence of step size on test time and overall accuracy

最终构建的模型参数如下:输入层 $P=17 \times 17 \times 48$ 、3D 残差单元数量 $N=5$ 、步长 $M=2$ 、卷积核大小 $Q=5 \times 5 \times 5$ 、学习率 $S=0.0004$,其他参数及超参数设置保持默认。

3.4 实验结果及精度评价

利用混淆矩阵,对所提方法(3D-RCNN)、传统机器学习算法支持向量机(SVM)、传统三维卷积神经网络算法(3D-CNN)进行精度评价和对比。使用交叉验证的方式确定 SVM 模型参数,即采用径向

基核函数、样本的影响半径的倒数设置为 0.0067、惩罚因子 C 设置为 1500。3D-CNN 模型使用双卷积结构,参考 3D-RCNN 模型进行参数验证和调整,结构由 P 为 $21 \times 21 \times 48$ 的输入层、3 个双卷积单元、1 个全连接层、1 个 Softmax 分类器组成。其中,卷积核尺寸为 $5 \times 5 \times 5$,步长为 1,卷积核数量随双卷积单元数量变化而变化,依次为 16, 32, 64, 最大池化尺寸为 $2 \times 2 \times 2$ 。在全连接层后使用 BN 和 dropout 策略,dropout 系数设置为 0.4,学习率为

表5 学习率对收敛速度和总体精度的影响

Table 5 Influence of learning rate on convergence rate and overall accuracy

S	Number of iterations	OA / %
0.0001	986	90.56
0.0002	776	90.92
0.0003	602	91.47
0.0004	474	91.69
0.0005	522	91.22
0.0006	439	90.03
0.0007	659	89.87
0.0008	755	90.21
0.0009	710	90.42
0.0010	928	89.15

0.0001。使用预留的验证集进行混淆矩阵精度验证,评价指标有 OA、Kappa 系数、用户精度(UA)及制图精度(PA)。表 6 为各算法精度混淆矩阵,Species 1~7 分别是麻栎、朴树、黄檀、马尾松、湿地松、杉木及其他树种。从表 6 可以看出:3D-RCNN 的总体精度达 91.72%,明显优于 3D-CNN (85.65%)和 SVM 算法(85.22%),识别准确率更高;在算法的稳定性上,3D-RCNN 的 Kappa 系数为 0.849,在 3 种算法中最高,说明识别结果更加稳定;从单一树种的识别精度上看,在 7 类树种的用户精度和制图精度共 14 个精度指标中,3D-RCNN 在 12 个精度指标上均取得了最优值,识别精度明显优于 3D-CNN 和 SVM 算法。

表6 各算法分类精度评价矩阵

Table 6 Classification accuracy evaluation matrix of each algorithm

Algorithm	Parameter	Species 1	Species 2	Species 3	Species 4	Species 5	Species 6	Species 7
3D-RCNN	UA / %	90.29	88.53	93.85	97.14	93.24	90.26	91.63
	PA / %	95.63	92.26	97.46	86.97	92.76	89.57	94.96
	OA / %				91.72			
	Kappa				0.849			
3D-CNN	UA / %	80.72	87.51	90.23	88.76	87.57	87.69	81.11
	PA / %	90.12	83.66	87.55	84.84	86.89	80.19	82.68
	OA / %				85.65			
	Kappa				0.820			
SVM	UA / %	87.50	77.50	92.50	95.00	72.50	82.50	91.67
	PA / %	94.59	83.78	94.87	86.36	72.50	86.84	75.86
	OA / %				85.22			
	Kappa				0.827			

为进一步验证 3D-RCNN 的应用精度,基于皇甫山林场提供的 2018 年森林资源一类清查数据内小班“树种组成”字段,对研究区内各树种面积进行统计,并与模型识别面积进行对比分析。以相对精度(RA)作为评价指标,从各树种识别面积的角度对模型精度进行评价,相对精度的计算公式为

$$R_{A,t,r} = [1 - (|S_{t,r} - S_{m,r}| \div S_{t,r})] \times 100, \quad (3)$$

式中: $S_{t,r}$ 为树种 r 的一类清查数据统计面积; $S_{m,r}$ 为树种 r 模型识别面积。

结果如表 7 所示,可以看出:3D-RCNN 树种识别模型能够很好地对各类树种分布面积进行提取;7 种林地类型树种调查相对精度都在 80%以上,平均相对精度达 92.92%;对于面积较小、样本丰富的树种,如黄檀、杉木和马尾松,树种调查相对精度可以达 97%以上;朴树在林场内分布较广,面积占比也

较大,而采集制作的样本相对而言较少,导致特征提取不充分,相对精度仅为 88.94%;其他树种由于包含的树种较为复杂,且分布零散,所以识别相对精度最低,只有 81.89%。在这里需要说明的是,基于一类清查数据的统计面积之和大于基于模型的识别面积,这是因为基于模型的树种识别是在林地提取的基础上进行的,对林场内的道路、迹地等未种植树木或未成林的地进行了剔除。

基于一类清查数据和模型识别结果绘制树种分布图,如图 9 所示。一类清查数据树种分布图是以小班为单位的,将优势树种(占比大于 50%)作为小班的树种属性,再因为亚热带森林林分结构复杂,存在多样的混交林,一类清查数据无法准确描述树种的空间分布。3D-RCNN 模型识别结果更加精细,能够识别出小班范围内的不同树种,准确地描述树种空间分布情况,对林业生产具有工程化的应用意义。

表 7 树种面积精度验证

Table 7 Accuracy verification of tree species area

Parameter	Species 1	Species 2	Species 3	Species 4	Species 5	Species 6	Species 7
Statistical area /km ²	7.815	10.418	1.574	4.227	4.754	1.708	4.721
Identified area /km ²	7.156	9.266	1.590	4.350	5.047	1.677	3.866
RA /%	91.57	88.94	99.00	97.08	93.83	98.15	81.89
Average RA /%	92.92						

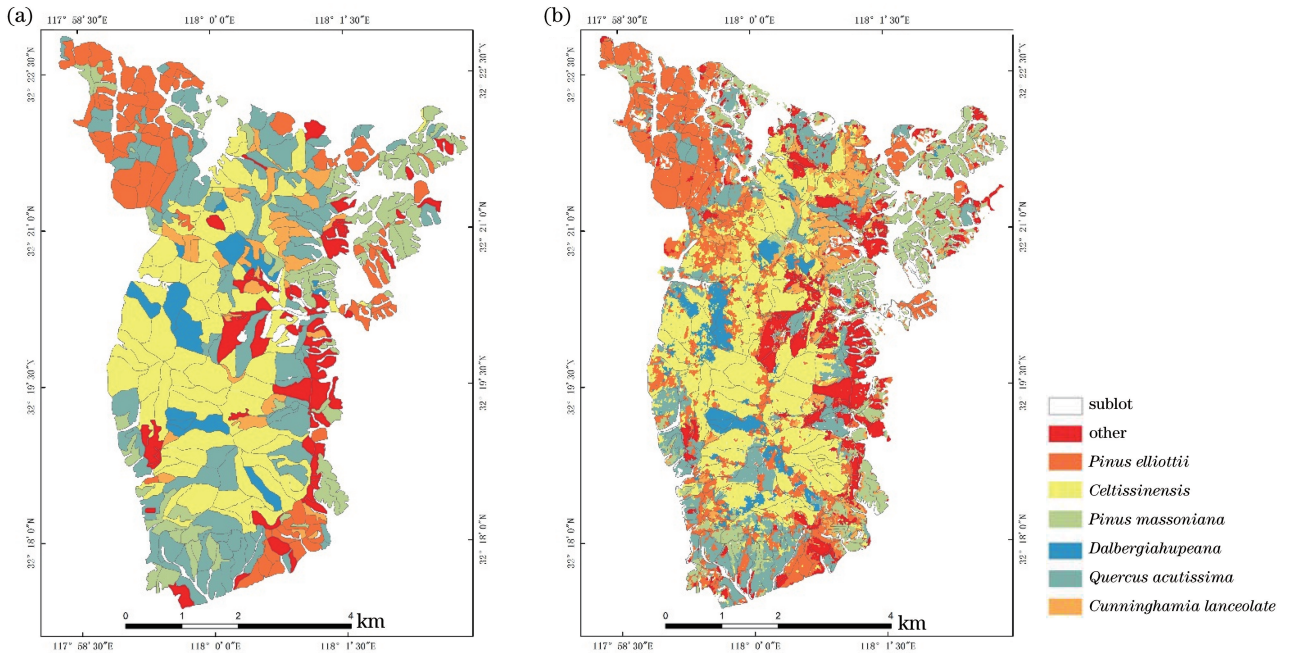


图 9 树种分布图。(a)森林清查；(b)算法提取

Fig. 9 Tree species distribution. (a) Forest inventory; (b) algorithm extraction

4 结 论

基于残差网络思想对 3D-CNN 进行改进,将模型深度增加了 6 层,总体精度和 Kappa 系数分别提高 6.07 个百分点和 0.029(详见表 6),表明改进后的 3D-RCNN 可以更充分地提取高维数据的立体特征,有效地深化了网络结构,减小额外网络层数带来的误差。同时对比 SVM 算法,可以看出,3D-RCNN 在精度和稳定性上要优于传统机器学习方法。在样本相对充足的朴树(99.00%)和杉木(98.15%)上,3D-RCNN 识别表现更加优异,这表明样本集越丰富,模型越有优化潜力。在联合 GF-5 AHIS 和 GF-6 PMS 数据的树种识别方面,传统机器学习算法 SVM、3D-CNN、3D-RCNN 所取得的总体精度均大于 85%,Kappa 系数也均超过 0.82,证明国产 GF-5 AHIS 和 GF-6 PMS 数据联合可以有效地对北亚热带常绿阔叶林和常绿针叶林进行识别,在林业的生产经营上有极大的应用潜力。

参 考 文 献

- [1] Cohen W B, Yang Z Q, Healey S P, et al. A LandTrendr multispectral ensemble for forest disturbance detection [J]. Remote Sensing of Environment, 2018, 205: 131-140.
- [2] Clark M L, Roberts D A, Clark D B. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales [J]. Remote Sensing of Environment, 2005, 96(3/4): 375-398.
- [3] Wang Z W, Sun J J, Yu Z Y, et al. Review of remote sensing image classification based on support vector machine[J]. Computer Science, 2016, 43(9): 11-17, 31.
王振武, 孙佳骏, 于忠义, 等. 基于支持向量机的遥感图像分类研究综述[J]. 计算机科学, 2016, 43(9): 11-17, 31.
- [4] Zheng Y, Chen Q Q, Zhang Y J. Deep learning and its new progress in object and behavior recognition [J]. Journal of Image and Graphics, 2014(2): 175-

- 184.
- 郑胤, 陈权崎, 章毓晋. 深度学习及其在目标和行为识别中的新进展[J]. 中国图象图形学报, 2014(2): 175-184.
- [5] Hu W, Huang Y Y, Wei L, et al. Deep convolutional neural networks for hyperspectral image classification [J]. Journal of Sensors, 2015 (2): 258619.
- [6] Chen W X, Picard R W. Predicting perceived emotions in animated GIFs with 3D convolutional neural networks [C] // 2016 IEEE International Symposium on Multimedia (ISM), December 11-13, 2016, San Jose, CA, USA. New York: IEEE Press, 2016: 367-368.
- [7] Yan M, Zhao H D, Li Y H, et al. Multi-classification and recognition of hyperspectral remote sensing objects based on convolutional neural network [J]. Laser & Optoelectronics Progress, 2019, 56(2): 021702.
- 闫苗, 赵红东, 李宇海, 等. 基于卷积神经网络的高光谱遥感地物多分类识别[J]. 激光与光电子学进展, 2019, 56(2): 021702.
- [8] Liu Y Z, Jiang Z Q, Ma F, et al. Hyperspectral image classification based on hypergraph and convolutional neural network [J]. Laser & Optoelectronics Progress, 2019, 56(11): 111007.
- 刘玉珍, 蒋政权, 马飞, 等. 基于超图和卷积神经网络的高光谱图像分类[J]. 激光与光电子学进展, 2019, 56(11): 111007.
- [9] Li Z Q, Zhu R F, Gao F, et al. Hyperspectral remote sensing image classification based on three-dimensional convolution neural network combined with conditional random field optimization[J]. Acta Optica Sinica, 2018, 38(8): 0828001.
- 李竺强, 朱瑞飞, 高放, 等. 三维卷积神经网络模型联合条件随机场优化的高光谱遥感影像分类[J]. 光学学报, 2018, 38(8): 0828001.
- [10] 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), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [11] Liu J W, Liu Y, Luo X L. Research and development on deep learning [J]. Application Research of Computers, 2014, 31(7): 1921-1930, 1942.
- 刘建伟, 刘媛, 罗雄麟. 深度学习研究进展[J]. 计算机应用研究, 2014, 31(7): 1921-1930, 1942.
- [12] Wang X P. Protection and utilization of medicinal plant resources in Huangfu mountain nature reserve [J]. Chinese Traditional and Herbal Drugs, 2001, 32(7): 669-670.
- 王晓鹏. 皇甫山自然保护区药用植物资源保护与利用[J]. 中草药, 2001, 32(7): 669-670.
- [13] Fan B, Chen X, Li B C, et al. Technical innovation of optical remote sensing payloads onboard GF-5 satellite[J]. Infrared and Laser Engineering, 2017, 46(1): 8-14.
- 范斌, 陈旭, 李碧岑, 等. “高分五号”卫星光学遥感载荷的技术创新[J]. 红外与激光工程, 2017, 46(1): 8-14.
- [14] Li G D, Zhang C J, Gao F, et al. Double convpool-structured 3D-CNN for hyperspectral remote sensing image classification [J]. Journal of Image and Graphics, 2019, 24(4): 639-654.
- 李冠东, 张春菊, 高飞, 等. 双卷积池化结构的 3D-CNN 高光谱遥感影像分类方法[J]. 中国图象图形学报, 2019, 24(4): 639-654.