

引用格式: LIANG Lianhui, LI Jun, ZHANG Shaoquan. Hyperspectral Images Classification Method Based on 3D Octave Convolution and Bi-RNN Attention Network[J]. Acta Photonica Sinica, 2021, 50(9):0910001

梁联晖, 李军, 张绍泉. 基于 3D Octave 卷积和 Bi-RNN 注意力网络的高光谱图像分类方法[J]. 光子学报, 2021, 50(9):0910001

基于 3D Octave 卷积和 Bi-RNN 注意力网络的 高光谱图像分类方法

梁联晖¹, 李军², 张绍泉¹

(1 湖南大学 电气与信息工程学院, 长沙 410082)

(2 中山大学 地理科学与规划学院, 广州 510275)

摘 要:传统卷积神经网络模型在高光谱图像分类生成特征图的空间维度中存在大量的空间特征信息冗余,而且把高光谱图像单个像元上的光谱带数据看作是无序高维向量进行数据处理,并不符合光谱数据的特性,极大影响了模型的运行效率和分类性能。针对该问题,提出一种三维 Octave 卷积和双向循环神经网络注意力网络相结合的高光谱图像分类方法。首先,利用三维 Octave 卷积获取高光谱图像的空间特征的同时减少空间特征冗余信息。其次,利用 Bi-RNN 光谱注意力网络将光谱带数据视为有序序列以获取高光谱图像的光谱信息。然后,通过全连接层将空间和光谱特征图连接起来实现特征融合。最后,经过 softmax 输出分类结果。实验结果表明,所提方法在 Pavia University 和 Botswana 两个数据集上的分类精度分别达到了 99.97% 和 99.79%,与其他主流算法相比,该方法可以充分利用空间和光谱特征信息,具有更佳分类性能。

关键词:高光谱图像分类;卷积神经网络;三维 Octave 卷积;双向循环神经网络;注意力网络

中图分类号: TP751.1

文献标识码: A

doi: 10.3788/gzxb20215009.0910001

Hyperspectral Images Classification Method Based on 3D Octave Convolution and Bi-RNN Attention Network

LIANG Lianhui¹, LI Jun², ZHANG Shaoquan¹

(1 College of Electrical and Information Engineering, Hunan University, Changsha 410082, China)

(2 School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China)

Abstract: The traditional convolutional neural network model has substantial spatial feature information redundancy exists in the spatial dimension of the feature maps in hyperspectral image classification, and the spectral band data on a single pixel of the hyperspectral image are regarded as a disordered high dimensional vector for data processing, which does not conform to the characteristics of spectral data, which greatly affects the operational efficiency of the model and the performance of classification. In order to address this problem, a hyperspectral images classification method combined three-dimensional Octave convolution with bi-directional recurrent neural network attention network is proposed. Firstly, the 3D Octave convolution is exploited to capture spatial features of the hyperspectral image, and reduce spatial feature redundant information. Secondly, Bi-RNN spectral attention network is applied to regard spectral bands

基金项目: 湖南省重点研发计划(No. 2019SK2102), 中国博士后科学基金面上项目(No. 2020M672483)

第一作者: 梁联晖(1993—), 男, 博士研究生, 主要研究方向为高光谱图像分析、信号处理和深度学习。Email: Lianglh@hnu.edu.cn

导师: 李军(1982—), 女, 教授, 博士, 主要研究方向为高光谱图像分析、信号处理、监督与半监督学习和主动学习。Email: ljun8206@hnu.edu.cn

通讯作者: 张绍泉(1990—), 男, 博士后, 主要研究方向为高光谱遥感图像处理、机器学习等。Email: zhangshaoquan1@163.com

收稿日期: 2021-01-04; 录用日期: 2021-05-06

<http://www.photon.ac.cn>

data as an ordered sequence to obtain spectral information of the hyperspectral image. Then, the spatial and spectral feature maps are connected by means of the fully connected layer to achieve features merge. Finally, the results of classification are outputted through softmax. Experimental results demonstrate that the classification accuracy of the method proposed reaches 99.97% and 99.79% in Pavia University and Botswana datasets. Compared with other mainstream methods, the proposed method can fully exploit spectral and spatial feature information, and own more competitive classification performance.

Key words: Hyperspectral image classification; Convolutional neural network; Three dimensional Octave convolution; Bi-directional recurrent neural network; Attention network

OCIS Codes: 100.5010; 100.4145; 100.3008; 110.4234

0 引言

高光谱成像也称为成像光谱,它捕获从同一区域内反射或发射的几十甚至数百个狭窄连续光谱带中的电磁能^[1-3]。高光谱图像分类是根据样本特征给每个像元分配一组特定的类别标签^[3-4],该技术被广泛应用于环境监测、城乡规划、矿物开采、国防建设和精准农业等领域^[5-6]。由于其具有丰富的空间信息和光谱信息,给分类任务带来了巨大挑战。这使得高光谱图像分类成为近年来一个热门的研究方向,受到国内外学者的广泛关注^[3-4]。

目前,高光谱图像分类方法大致可分为基于光谱信息的分类方法、基于空间-光谱特征联合的分类方法和深度学习分类方法三类。第一类方法仅利用高光谱图像中的光谱维度信息,忽略了空间上像素之间的相关性,如支持向量机(Support Vector Machine, SVM)^[7]、随机森林^[8]、稀疏表示等方法^[5,9];第二类方法在一定程度上提高了高光谱图像的分类性能,如边缘保留滤波^[10]、多尺度自适应^[11]、低秩 Gabor 滤波^[12]、分层引导滤波与最近邻正则化子空间^[13]等方法,但它们很大程度上取决于手工制作特征。即分类图的效果主要由低级特征决定,然而这并不能表示高光谱图像中复杂的内容,使得分类性能受到限制;第三类方法相较于前两类传统浅层分类方法,具有更强的表征和泛化能力,可提取到更深层的图像特征,获得更多判别特征以获得良好的分类结果。因此,近几年深度学习方法被广泛用于高光谱图像分类^[14-20],如卷积神经网络(Convolutional Neural Network, CNN)^[14]、双通道网络^[15-17]、光谱-空间注意力网络^[16-17]等方法。

尽管这些方法获得了较好的分类效果,但基于 CNN 的模型都伴随着大量空间维度信息的冗余,在某种程度上影响模型性能。有学者采用 Octave 卷积网络很好地解决了这一问题,先将高光谱数据输入 Octave 卷积模型中进行预处理,再分别加入光谱和空间注意力网络分别进行空间光谱特征提取,然后通过数据融合模块将特征信息融合,最后经 softmax 进行分类^[18],但该模型比较复杂。同时,高光谱图像的光谱数据本质上是一类序列数据,而基于 CNN 的方法是将单个像元上的光谱数据看作是无序高维向量进行数据处理,这不符合光谱数据的特性^[19]。由于循环神经网络(Recurrent Neural Network, RNN)具有记忆功能,是用于各种顺序问题建模的流行结构。有学者将 RNN 网络用于高光谱图像中光谱特征信息的提取,并获得优于基于 CNN 模型分类性能结果^[20]。

基于上述研究,本文提出一种基于 3D Octave 卷积和双向循环神经网络(Bi-directional Recurrent Neural Network, Bi-RNN)注意力网络的高光谱图像分类新框架,该框架利用四个 3D Octave 卷积模块减小空间冗余信息,获取高光谱图像的空间特征;光谱信息的提取采用基于 Bi-RNN 的光谱注意力网络,然后将空间和光谱特征图通过全连接层融合得到新的空间光谱联合特征,最后再经过 softmax 训练分类器来预测分类。该方法不仅可解决空间特征信息冗余的问题,降低模型的复杂度,还可增强空间-光谱特征信息的提取,获得良好的分类性能。

1 本文方法

本文设计了一个基于 3D Octave 卷积和 Bi-RNN 光谱注意力网络新框架,提取空间光谱联合特征以进行高光谱图像分类,该框架流程如图 1 所示。该框架的核心模块由 3D Octave 卷积网络模块、Bi-RNN 注意力网络模块和空间-光谱特征融合网络模块组成。本节将对这三个模块分别展开介绍。

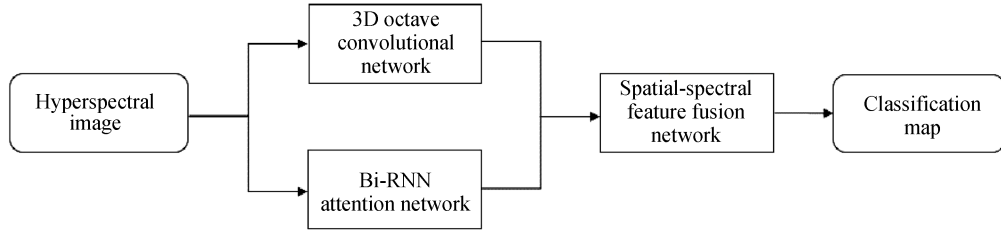


图1 3D Octave卷积和Bi-RNN注意力网络框架流程

Fig.1 Flow of framework for 3D Octave convolution and Bi-RNN attention network

1.1 3D Octave卷积网络

Octave卷积网络最初主要应用于自然图像处理中,假设自然图像信息以不同的频率传递,较高的频率通常表示局部细节特征信息,较低频率表示全局相对平滑的特征信息,卷积层的输出特征图可看作是不同的频率信息的叠加。因此,对于低频信息,可以采用共享相邻位置的信息来降低空间冗余信息。在Octave卷积网络中,从卷积层生成的特征图中分离出包含低频信息的特征图,通过池化或降采样操作,将低频分量压缩为原尺寸的四分之一,以达到减少参数数量和降低空间信息冗余的目的^[18,21]。另外,为确保信息的完整性,在两个频率之间建立了通信机制,以相互补充对应于高低部分的多样化信息。

假设一个Octave卷积网络的输入和输出数据分别为 $X = (X^H, X^L)$ 和 $Z = (Z^H, Z^L)$,其中,H和L分别表示为高频信息和低频信息。即输入和输出数据分别可表示为对应的高频信息和低频信息之和。Octave卷积模型表示为 $Z^H = Z^{H \rightarrow H} + Z^{L \rightarrow H}$ 和 $Z^L = Z^{L \rightarrow L} + Z^{H \rightarrow L}$ 。其中, $Z^{H \rightarrow H}, Z^{L \rightarrow L}$ 分别表示高频频率内和低频频率内的信息的更新, $Z^{L \rightarrow H}, Z^{H \rightarrow L}$ 分别表示低频到高频频率间和高频到低频频率间的信息转换。为了完成高频和低频特征信息的更新和转换,Octave卷积模型对应的权重参数为 $W = (W^H, W^L)$ 。同样,权重参数 W^H 和 W^L 分别可表示为 $W^L = (W^{L \rightarrow L}, W^{H \rightarrow L}), W^H = (W^{H \rightarrow H}, W^{L \rightarrow H})$,其中, $W^{H \rightarrow H}, W^{L \rightarrow L}$ 表示对应同频率内的信息更新权重, $W^{H \rightarrow L}, W^{L \rightarrow H}$ 表示对应频率间的信息转换权重。故 Z^H 和 Z^L 分别可以表示为

$$Z^H = Z^{H \rightarrow H} + Z^{L \rightarrow H} = \sum (W^H)^T X = \sum (W^{H \rightarrow H})^T X^H + \text{up} \left(\sum (W^{L \rightarrow H})^T X^L \right) \quad (1)$$

$$Z^L = Z^{L \rightarrow L} + Z^{H \rightarrow L} = \sum (W^L)^T X = \sum (W^{L \rightarrow L})^T X^L + \sum (W^{H \rightarrow L})^T \text{pool}(X^H) \quad (2)$$

式中,T表示矩阵转置,up表示上采样操作,pool表示平均池化操作。

为了获取高光谱图像更全面的空间光谱信息,采用3D Octave卷积网络,其流程如图2所示。3D Octave卷积模块包含四个3D Octave卷积块,一个平均池化运算层和上采样运算层。因为原始输入的高光谱图像数据具有完整的空间和光谱信息,故被视为高频信号 X^H 。因此,第一个3D Octave卷积网络只有高频数据

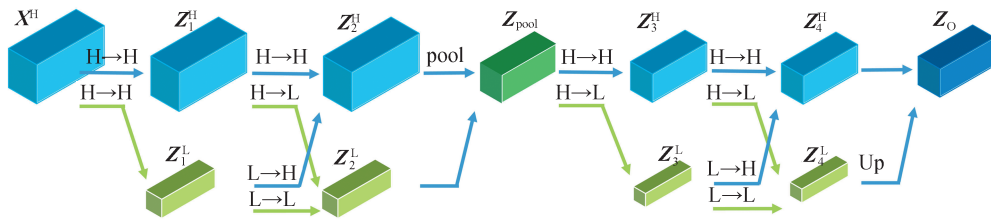


图2 3D Octave卷积网络流程

Fig.2 Flow of 3D Octave convolution network

X^H 输入到网络,输入数据 X 中的低频信号 X^L 为0。根据式(1)和(2)可得第一个3D Octave卷积网络输出 Z_1 的表达式为

$$Z_1 = (Z_1^L, Z_1^H) = [(0 + Z_1^{H \rightarrow L}), (Z_1^{H \rightarrow H} + 0)] = \left[\sum (W_1^{H \rightarrow L})^T \text{pool}(X^H), \sum (W_1^{H \rightarrow H})^T X^H \right] \quad (3)$$

式中, Z_1 作为第二个3D Octave卷积的输入数据 X_2 ,其中 $Z_1^{H \rightarrow H}$ 表示高频部分, $Z_1^{H \rightarrow L}$ 表示低频部分。同样可得第二个3D Octave卷积网络输出 Z_2 的表达式为

$$Z_2 = (Z_2^L, Z_2^H) = [(Z_2^{L \rightarrow L} + Z_2^{H \rightarrow L}), (Z_2^{H \rightarrow H} + Z_2^{L \rightarrow H})] = \left[\sum (W_2^L)^T Z_1, \sum (W_2^H)^T Z_1 \right] = \left[\sum (W_2^{L \rightarrow L})^T Z_1^L + \sum (W_2^{H \rightarrow L})^T \text{pool}(Z_1^H), \sum (W_2^{H \rightarrow H})^T Z_1^H + \text{up} \left(\sum (W_2^{L \rightarrow H})^T Z_1^L \right) \right] \quad (4)$$

为了降低高光谱图像的特征图的冗余信息,保留高光谱图像的重要特征。在第二个3D Octave卷积网络之后,采用池化操作对高频特征图 Z_2^H 进行下采样,再将下采样的结果和低频特征图 Z_2^L 融合为新的特征图 Z_{pool} ,其中 Z_{pool} 表示高频特征图,并将 Z_{pool} 作为第三个3D Octave卷积网络的输入。第三个3D Octave卷积的输入数据 X_3 为 Z_{pool} ,低频部分仍视为0。可得第三个3D Octave卷积网络输出 Z_3, Z_3 的表达式为

$$Z_3 = (Z_3^L, Z_3^H) = [(0 + Z_3^{H \rightarrow L}), (Z_3^{H \rightarrow H} + 0)] = \left[\sum (W_3^L)^T Z_{\text{pool}}, \sum (W_3^H)^T Z_{\text{pool}} \right] = \left[\sum (W_3^{H \rightarrow L})^T \text{pool}(Z_{\text{pool}}^H), \sum (W_3^{H \rightarrow H})^T Z_{\text{pool}}^H \right] \quad (5)$$

第四个3D Octave卷积的输入数据 X_4 为 Z_3 ,其中 $Z_3^{H \rightarrow H}$ 表示高频部分, $Z_3^{H \rightarrow L}$ 表示低频部分。同样,可得第四个3D Octave卷积网络输出 Z_4 表达式为

$$Z_4 = (Z_4^L, Z_4^H) = [(Z_4^{L \rightarrow L} + Z_4^{H \rightarrow L}), (Z_4^{H \rightarrow H} + Z_4^{L \rightarrow H})] = \left[\sum (W_4^L)^T Z_3, \sum (W_4^H)^T Z_3 \right] = \left[\sum (W_4^{L \rightarrow L})^T Z_3^L + \sum (W_4^{H \rightarrow L})^T \text{pool}(Z_3^H), \sum (W_4^{H \rightarrow H})^T Z_3^H + \text{up} \left(\sum (W_4^{L \rightarrow H})^T Z_3^L \right) \right] \quad (6)$$

此外,为了确保信息的完整性,需要将低频特征图 Z_4^L 经上采样操作后融合到 Z_4^H 中。由于该模型的输入特征图被认为是具有局部细节的高频信号,因此输出特征图 Z_0 也被视为高频。

1.2 Bi-RNN 注意力网络

与前馈神经网络不同,RNN被称为递归,是因为递归的隐藏状态,其每一步的激活取决于先前的计算,且Bi-RNN能充分利用序列中前者和后者的信息。将高光谱像素的光谱信息视为一个光谱序列时,Bi-RNN网络结构如图3所示。其接收输入光谱序列向量 X ,经过包含一个双向GRU层的隐藏层,生成输出向量 y 。对于Bi-RNN的隐藏层,它沿相反方向运行的两个隐藏层连接到单个输出,从而可以处理高光谱图像光谱序列中前后的光谱信息^[16,20]。

Bi-RNN将光谱序列一一输入,以学习前后状态的连续光谱特征。假设直接对每个光谱带的数据求和并求平均值,就意味着光谱带的每个光谱通道对分类任务的贡献权重值均相等。而实际中,光谱曲线并不是一个固定常数的直线,而是具有峰谷的连续曲线。因此,光谱中的某些重要的光谱通道应具有较大的权重,而那些次要的光谱段应赋予较小的权重。额外的注意力权重可增强光谱通道之间的光谱相关性,具有在序列中捕获上下文信息的强大功能。

为了给每个光谱通道分配合适的权重参数,突出和区分有效的特征,获取更多相关且值得注意的信息,且削弱不利于分类的信息。本文将注意力机制引入Bi-RNN中,使模型能够捕获内部光谱通道之间的相关性并进行更好地分类。

同样,将高光谱像素的整个光谱视为一个序列,采用带有注意力机制的Bi-RNN模型获取光谱信息,Bi-RNN注意力网络结构如图4所示。同样假设它的输入是一个高光谱向量 $X, X=(X_1, X_2, X_3, \dots, X_m)$,双向隐

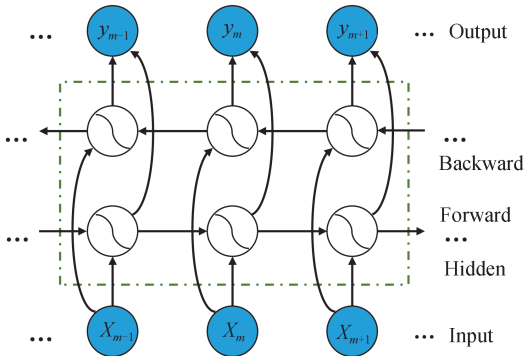


图3 Bi-RNN网络结构
Fig.3 Structure of Bi-RNN network

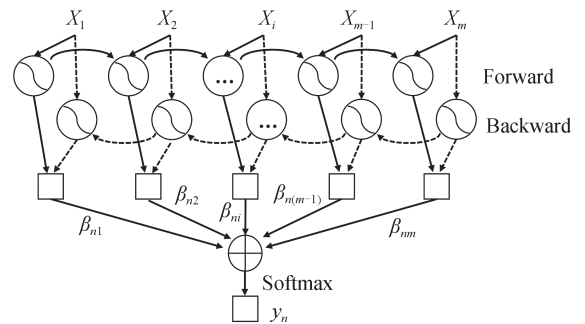


图4 Bi-RNN注意力网络结构
Fig.4 Structure of Bi-RNN attention network

藏层的计算可表示为

$$h_n^f = f(W^f X_n + U^f h_{n-1} + b^f) \quad (7)$$

$$h_n^b = f(W^b X_n + U^b h_{n+1} + b^b) \quad (8)$$

式中, n 表示光谱带的范围 $1 \sim m$,系数矩阵 W^f 和 W^b 来自当前隐藏层的输入 U^f 表示上一个隐藏状态 h_{n-1} , U^b 在后续隐藏状态中从 h_{n+1} 开始, f 是隐藏层的非线性激活函数,并且该编码器的输出作为 g_n 的输入, g_n 的表达式为

$$g_n = \text{concat}(h_n^f, h_n^b) \quad (9)$$

式中, $\text{concat}()$ 是正向隐藏状态函数和反向隐藏状态函数之间的串联函数。

在Bi-RNN模型上,添加了一个注意机制来获取不同的光谱信息的权重值以学习更多特征。该模型的注意力层可以表示为

$$e_{in} = \tanh(W_i g_n + b_i) \quad (10)$$

$$\beta_{in} = \text{softmax}(W'_i e_{in} + b'_i) \quad (11)$$

式中, W_i 和 W'_i 是门变换矩阵, b_i 和 b'_i 是偏差项, e_{in} 为一个中间参数, β_{in} 为注意力权重参数 β 中一个的组成部分,即第 i 个权重参数。而 $\text{softmax}()$ 是将非标准化输出值映射到概率分布,且将其输出值约束在 $(0,1)$ 区间内。

因此,由式(9)和式(11)可以计算出像素 X 的预测标签 y_n ,即

$$y_n = U[g_n, \beta_{in}] \quad (12)$$

式中, $U[\]$ 是对所有在对应的注意力权重加权下的状态向量求和函数。

式(10)是一层神经网络,该层重新排列了Bi-RNN的状态向量空间,然后 \tanh 激活将其转换为 e_{in} 作为 h_n 的一个新的隐藏表示。式(11)通过 softmax 层产生注意力权重 β ,其中,模型根据 e_{in} 与另一个通道矢量的相关性来衡量输入的重要性。在获得新学习的注意力权重之后,模型使用式(12)中软注意力机制来更新标签表示向量 y 。

通过该注意力机制,该模型针对高光谱图像光谱信息中的每个光谱通道对分类任务的贡献权重不同,赋予不同的权重,以更好地表征高光谱图像的光谱特征,更加关注相邻波段光谱信息的相关性,并使训练模型更加准确。

1.3 空间-光谱特征融合网络

为了充分利用光谱信息的相关性和空间特征信息并提取融合空间-光谱特征,本模型通过全连接层连接两个分支网络以进行空间-光谱信息的融合。所提3D Octave卷积和Bi-RNN注意力网络模型整体结构如图5所示,空间-光谱特征融合网络将经3D Octave卷积网络得到的三维信息数据展平为一维数据,输入全连接层,然后将该全连接层和Bi-RNN注意力网络的最后一层全连接层结合起来,形成一个新的全连接层。接着,通过一个全连接层,降低融合了空间-光谱特征的冗余信息。最后,将融合了空间和光谱特征信息的新的全连接层经过 softmax 层进行预测分类。

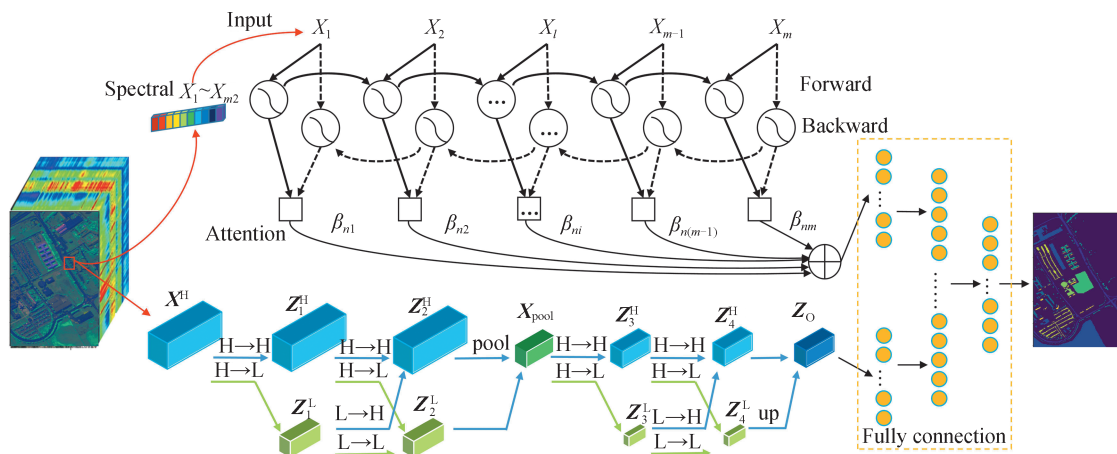


图5 3D Octave卷积和Bi-RNN注意力网络模型整体结构

Fig.5 The whole structure diagram of 3D Octave convolution and Bi-RNN attention network

2 实验结果与分析

实验硬件平台为一台高性能计算机,其配置为:英特尔 Core i9-9900K@3.60 GHz八核,32G内存,显卡为 Nvidia GeForce RTX 2080 Ti (11 GB)。软件平台为 Windows10 系统环境下的 Python3.6.0 和 TensorFlow1.14。

2.1 实验数据与样本划分

高光谱图像的分类精度评价指标采用常用的三个评价指标,总体分类精度(Overall Accuracy, OA),平均分类精度(Average Accuracy, AA)和 Kappa 系数来衡量分类的精度^[17]。为验证所提方法的分类效果,选择 Pavia University 和 Botswana 两个数据集来验证所提方法的性能。Pavia University 数据集是由反射光学成像光谱仪传感器在意大利北部帕维亚大学获得的遥感图像数据,像素大小为 610 像素×340 像素,在 430~860 nm 范围有 115 个原始光谱波段数,除去 12 个噪声波段,剩余 103 个光谱波段用于分类。在 Pavia University 数据集中定义了 9 个语义类别,每个类别样本的大小,以及实验训练样本和测试集样本数的划分如表 1 所示。

表 1 Pavia University 数据集的训练集和测试集样本数
Table 1 Number of training and testing samples of the Pavia University dataset

Code	Class name	Train	Test	Total
1	Asphalt	548	6 083	6 631
2	Meadows	540	18 109	18 649
3	Gravel	392	1 707	2 099
4	Trees	542	2 522	3 064
5	Painted metal sheets	256	1 089	1 345
6	Bare soil	532	4 497	5 029
7	Bitumen	375	955	1 330
8	Self-Blocking bricks	514	3 168	3 682
9	Shadows	231	716	947
Total		3 930	38 846	42 776

Botswana 数据集是美国航空局通过 EO-1 卫星上的 Hyperion 传感器成像光谱仪所获得的。该图像覆盖了 Botswana 三角洲地区长 7.7 km 的带状地带,空间分辨率为 30 m,光谱分辨率为 10 nm。该图像共包含 242 个波段,剔除受到噪声影响的 97 个波段后,剩余 145 个波段可用于高光谱图像的处理。同时,该图像尺寸为 1 476 像素×256 像素,含 14 个不同的类别,每个类别样本的大小,以及实验训练样本和测试集样本数的划分如表 2 所示。

表 2 Botswana 数据集的训练集和测试集样本数
Table 2 Number of training and testing samples of the Botswana dataset

Code	Class name	Train	Test	Total
1	Water	30	240	270
2	Hippo grass	30	71	101
3	Floodplain grasses1	30	221	251
4	Floodplain grasses1	30	185	215
5	Reeds1	30	239	269
6	Riparian	30	239	269
7	Firescar2	30	229	259
8	Island interior	30	173	203
9	Acacia woodlands	30	284	314
10	Acacia shrublands	30	218	248
11	Acacia grasslands	30	275	305
12	Short mopane	30	151	181
13	Mixed mopane	30	238	268
14	Exposed soils	30	65	95
Total		420	2 828	3 248

2.2 参数设置

实验中,学习率、空间大小、丢弃率三个参数会对实验造成显著的影响。以 Pavia University 数据集为例,对实验参数进行了详细的评估。

1)学习率:实验中测试了不同的学习率的影响。学习率决定着每次更新模型权重时的学习过程和分配错误量。学习率太大可能会导致训练的周期震荡,太小则会导致模型不能收敛。根据经验,学习率的大小既不能取值过大,也不能取值过小。当取大于0.01时,会导致模型训练严重不收敛;而学习率小于0.000 01时,模型收敛速度极慢,不利于模型的训练。因此,本文模型的学习率分别选择[0.01, 0.005, 0.001, 0.000 7, 0.000 5, 0.000 3, 0.000 1, 0.000 07, 0.000 05, 0.000 03, 0.000 01]进行实验,结果表明本文模型的学习率在0.000 1的时候分类效果最好。

2)空间大小:因为图像空间特征的提取,严重依赖空间邻域区域的大小。且更大的空间输入将提供更多的信息来学习更多的空间特征,但是空间区域越大,也会带来无用的干扰信息,造成图像细节信息的丢失和过度平滑现象的可能。因此,选择合适的空间大小对分类性能的提升非常重要。在光谱通道数固定不变,最优学习率,批量大小为32,训练迭代次数为400次的情况下,不同空间尺寸下的分类精度结果如表3所示。

表3 不同空间尺寸下的分类精度
Table 3 Classification accuracy of different spatial size

	Spatial size					
	11×11	13×13	15×15	17×17	19×19	21×21
Accuracy OA / %	99.86	99.93	99.97	99.96	99.96	99.91

3)丢失率:神经网络在训练过程中,神经元之间相互依赖,导致训练数据的过拟合。Dropout是一种有效的神经网络正则化方法,在神经网络的训练过程中,按照给定的丢失概率值随机将网络中的某些节点暂时丢弃,有助于减少网络相互依赖的学习和防止过拟合。通常,模型训练过程的丢失率在区间[0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0]内选取。因此,本文也不失一般性选用不同的丢失率来测试模型,不同丢失率下的精度如表4所示。

表4 不同丢失率下的分类精度
Table 4 Classification accuracy of different dropout

	Dropout							
	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Accuracy OA / %	99.82	99.87	99.91	99.94	99.97	99.96	99.96	99.95

从表3中可以看出,当选择空间尺寸为13×13时,由于空间尺寸的变小导致无法利用空间邻域之间更多的有用空间信息,从而导致精度下降。故当选择空间尺寸为11×11时,精度下降更加明显。而当空间尺寸大于21×21时,由于大的空间区域会使得图像空间细节丢失,使得模型无法提取到有效的空间细节信息,导致分类性能反而变差。因此,过大或过小的空间大小均无法充分有效地利用空间信息。从表3和表4可知,当输入数据的空间大小为15×15,丢失率为0.6时,分类效果最好。为使分类性能最佳,故本实验均选择最佳参数。

同样在 Botswana 数据集中进行相同的实验,为了使分类性能最佳,实验的学习率设置为0.000 1,空间大小为13×13,批量大小为16,丢失率为0.7,训练迭代次数为400次。

2.3 实验结果分析

在以上参数的设置下,为了验证所提方法(简称3DOC-RNN)的有效性和优越性,将3DOC-RNN方法与一些传统的方法及主流的深度学习方法(如SVM^[7]、ARNN^[16]、SSAN^[16]、DBMA^[17]、3DOC-SSAN^[18])进行实验对比。为了保证实验结果的准确性,不同方法在 Pavia University 数据集和 Botswana 数据集上重复进行10次实验,再取实验结果的平均值。表5为不同方法在 Pavia University 数据集上的分类性能对比实验结果。其中,加粗表示最高精度。

从表5可以得出,在Pavia University数据集上本文所提方法各项性能明显要比传统方法SVM好,本文所提方法的OA值、AA值和Kappa值均比其他主流深度学习分类方法精度高。其中,3DOC-SSAN与其他方法在Pavia University数据集上分类精度差值比较如表6所示。其中,“+”号表示3DOC-SSAN比该方法的精度值高,“-”号表示反之。

表5 不同方法在Pavia University数据集的分类性能
Table 5 Classification performance of different methods on Pavia University dataset

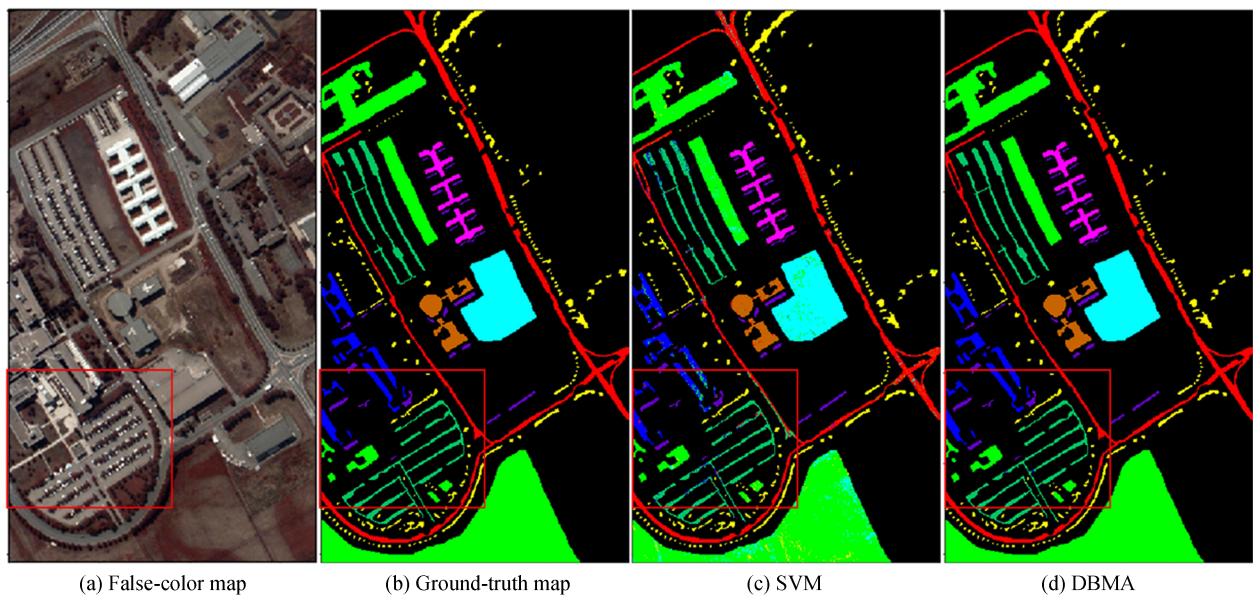
Class	SVM	DBMA	ARNN	SSAN	3DOC-SSAN	3DOC-RNN
1	89.33	99.98	98.93	99.28	99.82	99.93
2	93.84	99.96	96.82	98.66	99.93	99.99
3	85.82	99.91	97.71	98.53	99.67	99.88
4	97.86	99.18	99.72	98.66	99.81	99.96
5	98.99	99.90	100.00	100.00	100.00	100.00
6	94.95	99.30	99.84	99.40	100.00	100.00
7	94.14	99.92	99.69	99.79	99.76	100.00
8	89.96	96.82	99.72	99.75	99.80	99.94
9	99.98	99.33	100.00	99.72	100.00	100.00
OA/%	93.12	99.52	98.23	99.00	99.87	99.97
AA/%	90.70	99.37	99.17	99.30	99.87	99.96
Kappa/%	93.88	99.35	97.59	98.64	99.94	99.96

表6 与其他方法在Pavia University数据集上分类精度差值比较
Table 6 Comparison with other methods in the classification accuracy difference on Botswana dataset

Class	SVM	DBMA	ARNN	SSAN	3DOC-SSAN
OA/%	+6.85	+0.45	+1.74	+0.97	+0.10
AA/%	+9.26	+0.59	+0.79	+0.66	+0.09
Kappa/%	+6.08	+0.61	+1.47	+2.37	+0.02

从表6结果可以看出,三项指标都体现出了本文方法在Pavia University数据集上分类性能优于其他方法。同时,不同方法在Pavia University数据集上的分类图如图6所示。为更好地对该数据集上的分类情况进行比较,本文对分类图采用局部区域放大比较。对应图6中的红框区域,在Pavia University数据集上分类图局部放大对比图如图7所示。

从图6和7可以看出,SVM、DBMA、ARNN和SSAN的最终分类结果均有大量杂乱的斑点,并且有些



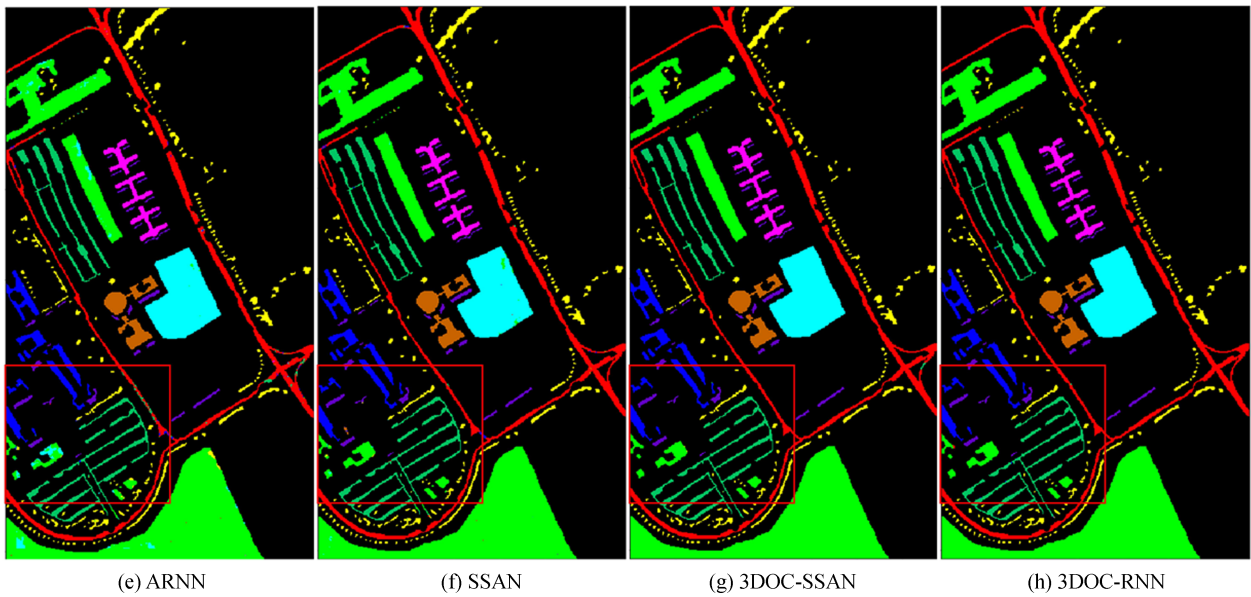


图6 不同方法在Pavia University数据集上的分类图
Fig.6 Classification maps of different methods on the Pavia University dataset

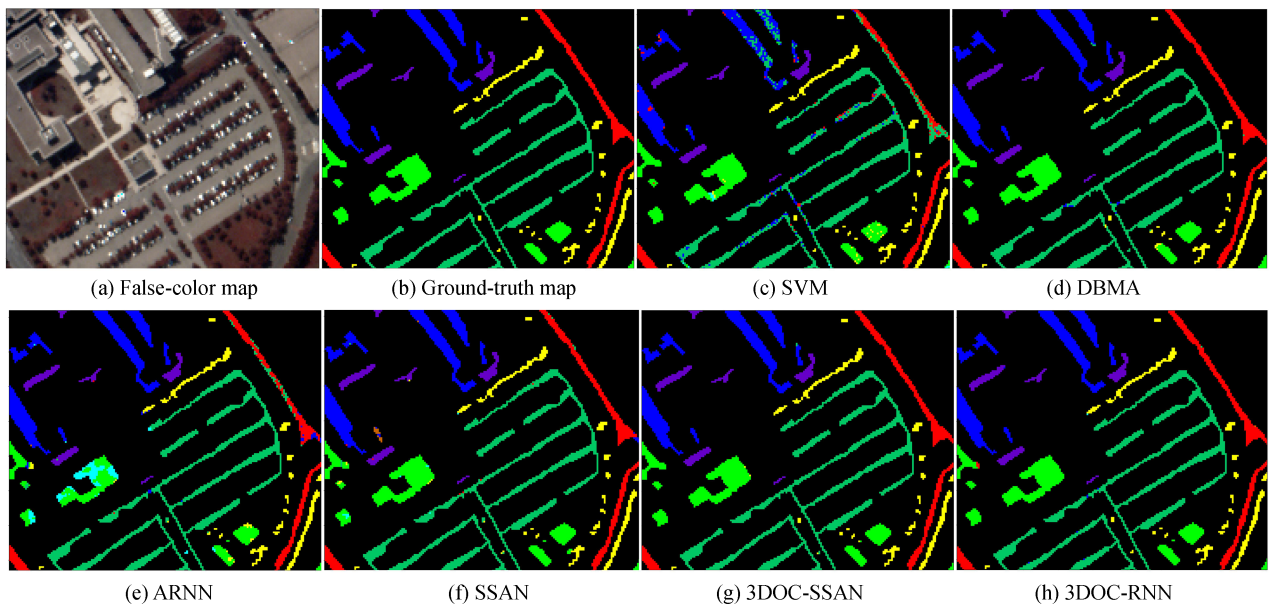


图7 在Pavia University数据集上的分类图局部放大对比
Fig.7 Partial enlargement comparison of classification maps on the Pavia University dataset

区域存在明显错误分类的现象。3DOC-SSAN和3DOC-RNN方法在图7对应的分类图上虽然均存在较少的噪点,但3DOC-SSAN方法在图6对应图中的右下角存在明显的一块错分类区域。而3DOC-RNN方法在其他区域的地物基本被完全正确分类,几乎看不出孤立噪声点,而且在同质区域相对平滑。

不同方法在Botswana数据集上的分类性能对比实验结果如表7所示。其中,加粗表示最高精度。

同时,在该数据集上的分类图如图8所示,为了更好地对Botswana数据集上的分类结果图进行比较,本文对分类图采用局部区域放大进行比较。对应图8中的红框区域,在Botswana数据集上分类图局部放大对比图如图9所示。

从表7可以得出,在Botswana数据集上本文所提3DOC-RNN方法在OA值、AA值和Kappa值三项指标上的精度都高于其他方法,且在11个类别上的分类精度都达到100%,除了漫滩草地1的分类精度为97.38%,其余两类的分类精度也都达到99.58%以上的分类精度。其中,本文方法与其他方法在Botswana

表7 不同方法在 Botswana 数据集的分类性能
Table 7 Classification performance of different methods

Class	SVM	DBMA	ARNN	SSAN	3DOC-SSAN	3DOC-RNN
1	100.00	98.94	99.60	99.19	100.00	100.00
2	100.00	100.00	100.00	100.00	100.00	100.00
3	98.19	100.00	99.13	99.15	100.00	97.38
4	96.76	99.03	88.89	100.00	100.00	100.00
5	76.15	98.81	84.62	85.49	96.65	100.00
6	74.48	98.73	99.19	98.39	97.91	99.58
7	95.63	100.00	97.05	100.00	100.00	100.00
8	98.84	100.00	100.00	99.45	100.00	99.94
9	85.21	100.00	98.97	96.56	100.00	100.00
10	89.91	100.00	94.74	100.00	100.00	100.00
11	95.27	100.00	100.00	100.00	100.00	100.00
12	92.05	100.00	100.00	100.00	98.01	100.00
13	88.24	100.00	100.00	100.00	100.00	100.00
14	100.00	100.00	97.14	95.38	100.00	100.00
OA/%	90.91	99.62	97.02	98.00	99.43	99.79
AA/%	90.14	99.67	97.10	98.12	99.47	99.81
Kappa/%	92.20	99.59	96.76	97.83	99.39	99.77

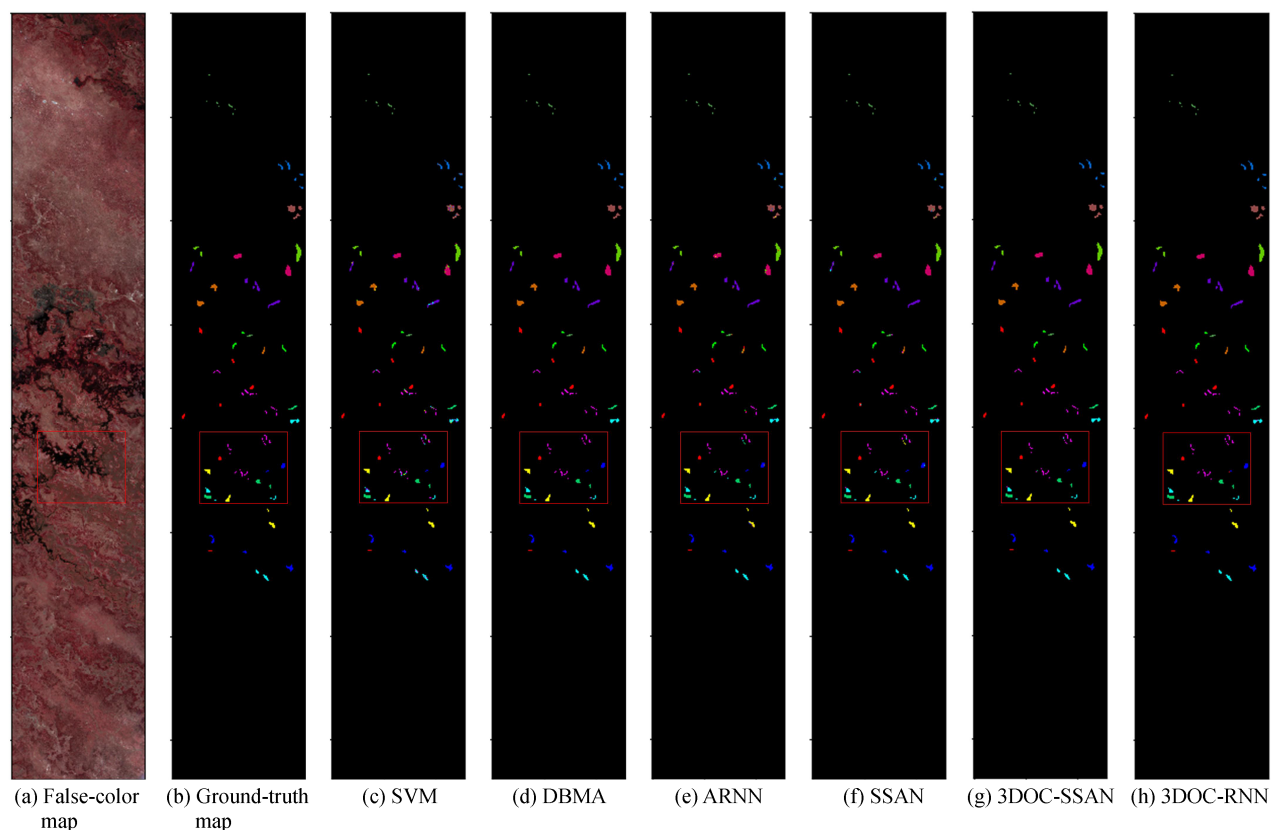


图8 不同方法在 Botswana 数据集上的分类图
Fig.8 Classification maps of different methods on the Botswana dataset

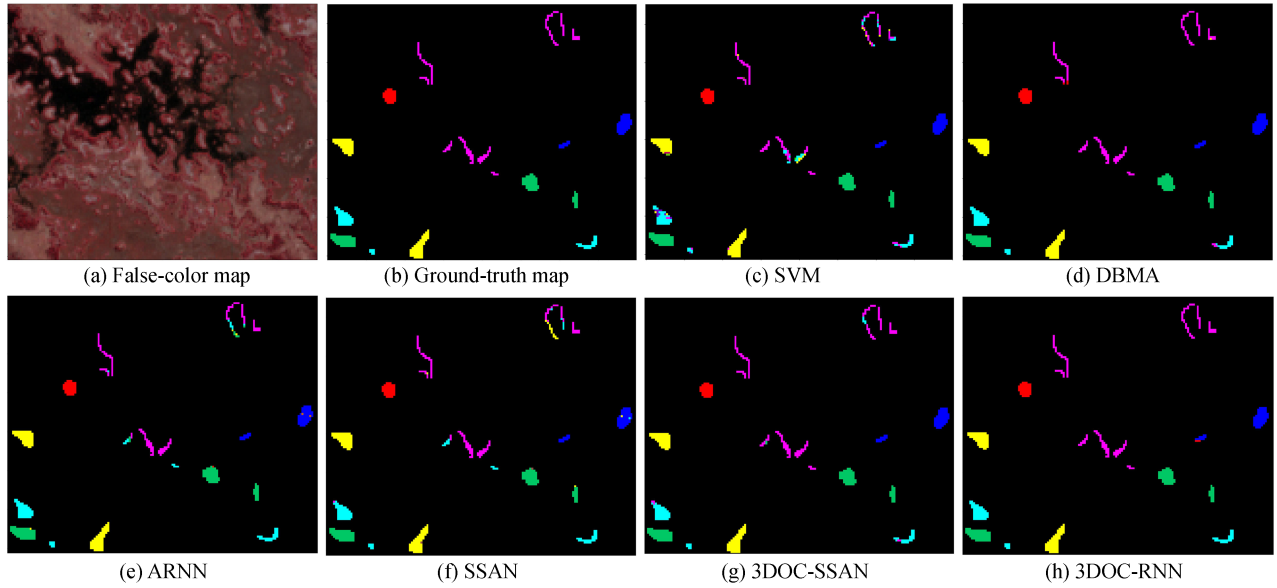


图9 Botswana数据集上的分类图局部放大对比

Fig.9 Partial enlargement comparison of classification maps on the Botswana dataset

数据集上分类精度差值比较如表8所示。其中,“+”号表示3DOC-SSAN比该方法的精度值高,“-”号表示反之。

表8 与其他方法在Botswana数据集上分类精度差值比较

Class	SVM	DBMA	ARNN	SSAN	3DOC-SSAN
OA/%	+8.88	+0.17	+2.77	+1.79	+0.36
AA/%	+9.67	+0.14	+2.71	+1.69	+0.34
Kappa/%	+7.57	+0.18	+3.01	+1.94	+0.38

从图8和9的分类图可以看出,3DOC-RNN方法的分类图虽然存在一处错分类点,但比其他几种方法的错分类处明显要少,分类效果更佳。

由表6和8可以得出,利用3D Octave卷积网络的3DC-SSAN分类方法和本文方法明显比DBMA、ARNN和SSAN方法的性能要好,这证明了3D Octave卷积在减少空间冗余信息,提高分类性能上的有效性。而3DOC-RNN方法又比不加Bi-RNN注意力网络的3DOC-SSAN方法的分类性能好,说明Bi-RNN注意力网络在增强光谱特征信息的提取方面具有一定的优势,有利于分类性能的提升。

此外,3DOC-RNN模型与3DOC-SSAN模型相比,3DOC-RNN方法模型待3D Octave卷积网络执行完后,就可直接将注入空间光谱特征信息融合网络,无需额外加入空间注意力网络模块,模型相对简单,且3DOC-RNN方法的空间光谱特征提取网络数据流是并行的。而3DOC-SSAN方法的数据流是串行的,需先将高光谱数据输入Octave卷积模型中进行预处理后,才能将数据再分别加入光谱和空间注意力网络分别进行空间光谱特征提取,然后经数据融合模块将特征信息融合,最后进行预测分类^[18]。因此,模型训练时前者可以采用并行处理,当程序采用并行处理方式运行时,相比之下可极大提高模型的运行效率。由于实验硬件条件的限制,无法采用并行方式进行实验,因此通过采用分模块运行的方式对两种方法功能模块的运行时间进行比较分析。以Pavia University数据集为例,在最佳参数设置下,迭代一次两种方法功能模块的运行时间对比如表9所示。其中,3DOC-Octave表示3DOC-RNN模型中的不包含Bi-RNN光谱注意力网络的模型;3D-OCM表示3DOC-SSAN模型中将高光谱数据输入Octave卷积模型中直接进行数据处理的网络,SSAM+SSICM表示3DOC-SSAN模型经Octave卷积模型进行预处理后,数据分别加入光谱和空间注意力网络分别进行空间光谱特征提取,然后经数据融合模块将特征信息融合模块。

由表9可得,3D Octave卷积网络的运行时间大于Bi-RNN的运行时间,故采用并行方式时,3D Octave

表9 两种方法的运行时间对比
Table 9 Comparison of the running time of the two methods

Class	3D Octave	Bi-RNN	3D-OCM	SSAM+SSICM
3DOC-RNN/s	26.212	16.062	—	—
3DOC-SSAN/s	—	—	26.635	0.766 9

网络对空间信息提取完毕就可以执行下一步程序,3DOC-RNN在Pavia University数据集上并行方式迭代一次运行时间约为26.21 s。而3DOC-SSAN网络的运行时间为3D-OCM与SSAM+SSICM网络运行时间之和,约为27.40 s。当采用并行方式训练模型时,同样迭代400次,3DOC-RNN方法比3DOC-SSAN方法快476 s,故3DOC-RNN方法训练模型运行效率更高。由于3DOC-RNN方法的主要时间开销在3D Octave卷积网络,因此,3DOC-RNN方法可通过优化用于提取空间特征信息的3D Octave卷积网络,以提高模型运行效率,这将作为下一步的研究内容。

4 结论

为了减少空间特征信息的冗余,增强光谱信息的获取,提高高光谱图像的分类性能,提出了一种基于3D Octave卷积和Bi-RNN注意力网络高光谱图像分类新模型。该模型结构简单,无需对高光谱图像数据进行复杂的预处理和后处理,可实现端到端的训练。实验结果表明,本文方法对比传统方法分类性能有很明显的提升,与当前主流深度学习算法相比,本文所提3DOC-RNN方法能更充分地提取空间和光谱特征信息,不论在分类精度还是在分类图上,都体现出更佳的性能。

参考文献

- [1] JOSE M B, ANTONIO P, GUSTAVO C, et al. Hyperspectral remote sensing data analysis and future challenges[J]. IEEE Geoscience Remote Sensing Magazine, 2013, 1(2): 6-36.
- [2] ZHANG Liangpei, ZAHGN Lefei, DU Bo. Deep learning for remote sensing data: A technical tutorial on the state of the art[J]. IEEE Geoscience Remote and Sensing Magazine, 2016, 4(2): 22-40.
- [3] HE Lin, LI Jun, LIU Chenyin, et al. Recent advances on spectral-spatial hyperspectral image classification: An overview and new guidelines[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(3): 1579-1597.
- [4] ANTONIO P, JON A B, JOSEPH W B, et al. Recent advances in techniques for hyperspectral image processing[J]. Remote Sensing of Environment, 2009, 113(1): S110-S122.
- [5] LIU Jiamin, ZHANG Limei, SHI Guangyao, et al. Hyperspectral image classification with combination of sparse characteristic and neighborhood similarity metrics[J]. Acta Photonica Sinica, 2018, 47(6): 0610001.
刘嘉敏, 张丽梅, 石光耀, 等. 联合稀疏特性和邻域相似度量的高光谱图像分类[J]. 光子学报, 2018, 47(6): 0610001.
- [6] XU Yonghao, DU Bo, ZHANG Fan, et al. Hyperspectral image classification via a random patches network[J]. Journal of Photogrammetry and Remote Sensing, 2018(142): 344-357.
- [7] PAL M. Ensemble of support vector machines for land cover classification[J]. International Journal of Remote Sensing, 2008, 29(10): 3043-3049.
- [8] HAM J, CHEN Yangchi, CRAWFORD M M, et al. Investigation of the random forest framework for classification of hyperspectral data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2005, 43(3): 492-501.
- [9] LI Wei, DU Qian. A survey on representation-based classification and detection in hyperspectral remote sensing imagery[J]. Pattern Recognition Letters, 2016, 83(2), 115-123.
- [10] KANG Xudong, LI Shutao, JOSEPH A B, et al. Spectral-spatial hyperspectral image classification with edge-preserving filtering[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(5): 2666-2677.
- [11] FANG Leyuan, LI Shutao, KANG Xudong, et al. Spectral-spatial hyperspectral image classification via multiscale adaptive sparse representation[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(12): 7738-7749.
- [12] HE Lin, LI Jun, ANTONIO P, et al. Discriminative low-rank gabor filtering for spectral-spatial hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(3): 1381-1395.
- [13] XU Dongdong, CHENG Deqiang, CHEN Liangliang, et al. Hyperspectral image classification based on hierarchical guidance filtering and nearest regularized subspace[J]. Acta Photonica Sinica, 2020, 49(4): 0410004.
徐冬冬, 程德强, 陈亮亮, 等. 基于分层引导滤波与最近邻正则化子空间的高光谱图像分类[J]. 光子学报, 2020, 49(4): 0410004.
- [14] CHEN Yushi, JIANG Hanlu, LI Chunyang, et al. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(10): 6232-6251.

- [15] BI Xiaojun, ZHOU Zengyu. Hyperspectral image classification algorithm based on two-channel generative adversarial network[J]. Acta Optica Sinica, 2019, 39(10): 297-308.
毕晓君,周泽宇. 基于双通道GAN的高光谱图像分类算法[J]. 光学学报, 2019, 39(10): 297-308.
- [16] PAN Erting, MA Yong, MEI Xiaoguang, et al. Spectral-spatial classification of hyperspectral image based on a joint attention network [C]. IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019: 413-416.
- [17] LI Rui, ZHENG Shunyi, DUAN Chenxi, et al. Classification of hyperspectral image based on double-branch dual-attention mechanism network[J]. Remote Sensing, 2020, 12(3): 582-607.
- [18] TANG Xu, MENG Fanbo, ZHANG Xiangrong, et al. Hyperspectral image classification based on 3-D octave convolution with spatial-spectral attention network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(3): 2430-2447.
- [19] GUAN Shihao, YANG Guang, LI Hao, et al. Hyperspectral image classification based on 3-D convolutional recurrent neural network[J]. Laser Technology, 2020, 44(4): 485-491.
关世豪,杨桃,李豪,等. 3维卷积递归神经网络的高光谱图像分类方法[J]. 激光技术, 2020, 44(4): 485-491.
- [20] MOU Lichao, GHAMISI Pedram, ZHU Xiaoxiang. Deep recurrent neural networks for hyperspectral image classification [J]. IEEE Transaction Geoscience and Remote Sensing, 2017, 55(7): 3639-3655.
- [21] CHEN Yunpeng, FAN Haoqi, XU Bing, et al. Drop an octave: Reducing spatial redundancy in convolutional neural networks with octave convolution [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 3435-3444.