

# 基于深度学习的高光谱影像分类方法研究

张彬<sup>1</sup>, 刘亮<sup>2</sup>, 李晓杰<sup>1</sup>, 周伟<sup>1\*</sup>

(1. 海军航空大学 航空作战勤务学院, 山东 烟台 264000;

2. 海军航空大学 岸防兵学院, 山东 烟台 264000)

**摘要:** 针对高光谱影像分类方法精度不足的问题, 提出一种基于空间-频谱变换(Spectral-Spatial Transformer, SST)网络的高光谱影像分类方法。首先, 将高光谱影像预处理为一维特征向量。然后, 设计了具有光谱-空间注意力模块和池化残差模块的SST高光谱影像分类网络。本文所提出的分类方法在Indian Pines数据集和Pavia University数据集上的总体分类精度分别为98.67%和99.87%, 表明此方法具有较高的分类精度, 为高光谱影像分类及应用提供了一种新方案。

**关键词:** 深度学习; 高光谱影像; 分类; 遥感图像

中图分类号: TP18 文献标识码: A

## Research on hyperspectral image classification method based on deep learning

ZHANG Bin<sup>1</sup>, LIU Liang<sup>2</sup>, LI Xiao-Jie<sup>1</sup>, ZHOU Wei<sup>1\*</sup>

(1. Aviation Operations and Service Institute, Naval Aviation University, Yantai 264000, China;

2. Coastal Defense College, Naval Aviation University, Yantai 264000, China)

**Abstract:** Targeting the issue of insufficient accuracy of hyperspectral image classification methods, a hyperspectral image classification method based on Spatial-spatial transformer (SST) network is proposed. Firstly, the hyperspectral images are preprocessed into one-dimensional feature vectors. Then, the SST hyperspectral image classification network with spectral-spatial attention module and pooled residual module is designed. The overall classification accuracy of the proposed classification method on Indian Pines dataset and Pavia University dataset is 98.67% and 99.87%, respectively, which indicates that this method has high classification accuracy and provides a new scheme for hyperspectral image classification and application.

**Key words:** deep learning, hyperspectral image, classification, satellite imagery

### 引言

高光谱影像分类作为高光谱影像处理、分析和应用的关键技术之一, 其目的是对每一个影像像素确定唯一的地物类别<sup>[1-3]</sup>。由于高光谱影像的数据量快速增长, 现有的分类技术难以满足对海量高光谱影像数据进行快速、精准和自动影像解释的需求<sup>[4]</sup>。目前高光谱影像分类方法主要分为两大类: 基于传统方法的高光谱影像分类方法和基于深度学习的高光谱影像分类方法。

基于传统方法的高光谱影像分类方法是采用

统计学方法对光谱信息的相似性进行度量, 进而对高光谱像素加以区分。这种方法的缺点是效率低, 需要耗费大量的人工劳动。常见的匹配分类方法有: 二值编码匹配法<sup>[5]</sup>、光谱角度匹配法、交叉相关光谱匹配和聚类分析技术。由于高光谱影像存在“同物异谱、异物同谱”的现象, 这类统计学方法能够达到的精度有限。随着机器学习的发展, 研究者不断尝试在高光谱影像分类中使用机器学习方法。机器学习方法通常需要先对数据作特征工程, 然后对预处理后的特征使用分类器分类。常用的特征

收稿日期: 2023-01-06, 修回日期: 2023-07-22

基金项目: 国家自然科学基金(62005318)

Foundation items: Supported by the National Natural Science Foundation of China (62005318)

作者简介(Biography): 张彬(1981—), 男, 山东临沂人, 讲师, 主要研究方向为高光谱成像、图像处理. E-mail: gfkdzhb@163.com

\*通讯作者(Corresponding author): E-mail: yeaweam@163.com

Received date: 2023-01-06, Revised date: 2023-07-22

工程方法有主成分分析<sup>[6]</sup>、独立成分分析<sup>[7]</sup>、线性判别分析<sup>[8]</sup>等等。上述方法具有减少特征冗余、增大类间差异的作用,又因为其较好的性能,这些方法也常常用作深度学习方法的特征预处理。常用的分类器有K-最近邻<sup>[9]</sup>、支持向量机<sup>[10]</sup>、随机森林<sup>[11]</sup>、极限学习机<sup>[12]</sup>、决策树<sup>[13]</sup>等等。这些机器学习方法和之前的统计学方法相比,均获得了不小的性能提升。

在利用深度学习进行高光谱影像分类方面,Chen等人<sup>[14]</sup>的方案首先利用PCA降低光谱维数,其次设计了三种基于卷积神经网络(Convolutional Neural Networks, CNN)的深度有限元架构,分别提取光谱、空间以及光谱-空间特征,相较于传统分类技术,利用CNN网络提高了影像的分类性能。Li等人<sup>[15]</sup>利用中心像素块与其邻域内相邻像素块之间具有高概率属于同一类成像单位的特点,构建像素对联合分类模型,提高分类精度。Xie等人<sup>[16]</sup>提出了一种多尺度密集连接卷积网络,通过提取多尺度patch构造多尺度样本,利用dense block融合多尺度样本特征,相较于以上单尺度CNN网络,MS DenseNET能够更好地利用高光谱影像中复杂的空间结构信息进行分类。除提高分类精度之外,许多学者也在其他方面进行深入研究。Zheng等人<sup>[17]</sup>通过引入注意力机制模块去除冗余光谱并基于构建矫正空间注意力模块提出一种旋转不变注意网络,缓解由于成像仪旋转导致的分类精度下降问题。由于高光谱影像中包含大量不同波段的频谱信息,现有的基于深度学习的高光谱影像分类方法大多数采用卷积神经网络提取特征,其感受野取决于卷积核的大小,难以获得全局的感受野,从而限制了网络学习更多的特征,导致其分类精度不足。

本文针对现有高光谱影像分类方法其分类精度不足的问题,提出了一种基于空间-频谱Transformer(Spectral-Spatial Transformer, SST)网络的高光谱影像分类方法,通过引入transformer以获得全局感受野,从而提取更多的高光谱影像特征,并分别采用不同形状的卷积核构建查询向量 $Q$ 、键向量 $K$ 和值向量 $V$ ,以更好地挖掘出高光谱影像中的频域信息和空间域信息。首先将高光谱影像预处理为一维特征向量,然后设计了具有光谱-空间注意力模块(Spectral-Spatial Attention Module, SSAM)和池化残差模块(Pooling Residual Block, Pool Res Block)的SST高光谱影像分类网络。其中,光谱-空间注意

力模块用于捕获光谱通道和空间位置之间的相关性,池化残差模块在残差模块之前加入池化层提取光谱通道上的Patch全局特征,并同时采用最大池化和平均池化以提高模型的鲁棒性。

## 1 研究方法

目前,传统的图像分类方法并不能充分利用高光谱影像中的光谱信息和空间信息。一些基于深度学习的方法利用卷积神经网络,通过制作特定的数据集可以增加网络对于高光谱信息的利用率。Transformer最开始应用在自然语言处理领域,通常用于处理文本分类、情感分析、机器翻译、对话分析、阅读理解等任务。视觉转换器(Vision Transformer, ViT)使其在图像处理中也大放异彩,Transformer的核心是注意力机制,其通过生成成对的查询向量 $Q$ 和键值向量 $K$ ,并计算他们之间的内积得到原始向量不同元素之间的相似度,从而产生新的向量。本文结合高光谱影像的特点,设计SSAM和Pool Res Block,提出如图1所示的SST高光谱影像分类网络,该网络主要包括高光谱影像预处理、光谱-空间注意力模块、池化残差模块和损失函数等。高光谱影像预处理的目的是将二维的高光谱影像展开成一维像素点;光谱-空间注意力模块用于捕获光谱通道和空间位置之间的相关性;池化残差模块中同时采用最大池化和平均池化,以鲁棒提取光谱通道上的Patch全局特征;损失函数采用交叉熵函数,以提高网络训练速度。

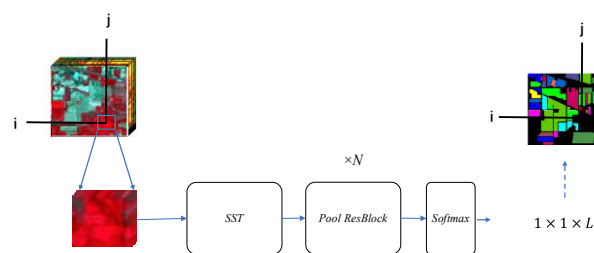


图1 SST高光谱影像分类网络

Fig. 1 The hyperspectral image classification network of SST

### 1.1 预处理

在高光谱影像中,结合每个像素点的邻域信息,能够判断出该像素点的地物类别。一般来说,离像素点越远的区域与像素点地物类别的关系越小,为了减少计算量,取像素点周围 $S \times S$ 大小的区域作为该像素点的邻域,称之为一个Patch。结合该Patch内像素点的光谱信息,综合判断该像素点的地

物类别。具体而言,一张长为 $H \times W$ 的光谱波段数为 $B$ 的高光谱影像,首先将其预处理为 $HW \times (S \times S \times B)$ 的数组,再以单个Patch( $S \times S \times B$ )为单位进行类别预测。

### 1.2 光谱-空间注意力模块

本文采用光谱-空间注意力模块(SSAM)用于捕获光谱通道和空间位置之间的相关性。如图2所示,SSAM将预处理后的特征Patch  $P \in \mathcal{R}^{S \times S \times B}$  作为输入,用不同大小的卷积核生成查询向量 $Q$ 、键向量 $K$ 和值向量 $V$ :

$$\begin{aligned} Q &= P * W_{(1 \times 1 \times 7)}^Q, K = P * W_{(3 \times 3 \times 7)}^K \\ V &= P * W_{(3 \times 3 \times 7)}^V \end{aligned}, \quad (1)$$

其中,\*表示三维卷积操作, $W_{(1 \times 1 \times 7)}^Q$ 表示大小为 $1 \times 1 \times 7$ 的卷积核,其与 $P$ 卷积得到的向量 $Q$ 表示光谱层面的特征图, $W_{(3 \times 3 \times 7)}^K$ 表示大小为 $3 \times 3 \times 7$ 的卷积核,其与 $P$ 卷积得到的向量 $K$ 表示空间层面的特征图。

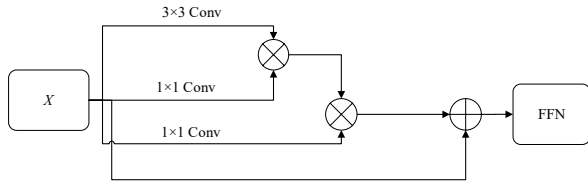


图2 光谱-空间注意力模块具体结构

Fig. 2 The specific structure of the Spectral-Spatial Attention Module

计算 $Q$ 和 $K$ 的内积得到光谱-空间注意力矩阵 $A$ :

$$A = QK^T \quad (2)$$

再将注意力矩阵 $A$ 与向量 $V$ 相乘,生成具有光谱-空间信息的特征Patch  $\hat{P}$ :

$$\hat{P} = AV \quad (3)$$

注意力的主要操作是矩阵乘法,均是线性变换,线性变换的学习能力远低于非线性变换,Transformer结构中需引入前馈网络(Feed-Forward Network, FFN)来实现非线性变换,通常采用卷积和ReLU非线性激活函数搭建FFN。

### 1.3 池化残差模块

池化操作具有一定的特征提取作用,同时不会产生新的参数、不会产生过拟合的现象。本文在残差模块之前加入池化层提取光谱通道上的Patch全局特征。为了提高模型的鲁棒性,本文同时采用了

最大池化和平均池化。首先,分别用最大池化和平均池化将特征图 $\hat{P} \in \mathcal{R}^{S \times S \times B}$ 映射为最大池化特征 $y_{max} \in \mathcal{R}^{1 \times 1 \times B}$ 及平均池化特征 $y_{avg} \in \mathcal{R}^{1 \times 1 \times B}$ 。

$$\begin{aligned} y_{max} &= \text{MaxPool}(\hat{P}) = \max_{(ij) \in (S,S)} \{x_{ij,c}\} y_{avg} = \\ \text{AveragePool}(\hat{P}) &= \frac{1}{S \times S} \sum_{ij=1}^S x_{ij,c} \end{aligned}, \quad (4)$$

其中, $x_{i,j,c}$ 表示特征图的第 $c$ 个通道。为了减少过拟合,将上述两个池化特征作为权重赋予特征图 $\hat{P}$ ,并将两个新的特征图在光谱通上拼接:

$$\begin{aligned} y'_{max} &= \alpha(\text{Reshape}(y_{max})) y'_{avg} = \beta(\text{Reshape}(y_{avg})) y = \\ y'_{max} \odot \hat{P} &+ y'_{avg} \odot \hat{P} + \hat{P} \end{aligned}, \quad (5)$$

其中,Reshape表示将大小为 $\mathcal{R}^{1 \times 1 \times B}$ 的矩阵复制为 $\mathcal{R}^{S \times S \times B}$ 的矩阵, $\alpha \in \mathcal{R}^{1 \times 1 \times B}$ 和 $\beta \in \mathcal{R}^{1 \times 1 \times B}$ 为两个可学习的参数, $\odot$ 表示哈达玛乘积。

残差网络(Residual Network, ResNet)是一种基于CNN提出的能够有效防止梯度消失、并能够解决由于网络过深引起退化问题的神经网络。原始ResNet是由残差块(Resblock)堆叠在一起构建而成,为了使网络对光谱空间特征更具鲁棒性和鉴别能力,本文在池化特征后引入残差网络。

### 1.4 损失函数

损失函数采用常见的交叉熵函数:

$$\text{Loss}^{CE} = -\frac{1}{M} \sum_{m=1}^M y_{L_c}^m \log \hat{y}_{L_c}^m, \quad (6)$$

其中, $y_{L_c}^m$ 表示实际类别标签, $\hat{y}_{L_c}^m$ 表示预测类别标签, $M$ 表示小批量样本总数, $L_c$ 表示土地类别覆盖总数。

### 1.5 评价指标

常用的高光谱影像分类方法评价指标包括总体分类精度(Overall Accuracy, OA)、平均精度(Average Accuracy, AA)和Kappa系数。OA为正确标记的高光谱像素数与测试样本中高光谱像素总数的比值,AA为不同土地覆盖类别的精度均值,Kappa系数为测量分类结果与地面实况之间的一致性。OA、AA和Kappa系数越大,分类效果越好。

设 $M \in R^{n \times n}$ 为分类结果的误差矩阵,其中 $n$ 为土地覆盖类目个数, $(i, j)$ 位置上 $M$ 的值表示第 $i$ 类分类到第 $j$ 类的样本个数,则OA、AA和Kappa系数的定义如下:

$$OA = \frac{\text{sum}(\text{diag}(M))}{\text{sum}(M)}, \quad (7)$$

$$AA = \text{mean}(\text{diag}(M) ./ \text{sum}(M, 2)) \quad (8)$$

$$Kappa = \frac{OA - (\sum(M,1)\sum(M,2))/\sum(M)^2}{1 - (\sum(M,1)\sum(M,2))/\sum(M)^2}, (9)$$

其中,  $diag(M) \in \mathbb{R}^{n \times 1}$  表示  $M$  的对角元素的向量,  $\sum(\cdot) \in \mathbb{R}^1$  表示所有元素的和,  $\sum(\cdot, 1) \in \mathbb{R}^{1 \times n}$  表示每一列元素和的向量,  $\sum(\cdot, 2) \in \mathbb{R}^{n \times 1}$  表示每一行元素和的向量,  $mean(\cdot) \in \mathbb{R}^1$  表示所有元素的平均值,  $./$  表示逐元素除法。

## 2 实验

### 2.1 数据集选择与处理

本文分别采用 Indian Pines(IP)数据集和 Pavia University(UP)数据集对所提高光谱分割方法进行实验。Indian Pines(IP)数据集是由机载可见光/红外成像光谱仪(Airborne Visible Infrared Imaging Spectrometer, AVIRIS)传感器在美国印第安纳州一个单一景观上收集的。该数据集包含 220 个光谱波段,波长范围为 400~2500 nm。高光谱影像的空间维度为 145×145,包含 16 个不同的地表植被类别。UP 数据集是由德国的机载反射光学光谱成像仪(Reflective Optics Spectrographic Imaging System, ROSIS-03)在 2003 年对意大利的帕维亚城所成的像的一部分高光谱数据。该成像光谱仪对 0.43~0.86  $\mu\text{m}$  波长范围内的 115 个光谱波段连续成像,所成图像的空间分辨率为 1.3 m。该数据的空间尺寸为 610×340,因此共包含 2207400 个像素,但是其中包含大量的背景像素,包含地物的像素只有 42776 个,这些像素中共包含 9 类地物,包括树(Trees)、沥青道路(Asphalt)、砖块(Bricks)、牧场(Meadows)等。按照 1:9 的比例随机地将全部像素点分成训练集和测试集,每个像素点和周围 9×9 区域组成一个 Patch。

### 2.2 实验方法

本文实验所采用的 SST 高光谱影像分类网络如图 3 所示,网络由 SSAM 模块和 FFN 组成 SST,并与三个级联的 Pool Res Block 相连,SST 模块网络参数设置如表 1 所示,池化残差模块具体参数如表 2 所示。训练流程为:对高光谱图像的每个像素,取其周围  $S \times S$  个像素作为其邻域 Patch,其经过端到端的网络输出预测结果,计算该预测值(Predict)与真实标签(Label)的交叉熵损失,再利用该损失梯度反向传播更新参数,重复该过程直至网络收敛,并采用 OA、AA 以及 Kappa 系数对网络的性能进行评价。本文实验所采用算法模型基于 Pytorch 框架搭建,采

用 Adam 优化器,使用余弦退火的方式实现学习率的更新,初始学习率为 0.001,每轮训练 200 个 epoch。为了使结果更公正,每组实验重复训练了 5 轮,并报告了这 5 轮的均值和方差作为实验结果。

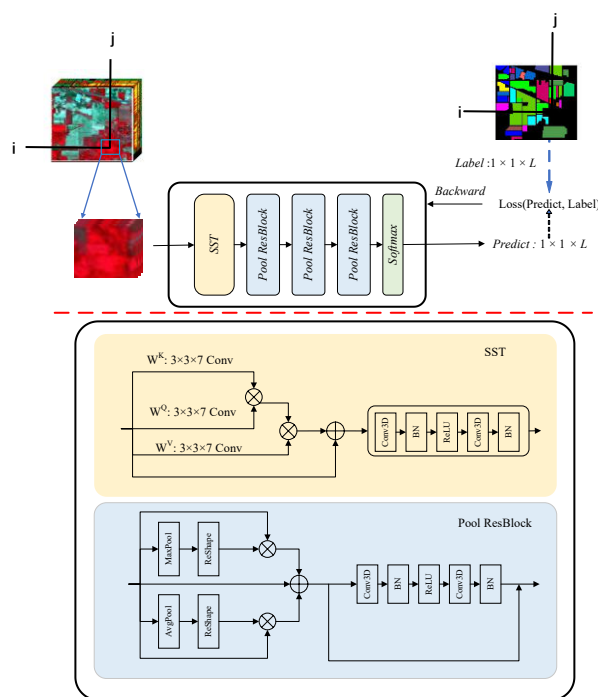


图3 实验采用的 SST 高光谱影像分类网络架构

Fig. 3 The SST hyperspectral image classification network architecture adopted for the experiments

表1 SST 模块网络参数

Table 1 SST module network parameters

Layer	Layer Type
$W^Q$	Conv3d( $i=1, o=24, k=(1 \times 1 \times 7), s=(1, 1, 2), p=(1, 1, 0)$ )
$W^K$	Conv3d( $i=1, o=24, k=(3 \times 3 \times 7), s=(1, 1, 2), p=(1, 1, 0)$ )
$W^V$	Conv3d( $i=1, o=24, k=(3 \times 3 \times 7), s=(1, 1, 2), p=(1, 1, 0)$ )
	Conv3d( $i=24, o=24, k=(1 \times 1 \times 7), s=1, p=(0, 0, 3)$ )
	BatchNorm3d(24)
FFN	ReLU
	Conv3d( $i=24, o=24, k=(1 \times 1 \times 7), s=1, p=(0, 0, 3)$ )
	BatchNorm3d(24)

备注:表中  $i$  代表输入通道数,  $o$  代表输出通道数,  $k$  代表卷积核大小,  $s$  代表步长大小,  $p$  代表边界补充参数大小。

为验证本文所提方法的有效性,分别对两个数据集 The Indian Pines (IP) 和 The University of Pavia (UP) 采用不同方法进行 5 次独立的分类实验,并与经典的机器学习方法和代表性的深度学习方法进行对比,其中包括:多项逻辑回归(MLR)<sup>[18]</sup>、随机森林(RF)<sup>[19]</sup>、支持向量机(SVM)<sup>[20]</sup>、长短期记忆

表2 池化残差网络参数

Table 2 Pooled residual network parameters

Model Type	Module Type	Layer Type	
PoolRes1	Pool	Maxpool	
		Avgpool	
		Conv2d( $i=1, o=1, k=(3 \times 3), s=1, p=(1, 1)$ )	
		Conv2d( $i=1, o=1, k=(3 \times 3), s=1, p=(1, 1)$ )	
		Sigmoid	
		Conv3d( $i=24, o=24, k=(1 \times 1 \times 7), s=1, p=(0, 0, 3)$ )	
	ResBlock1		BatchNorm3d(24)
			ReLU
			Conv3d( $i=24, o=24, k=(1 \times 1 \times 7), s=1, p=(0, 0, 3)$ )
			BatchNorm3d(24)
PoolRes2	Pool	Maxpool	
		Avgpool	
		Conv2d( $i=1, o=1, k=(3 \times 3), s=1, p=(1, 1)$ )	
		Conv2d( $i=1, o=1, k=(3 \times 3), s=1, p=(1, 1)$ )	
		Sigmoid	
		Conv3d( $i=24, o=24, k=(3 \times 3 \times 1), s=1, p=(1, 1, 0)$ )	
	ResBlock2		BatchNorm3d(24)
			ReLU
			Conv3d( $i=24, o=24, k=(3 \times 3 \times 1), s=1, p=(1, 1, 0)$ )
			BatchNorm3d(24)
PoolRes3	Pool	Maxpool	
		Avgpool	
		Conv2d( $i=1, o=1, k=(3 \times 3), s=1, p=(1, 1)$ )	
		Conv2d( $i=1, o=1, k=(3 \times 3), s=1, p=(1, 1)$ )	
		Sigmoid	
		Conv3d( $i=24, o=24, k=(3 \times 3 \times 1), s=1, p=(1, 1, 0)$ )	
	ResBlock3		BatchNorm3d(24)
			ReLU
			Conv3d( $i=24, o=24, k=(3 \times 3 \times 1), s=1, p=(1, 1, 0)$ )
			BatchNorm3d(24)

备注:表中*i*代表输入通道数,*o*代表输出通道数,*k*代表卷积核大小,*s*代表步长大小,*p*代表边界补充参数大小。

(LSTM)<sup>[21]</sup>、ResNet<sup>[22]</sup>、ContextNet<sup>[23]</sup>、MS-3DNet<sup>[24]</sup>、ENL-FCN<sup>[25]</sup>、DPyResNet<sup>[26]</sup>和SSRN<sup>[27]</sup>。所有实验均在Nvidia RTX3080Ti上进行。

### 3 结果和讨论

图4和图5分别为各分类方法在IP数据集和UP数据集上的分类可视化结果。IP数据集各地物

之间距离较小,因此在不同地物分界不容易正确分类,而UP数据集尺寸较大,但其地物类别更少,且彼此之间泾渭分明更好区分。由图4和图5可知,本文所提出的高光谱影像分类分法能够对两类数据集中不同地物进行精确分类。

为了定量地比较各种高光谱分类方法的性能,对于IP数据集和UP数据集,采用不同分类方法分别进行5次独立的分类实验,5次独立实验的OA、AA和Kappa系数性能指标的均值和标准差如表3和表4所示。由表3和表4可知,本文提出的SST高光谱影像分类方法对于IP数据集和UP数据集的OA、AA和Kappa系数这三个性能指标均值均优于其他比较方法,其原因主要包括两个方面。首先,本文提出的SST高光谱影像分类方法针对光谱和空间设计了不同大小的卷积核,利用注意力机制更新两个卷积核所提取的特征,以自适应地提取光谱信息和空间信息。其次,利用池化残差模块有效避免网络训练的过拟合现象,并且残差机制能够防止梯度消失,从而增强SST网络模型的鲁棒性。因此,本文提出的SST高光谱分类方法能够实现高光谱影像的高精度分类。

### 4 结论

本文深入研究了基于深度学习的高光谱影像分类方法,提出了一种基于SST网络的高光谱影像分类方法。在对高光谱影像展开为一维像素点的预处理基础上,提出了SST高光谱影像分类网络。该网络包括SSAM和Pool Res Block两个模块,前者用于捕获光谱通道和空间位置之间的相关性,而后者用于提取光谱通道上的Patch全局特征。实验结果表明,本文所提出方法能够实现高精度的高光谱影像分类,在IP数据集和UP数据集上,其OA性能指标分别为98.67%和99.87%,AA性能指标分别为96.63%和99.83%,Kappa系数分别为0.9849和0.9983。

本网络结构设计的参数量较大,如何对网络进行轻量化改进,如开发减枝压缩算法模型,进而手持设备上实际应用,是未来需要研究的一个重要方向。此外,目前网络在提取特征方面仍采用“黑盒子”的处理方式,如何通过物理属性知道每一个特征对于分类任务的重要程度,其关乎到所设计模型的可解释性,因此,网络可解释性也是一个重要的研究方向。

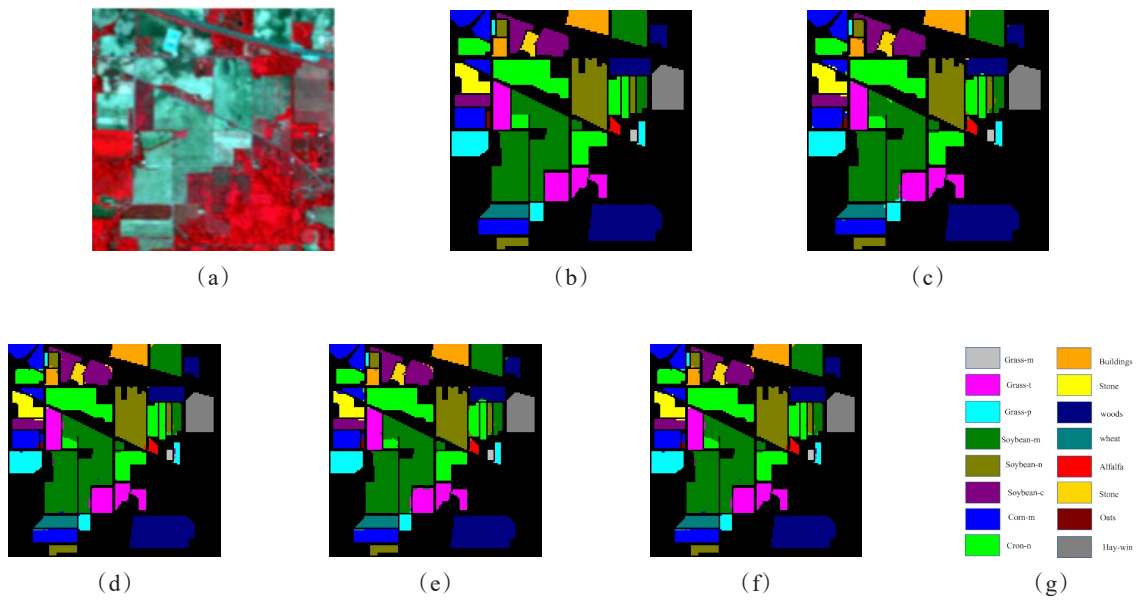


图4 IP数据集的地物分类结果(a)伪彩色合成影像,(b)地表真实数据,(c)DPyResNet方法,(d)SSRN方法,(e)ContextNet方法,(f)本文所提方法,(g)分类图例

Fig. 4 Feature classification results for the IP dataset (a) pseudo-colour composite image, (b) surface real data, (c) DPyResNet method, (d) SSRN method, (e) ContextNet method, (f) method proposed in this paper, (g) classification legend

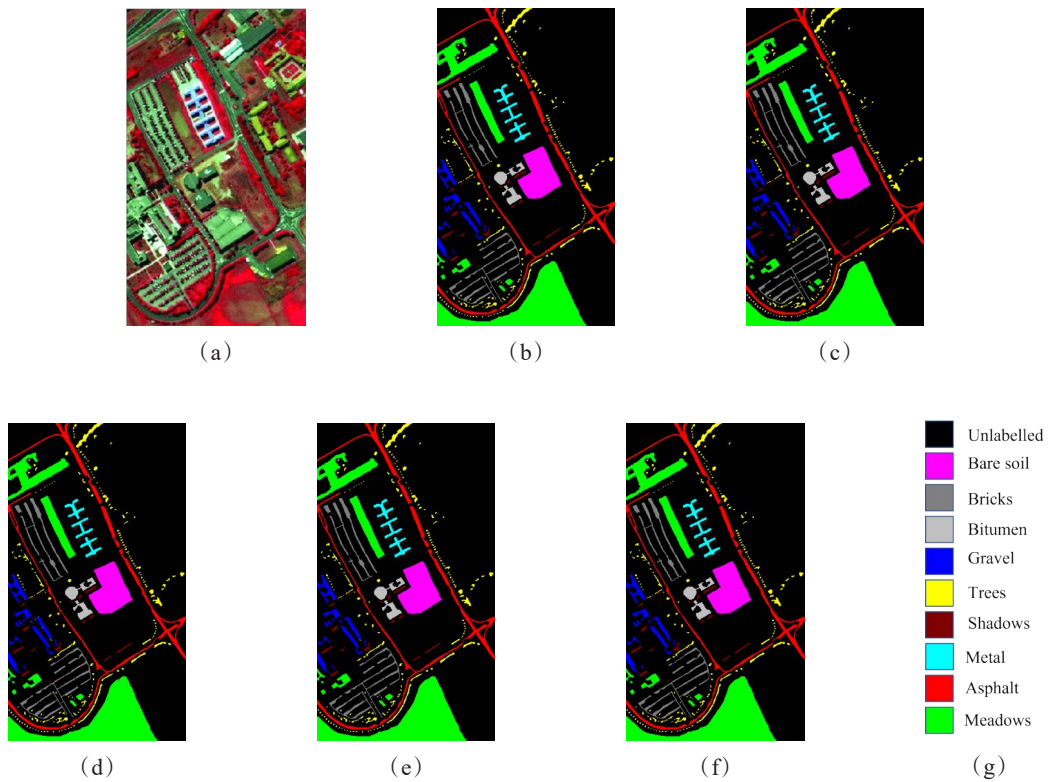


图5 UP数据集的地物分类结果(a)伪彩色合成影像,(b)地表真实数据,(c)DPyResNet方法,(d)SSRN方法,(e)ContextNet方法,(f)本文所提方法,(g)分类图例

Fig. 5 Feature classification results for the UP dataset (a) pseudo-colour composite image, (b) surface real data, (c) DPyResNet method, (d) SSRN method, (e) ContextNet method, (f) method proposed in this paper, (g) classification legend

表3 不同分类方法在IP数据集上的分类性能指标OA(%)、AA(%)和Kappa系数( $\kappa$ )的比较结果Table 3 Comparison results of classification performance metrics OA (%), AA (%) and Kappa coefficient ( $\kappa$ ) of different classification methods on IP dataset

Class	Training	Test	Classical Models				Deep Neural Networks						
			MLR	RF	SVM	LSTM	ResNet	ContextNet	Ms-3DNet	ENL-FCN	DPyResNet	SSRN	SST
1	4	42	15.45±	28.46±	51.22±	69.11±	98.66±	88.78±	66.67±	97.56±	94.69±	57.78±	99.07±
			0.023	0.061	0.190	0.090	0.018	0.080	0.471	0.000	0.076	0.423	0.013
2	142	1286	73.77±	56.63±	81.22±	74.22±	87.85±	98.19±	75.94±	93.15±	93.83±	98.37±	98.73±
			0.006	0.024	0.037	0.016	0.020	0.005	0.080	0.000	0.040	0.012	0.006
3	83	747	51.14±	48.42±	65.82±	71.49±	92.71±	95.37±	81.39±	97.59±	89.30±	97.47±	98.83±
			0.027	0.013	0.013	0.030	0.007	0.028	0.007	0.000	0.003	0.010	0.006
4	23	214	43.97±	33.49±	57.75±	60.72±	95.43±	97.04±	88.63±	91.55±	93.51±	99.12±	98.63±
			0.051	0.025	0.041	0.041	0.046	0.021	0.063	0.000	0.055	0.0099	0.020
5	48	435	83.52±	85.21±	90.04±	87.51±	98.23±	97.78±	95.61±	97.47±	99.26±	97.79±	99.03±
			0.034	0.025	0.014	0.015	0.015	0.015	0.054	0.000	0.004	0.013	0.006
6	73	657	94.82±	92.64±	96.25±	94.77±	97.98±	98.60±	96.78±	99.24±	98.52±	98.50±	98.55±
			0.009	0.027	0.006	0.015	0.011	0.008	0.026	0.000	0.007	0.010	0.009
7	2	26	41.33±	2.67±	73.33±	85.33±	92.98±	90.35±	100.00±	100.00±	83.08±	66.67±	90.70±
			0.186	0.038	0.019	0.094	0.099	0.098	0.000	0.000	0.178	0.471	0.193
8	47	431	98.53±	97.67±	97.98±	97.83±	95.06±	97.76±	89.51±	97.44±	97.63±	96.45±	100.00±
			0.006	0.015	0.006	0.009	0.014	0.026	0.091	0.000	0.022	0.029	0.000
9	2	18	5.56±	9.26±	50.00±	53.70±	60.83±	86.90±	66.67±	72.22±	66.66±	56.25±	78.34±
			0.045	0.094	0.045	0.139	0.283	0.102	0.471	0.000	0.471	0.418	0.091
10	97	875	65.41±	60.91±	73.87±	73.68±	96.05±	96.08±	87.41±	94.74±	93.77±	98.33±	97.49±
			0.041	0.047	0.018	0.025	0.013	0.018	0.070	0.000	0.029	0.009	0.002
11	245	2210	80.37±	87.88±	82.90±	84.93±	93.32±	97.35±	76.69±	95.61±	89.78±	99.08±	99.26±
			0.015	0.019	0.012	0.024	0.041	0.004	0.096	0.000	0.040	0.005	0.002
12	59	534	55.68±	41.26±	74.91±	73.68±	86.65±	94.00±	88.65±	97.00±	83.43±	98.46±	98.03±
			0.007	0.030	0.043	0.025	0.077	0.012	0.036	0.000	0.107	0.009	0.002
13	20	185	97.66±	90.09±	96.94±	98.74±	82.16±	95.01±	99.78±	97.83±	98.19±	100.00±	99.59±
			0.005	0.040	0.021	0.005	0.076	0.03	0.003	0.000	0.021	0.000	0.005
14	126	1139	95.61±	95.46±	93.82±	96.22±	95.39±	98.49±	90.06±	99.12±	96.00±	98.63±	99.38±
			0.004	0.014	0.010	0.004	0.016	0.014	0.087	0.000	0.021	0.010	0.005
15	38	348	56.00±	41.11±	60.42±	60.04±	90.96±	94.10±	88.21±	92.80±	91.22±	99.24±	98.08±
			0.045	0.029	0.044	0.029	0.127	0.031	0.044	0.000	0.040	0.005	0.009
16	9	84	84.92±	79.37±	91.27±	90.87±	94.73±	93.57±	98.53±	100.00±	70.90±	95.63±	96.94±
			0.020	0.030	0.054	0.022	0.038	0.046	0.021	0.000	0.388	0.062	0.023
OA			76.23±	72.98±	82.00±	82.13±	92.44±	96.98±	83.44±	96.15±	91.47±	98.38±	98.67±
			0.008	0.006	0.006	0.004	0.006	0.006	0.060	0.054	0.029	0.004	0.001
AA	1018	9231	65.23±	59.41±	77.36±	79.53±	91.19±	94.96±	86.91±	95.21±	94.14±	91.11±	96.63±
			0.019	0.005	0.019	0.005	0.025	0.003	0.084	0.028	0.006	0.080	0.014
$\kappa$			0.7266±	0.6862±	0.7941±	0.7954±	0.9137±	0.9655±	0.8082±	0.9560±	0.9020±	0.9815±	0.9849±
			0.010	0.007	0.007	0.004	0.006	0.007	0.070	0.030	0.034	0.005	0.002

## References

- [1] Cen Y, Zhang L F, Zhang X, *et al.* Aerial hyperspectral remote sensing classification dataset of Xiongan New Area (Matiwan Village) [J]. Journal of Remote Sensing (Chinese), 2020, **24** (11): 1299–1306. (岑奕, 张立福, 张霞, 等. 雄安新区马蹄湾村航空高光谱遥感影像分类数据集 [J]. 遥感学报). 2020, **24** (11): 1299–1306.
- [2] Zhang F, Zhang J S, Pei C X. A classification algorithm of hyperspectral images based on double channel temporal dense network [J]. Journal of Xian Jiao tong University, 2020, **54** (4): 126–132 (张帆, 张景森, 裴昌幸. 采用双通道时间稠密网络的高光谱影像分类方法 [J]. 西安交通

表4 不同分类方法在UP数据集上的分类性能指标OA(%)、AA(%)和Kappa系数( $\kappa$ )的比较结果Table 4 Comparison results of classification performance metrics OA (%), AA (%) and Kappa coefficient ( $\kappa$ ) of different classification methods on UP dataset

Class	Training	Test	Classical Models				Deep Neural Networks						
			MLR	RF	SVM	LSTM	ResNet	ContextNet	Ms-3DNet	ENL-FCN	DPyResNet	SSRN	SST
1	663	5968	92.30±	91.11±	94.30±	95.47±	96.82±	99.56±	99.36±	99.98±	98.35±	99.85±	99.91±
			0.004	0.007	0.008	0.005	0.023	0.002	0.001	0.000	0.017	0.001	0.001
2	1864	16785	96.18±	98.11±	97.65±	96.91±	98.59±	99.85±	99.80±	100.00±	98.76±	99.98±	99.98±
			0.003	0.003	0.002	0.002	0.008	0.002	0.000	0.000	0.008	0.000	0.000
3	209	1890	72.75±	67.71±	81.26±	78.01±	90.01±	99.19±	98.02±	99.68±	94.22±	99.68±	99.75±
			0.013	0.014	0.018	0.011	0.061	0.001	0.017	0.000	0.034	0.003	0.002
4	306	2758	89.28±	88.20±	94.63±	94.92±	99.32±	99.80±	99.71±	98.94±	99.20±	99.92±	99.97±
			0.002	0.006	0.004	0.007	0.003	0.002	0.001	0.000	0.005	0.001	0.004
5	134	1211	99.42±	98.93±	99.20±	99.26±	99.81±	99.91±	99.94±	100.00±	99.72±	99.94±	99.9±
			0.003	0.002	0.002	0.003	0.000	0.001	0.000	0.000	0.003	0.000	0.003
6	502	4527	77.45±	72.14±	90.58±	87.85±	99.41±	99.75±	99.43±	99.87±	98.52±	99.95±	99.95±
			0.005	0.022	0.008	0.012	0.002	0.003	0.003	0.000	0.006	0.001	0.005
7	133	1197	55.69±	75.68±	85.71±	80.23±	96.90±	98.37±	99.18±	100.00±	97.37±	100.00±	100.00±
			0.043	0.017	0.011	0.007	0.017	0.022	0.005	0.000	0.004	0.000	0.000
8	368	3314	87.04±	89.64±	88.20±	88.49±	92.00±	98.48±	97.13±	99.69±	97.37±	98.28±	99.02±
			0.004	0.013	0.003	0.008	0.044	0.008	0.005	0.000	0.004	0.015	0.005
9	94	853	99.77±	99.77±	99.84±	99.88±	98.88±	99.26±	99.74±	100.00±	84.51±	99.39±	99.94±
			0.001	0.002	0.001	0.001	0.012	0.005	0.002	0.000	0.071	0.003	0.006
OA			89.87±	90.41±	94.19±	93.45±	97.38±	99.57±	99.35±	99.76±	99.60±	99.77±	99.87±
			0.001	0.001	0.002	0.001	0.007	0.001	0.001	0.002	0.001	0.001	0.004
AA	4273	38503	85.54±	86.81±	92.38±	91.23±	96.86±	99.35±	99.15±	99.70±	97.05±	99.66±	99.83±
			0.004	0.002	0.003	0.001	0.005	0.002	0.002	0.002	0.010	0.01	0.005
$\kappa$			0.8646±	0.8710±	0.9229±	0.9130±	0.9652±	0.9943±	0.9913±	0.9972±	96.69±	0.9969±	0.9983±
			0.001	0.002	0.002	0.001	0.009	0.001	0.002	0.001	0.006	0.01	0.006

大学学报), 2020, 54(4): 126-132.

- [3] Gao K L, Yu X C, Zhang P Q, *et al.* Hyperspectral image spatial-spectral classification using capsule network-based method [J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(3): 428-437. (高奎亮, 余旭初, 张鹏强, 等. 利用胶囊网络实现高光谱影像空谱联合分类 [J]. *武汉大学学报(信息科学版)*, 2022, 47(3): 428-437.
- [4] Xu Y H. Research on deep learning and adversarial defense methods for hyperspectral remote sensing image classification [J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(1): 157. (徐永浩. 面向高光谱遥感影像分类的深度学习与对抗防御方法研究 [J]. *武汉大学学报(信息科学版)* 2022, 47(1): 157.
- [5] Xie H, Tong X H. A probability-based improved binary encoding algorithm for classification of hyperspectral images [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, 7(6): 2108-2118.
- [6] Prasad S, Bruce L M. Limitations of principal components analysis for hyperspectral target recognition [J]. *IEEE Geoscience and Remote Sensing Letters*, 5(4): 625-629.
- [7] Villa A, Benediktsson J A, Chanussot J, *et al.* Hyperspectral image classification with independent component discriminant analysis [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, 49(12): 4865-4876.
- [8] Bandos T V, Bruzzone L, Camps-Valls G. Classification of hyperspectral images with regularized linear discriminant analysis [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2009, 47(3): 862-873.
- [9] Blanzieri E, Melgani F. Nearest neighbor classification of remote sensing images with the maximal margin principle [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46(6): 1804-1811.
- [10] Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, 42(8): 1778-1790.
- [11] Gislason P O, Benediktsson J A, Sveinsson J R. Random forests for land cover classification [J]. *Pattern Recognition Letters*, 2006, 27(4): 294-300.
- [12] Pal M. Extreme-learning-machine-based land cover classification [J]. *International Journal of Remote Sensing*, 2009, 30(14): 3835-3841.
- [13] Friedl M A, Brodley C E. Decision tree classification of



- land cover from remotely sensed data [J]. *Remote Sensing of Environment*, 1997, **61**(3): 399–409.
- [14] Chen Y, Jiang H L, Li C Y, *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] Li W, Wu G D, Zhang F, *et al.* Hyperspectral image classification using deep pixel-pair features [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, **55**(2): 844–853.
- [16] Xie J, He N J, Fang L Y, *et al.* Multiscale densely-connected fusion networks for hyperspectral images classification [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, **31**(1): 246–259.
- [17] Zheng X T, Sun H, Lu X Q, *et al.* Rotation-invariant attention network for hyperspectral image classification [J]. *IEEE Transactions on Image Processing*, 2022, **31**: 4251–4265.
- [18] Haut J, Paoletti M, Paz-Gallardo A, *et al.* Cloud implementation of logistic regression for hyperspectral image classification. *17th International Conference on Computational and Mathematical Methods in Science and Engineering (CMMSE)*, Jul. 2017.
- [19] Ham J, Chen Y, Crawford M M, *et al.* Investigation of the random forest framework for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, **43**(3): 492–501.
- [20] Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, **42**(8): 1778–1790.
- [21] Paoletti M E, Haut J M, Plaza J, *et al.* Scalable recurrent neural network for hyperspectral image classification. *Journal of Supercomputing*, 2020, **76**(11): 8866–8882.
- [22] He K, Zhang X Y, Ren S Q, *et al.* Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, Jun.2016.
- [23] Lee H, Kwon H. Going deeper with contextual CNN for hyperspectral image classification. *IEEE Transactions on Image Processing*, 2017, **26**(10): 4843–4855.
- [24] He M Y, Li B, Chen H H. Multi-scale 3D deep convolutional neural network for hyperspectral image classification. *IEEE International Conference on Image Processing (ICIP)*, China, Sep.2017.
- [25] Shen Y, Zhu S J, Chen C, *et al.* Efficient deep learning of nonlocal features for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, **59**(7): 6029–6043.
- [26] Paoletti M E, Haut J M, Fernandez-Beltran R, *et al.* Deep pyramidal residual networks for spectral-spatial hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, **57**(2): 740–754.
- [27] Zhong Z L, Li J, Luo Z M, *et al.* Spectral-spatial residual network for hyperspectral image classification: a 3-D deep learning framework. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, **56**(2): 847–858.