

结合协方差池化与跨尺度特征提取的高光谱分类

汪乐¹, 彭江涛^{1,3}, 陈娜¹, 孙伟伟^{2,3}

1. 湖北大学 数学与统计学学院 应用数学湖北省重点实验室, 武汉 430062;

2. 宁波大学 地理与空间信息技术系, 宁波 315211;

3. 宁波大学 遥感遥测产业技术研究院 宁波拾烨智能科技有限公司, 宁波 315211

摘要: 深度卷积神经网络在高光谱图像分类任务上取得了优越性能。但是, 主流深度学习算法通常采用一阶池化运算, 容易忽略光谱之间的相关性, 因而难以获取高阶统计判别特征。另外, 这类算法往往难以选择最优的窗口大小去捕获不同感受野信息。针对上述问题, 本文提出了一种结合协方差池化和跨尺度特征提取的高光谱影像分类方法。该方法设计了跨尺度自适应特征提取模块, 能够自动提取多尺度特征, 获取不同视野的互补信息, 避免了尺度选择问题; 进一步利用平均池化和快速协方差池化的联合池化操作, 得到一阶统计量和结合空间光谱信息的二阶统计量; 最终, 将一阶和二阶池化特征进行融合用于分类。在3个公开高光谱数据集Indian Pines、Houston和Pavia University上分别随机选取5%、5%和1%标记样本进行训练, 本文算法得到的总体分类精度分别达到97.63%、98.48%和98.21%, 分类性能优于主流深度学习方法。

关键词: 高光谱图像分类, 协方差池化, 多尺度, 特征融合, 卷积神经网络

中图分类号: P407.8/P2

引用格式: 汪乐, 彭江涛, 陈娜, 孙伟伟. 2024. 结合协方差池化与跨尺度特征提取的高光谱分类. 遥感学报, 28(1): 203–218

Wang L, Peng J T, Chen N and Sun W W. 2024. Hyperspectral classification algorithm based on covariance pooling and cross scale feature extraction. National Remote Sensing Bulletin, 28(1): 203–218 [DOI: 10.11834/jrs.20242326]

1 引言

高光谱图像 HSI (Hyperspectral image) 包含丰富的空间信息和光谱信息。相比于传统的 RGB 图像和多光谱图像, 高光谱图像的光谱分辨率更高、光谱波段更多, 因此能更好表征地物的结构和成分, 适用于鉴别地物之间的细微差异。分类是许多高光谱遥感处理与应用的基础任务 (杜培军等, 2016; 童庆禧等, 2016; 彭江涛等, 2020)。分类目的就是将 HSI 中每个像素分类到已知的地物类别中。目前高光谱图像分类已广泛应用于植被研究 (Demarez, 1999)、农业和资源管理 (Murphy 等, 2019) 和军事 (Shimoni 等, 2019) 等多个领域。

早期基于像素的高光谱分类方法更关注光谱信息, 如支持向量机 SVM (Support Vector Machine) (Melgani 和 Bruzzone, 2004)、随机森林 (Ham 等,

2005)、K 均值聚类 (Ma 等, 2010) 等分类。这些方法仅利用光谱特征而未考虑空间信息, 其特征表征能力有限。特征提取的好坏直接影响分类效果 (孙伟伟等, 2018), 联合利用空间和光谱信息进行特征提取与分类是提升高光谱分类性能的有效途径。Benediktsson 等提出一种结合光谱和空间信息的形态学轮廓 EMP (Extended Morphological Profiles) 特征提取与分类方法 (Benediktsson 等, 2005); Mura 等 (2010) 提出了扩展多属性剖面方法 EMAP (Extended Multi-Attribute Profile)。Li 等 (2015) 采用局部二值模式 LBP (Local Binary Patterns) 提取高光谱影像的局部特征。Jia 等 (2015) 提出一种基于三维 Gabor 特征的协同表示高光谱影像分类方法。Camps-Valls 等 (2006) 提出一种结合空间和光谱信息的支持向量机复合核框架 SVM-CK (SVM with Composite Kernel)。Peng 等 (2017) 提

收稿日期: 2022-07-06; 预印本: 2023-01-11

基金项目: 国家重点研发计划 (编号: 2020YFA0714200); 湖北省自然科学基金 (编号: 2021CFA087); 国家自然科学基金 (编号: 42171351, 42122009); 浙江省“尖兵”“领雁”研发攻关计划 (编号: 2023C01027)

第一作者简介: 汪乐, 研究方向为图像处理。E-mail: wangle@stu.hubu.edu.cn

通信作者简介: 孙伟伟, 研究方向为多源遥感协同智能解译与海岸带健康监测。E-mail: sunweiwei@nbu.edu.cn

出了一种理想正则化复合核分类框架,该方法在学习复合核时同时利用了光谱信息、空间信息和标签信息。Chen等提出了联合稀疏表示JSR (Joint Sparse Representation) 高光谱分类方法 (Chen等, 2011; Peng等, 2019)。这些空谱联合特征提取与分类方法能够有效地提取高光谱图像的空间—光谱特征,但是这些人为设定的特征如EMP特征、LBP特征、Gabor特征并不具备数据自适应变化的特性。

近年来,深度学习的发展极大促进了HSI分类的研究。深度学习网络能够根据数据的特点自动学习相应的判别特征,已被广泛应用于HSI分类领域。典型的深度学习框架包括堆叠式自动编码器SAE (Stacked AutoEncoder)、深度信念网络DBN (Deep Belief Networks)、卷积神经网络CNN (Convolutional Neural Networks)等。Chen等 (2014) 首先将深度学习的概念引入高光谱图像分类中,应用SAE构建基于深度架构的模型,以无监督的方式学习深度特征。随后,Tao等 (2015) 提出了一种基于稀疏约束的改进自编码器。Li等 (2014) 基于DBN方法提取深度光谱特征图。然而,这2类深度学习的方法依然存在空间信息缺失的问题。为了充分利用高光谱图像的空间和光谱信息,Zhao和Du (2016) 将CNN用于高光谱图像分类,通过二维卷积神经网络 (2D-CNN) 提取空间特征,利用平衡局部判别嵌入算法提取光谱特征。CNN具有局部连接和权重共享的优点,在提取空间信息的同时能够有效减小参数量。考虑到HSI本质是一个三维张量,(Chen等 (2016) 及Zhang等 (2017) 直接利用三维卷积神经网络 (3D-CNN),从原始三维数据中同时提取光谱和空间特征,取得了较好的分类性能。Zhang等 (2019) 提出了三维DenseNet来处理HSI分类,采用串联的方式进行密集连接,通过各层之间的直接连接,确保各层之间信息的有效传输。鉴于仅使用二维卷积或三维卷积分别存在缺失通道关系信息或模型复杂的缺点,Roy等 (2020) 提出了混合光谱卷积神经网络HybridSN (Hybrid Spectral-Net),利用三维卷积进行联合空间光谱特征表示,二维卷积在三维卷积的基础上进一步学习更抽象的空间表示。网络层数的增加能够提升CNN的特征表达能力。但是,网络加深可能会出现梯度消失或梯度爆炸问题。为此,Zhong等 (2018) 提出了空间光谱残差网络SSRN (Spectral-

Spatial Residual Network),借助残差连接对网络进行深化。Roy等提出了一种基于注意力机制的自适应光谱空间核改进残差网络A2S2K-ResNet (Attention-based Adaptive Spectral-Spatial Kernel ResNet),该网络通过选择性学习三维卷积核,利用改进的三维残差块联合提取光谱空间特征 (Roy等, 2021)。以上网络提取单尺度邻域数据作为输入,只利用单尺度下的空间—光谱特征融合,没有考虑多尺度特征融合提供不同视野丰富的互补信息。Lee和Kwon (2017) 设计了上下文交互的深度卷积网络 (Contextual deep CNN),将多尺度卷积获得的初始空间特征和光谱特征图组合,最后得到联合空间—光谱特征图。类似地,Gong等 (2019) 提出具有多尺度卷积 (Multiscale-CNN) 的新型卷积神经网络,从高光谱图像中提取深层多尺度特征,并结合多种度量提高分类性能。以上方法仅在网络某一层考虑执行多尺度特征提取,可能不足以全面学习不同尺度下的光谱空间信息。充分利用不同层次之间的互补信息也是一种特征融合的思路。Song等 (2018) 提出深度特征融合网络DFFN (Deep Feature Fusion Network),引入残差学习简化网络的训练,并融合不同层次的输出,进一步提高了分类精度。Li等 (2019) 提出带有辅助分类器的分层多尺度卷积神经网络,多尺度CNN同时提取不同尺度下的光谱—空间特征,并采用双向长短时记忆网络来捕获多尺度特征之间的相关性。现在多尺度方法仍存在问题,如空间信息冗余,计算成本较高,多尺度特征利用不足等。

在上述的深度神经网络框架中,一般采用一阶池化运算来提取特征。近年来,高阶统计量在机器视觉领域显示出巨大的潜力,特别是采用二阶池化得到的二阶统计量,可以有效利用不同通道之间的相关性信息,在图像分类领域取得了优越性能。在经典的浅层结构中,Tuzel等 (2006) 提出区域协方差矩阵的特征描述方法;Carreira等 (2015) 提出二阶非中心矩池化方法 O_2P (Second-order pooling)。在深度结构中,Ionescu等 (2015) 提出了Deep O_2P 网络,在深度卷积网络的最后一个卷积层后插入二阶池化层,取得了较好分类结果。虽然二阶池化的研究在深度结构中有一定的进展,但仍存在2个难题:一是稳健协方差估计,二是如何充分利用协方差的黎曼几何结构特性。Li等 (2017a) 提出了矩阵幂归一化协方差池化

MPN-COV (Matrix power normalized covariance) 方法来解决上述2个问题。为使算法能够在GPU上进行有效运算,他们进一步构造了迭代矩阵平方根归一化的快速协方差池化 iSQRTCOV (Iterative matrix square root normalization of covariance pooling) 方法 (Li 等, 2018)。在HSI分类应用上, Lin 等 (2015) 提出双分支卷积神经网络B-CNN (Bilinear CNN), 通过聚类局部相邻像素来构造不同光谱波段之间的协方差矩阵。利用协方差池化, 从不同的卷积层生成多个叠加特征。该矩阵的每个元素代表2个特征映射之间的协方差, 进一步提供更多的互补信息 (He 等, 2018)。Xue 等 (2021) 提出了一种基于注意力的二阶池化网络 A-SPN (Attention based second order pooling network), 计算基于注意力的二阶池化特征, 进行HSI分类。Zheng 等 (2021) 提出了协方差池化的三维二维混合卷积提取光谱空间特征 MCNN-CP (Mixed Convolutions and Covariance Pooling), 通过协方差池化从空谱特征图中提取二阶信息。需要注意, 上述算法只使用二阶池化, 可能会忽略一阶池化所具有的特有信息。此外, 采用固定尺度的特征提取器没有考虑到不同感受野获取的互补信息 (Lee 和 Kwon, 2017; Xie 等, 2021)。

因此, 为充分利用一阶二阶统计量信息和多尺度互补空谱特征, 本文提出了一种结合协方差池化和跨尺度特征提取的高光谱影像分类算法。该方法具有以下几个特点: 首先, 设计了跨尺度自适应特征提取模块 CSAF (Cross Scale Adaptive Feature Extraction), 利用多尺度卷积核, 进行横向跨尺度特征提取, 自适应获取不同尺度在相同深度的互补信息; 其次, 设计了一阶二阶联合池化与融合模块, 其中二阶协方差矩阵可以同时利用空间和光谱信息, 矩阵中每个元素代表1个局部空间领域内的2个不同光谱波段的协方差, 一定程度反映样本光谱波段之间的相关性。并且, 得到的协方差矩阵大小只与光谱波段数量有关, 这使得不同尺度构建的协方差矩阵大小相同, 在训练过程中不会产生维度障碍。

2 研究方法

本文提出了一种结合协方差池化和跨尺度特征提取 CP-CSFE (Covariance Pooling and Cross Scale Feature Extraction) 的高光谱分类模型, 如图1所示。该方法由跨尺度自适应特征提取模块和联合池化特征融合模块2个主要部分组成。

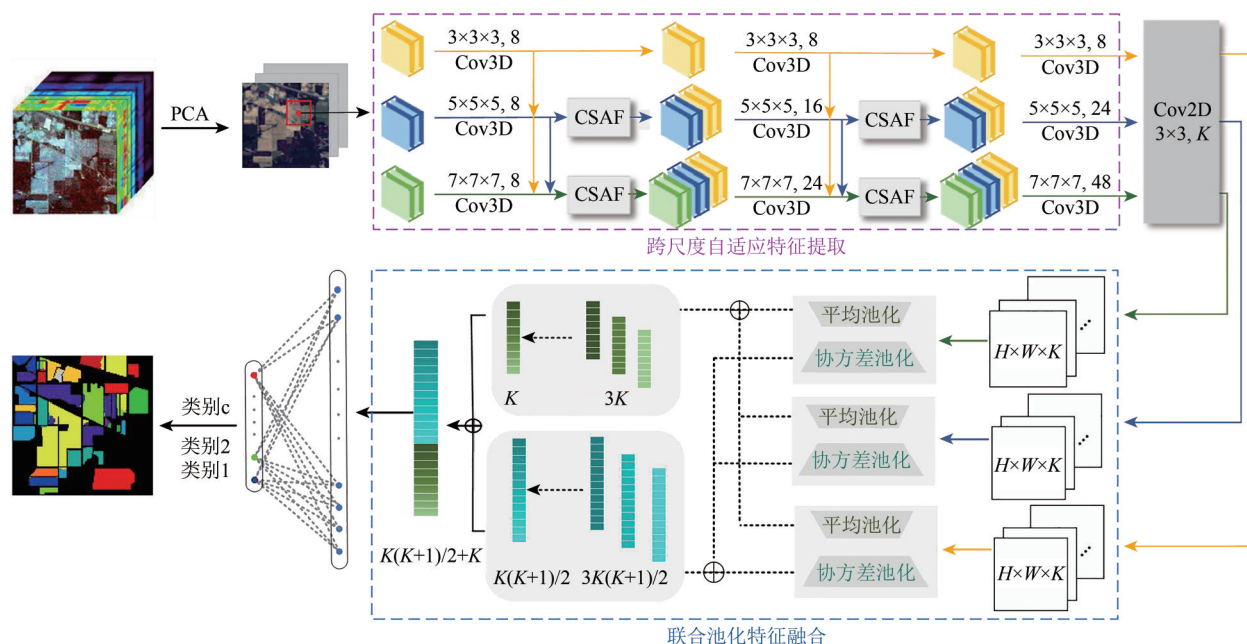


图1 结合协方差池化和跨尺度特征提取的高光谱影像分类算法框架

Fig. 1 The structure of hyperspectral classification algorithm based on covariance pooling and cross scale feature extraction

2.1 跨尺度自适应特征提取

多尺度能够带来不同视野的信息。为了更好地利用不同尺度之间的互补信息，本文设计了CSAF模块，通过跨尺度提取信息获取丰富的互补空谱特征。采用了3种不同的尺度，在第1个三维卷积层中分别设置：8个大小为 $(3 \times 3 \times 3)$ 的三维卷积核；8个大小为 $(5 \times 5 \times 5)$ 的三维卷积核；8个大小为 $(7 \times 7 \times 7)$ 的三维卷积核。剩下2个卷积层与第一个类似，但卷积核数量设置不同。三维卷积能够同时学习高光谱图像的空间和光谱特征，在每次卷积操作之后应用批量归一化（BN）和ReLU激活函数，提高训练收敛速率。

以尺度7为例，CSAF模块的结构如图2所示。输入该模块的张量为 $X_i \in \mathbb{R}^{h \times w \times c \times m}$ 其中， h 表示空间的长度， w 表示空间的宽度， c 表示通道数， m 表示特征图个数。首先从上到下分别将输入通过三维卷积映射得到特征图 $f(X_i) \in \mathbb{R}^{h \times w \times c \times m_1}$ ， $g(X_i) \in \mathbb{R}^{h \times w \times c \times m_2}$ 和 $h(X_i) \in \mathbb{R}^{h \times w \times c \times m_3}$ 。然后，应用自适应全局池化操作将映射特征图转换成体现特征图重要性值的一维向量，3个映射特征图对应得到的结果为 $\text{Adp}(f(X_i)) \in \mathbb{R}^{1 \times 1 \times 1 \times m_1}$ ， $\text{Adp}(g(X_i)) \in \mathbb{R}^{1 \times 1 \times 1 \times m_2}$ 和 $\text{Adp}(h(X_i)) \in \mathbb{R}^{1 \times 1 \times 1 \times m_3}$ 。随后，依次使用大小 $(1 \times 1 \times 1)$ 的卷积核和用于归一化的Sigmoid激活函数，进行信息整合。根据具体的任务情况，不同尺度提供信息的重要程度不同。因此，CSAF模块根据特征映射的重要性分配不同的权重。自适应权重参数 α 计算公式如下：

$$\alpha = \sigma(W_f * \text{Aap}(f(X_i)) + b_f) \quad (1)$$

式中， W_f 和 b_f 分别代表 $(1 \times 1 \times 1)$ 卷积的权重参数和偏置参数， σ 表示Sigmoid激活函数。

类似地，自适应权重参数 β 和 γ 由 $g(X_i)$ 和 $h(X_i)$ 计算得到，具体公式如下：

$$\beta = \sigma(W_g * \text{Aap}(g(X_i)) + b_g) \quad (2)$$

$$\gamma = \sigma(W_h * \text{Aap}(h(X_i)) + b_h) \quad (3)$$

最后将尺度3、尺度5和尺度7本身的特征，结合自适应参数进行串联连接。提取到跨尺度特征，并作为下一个尺度为7的卷积层的输入，形式为

$$\text{Concat} = [\alpha f(X_i), \beta g(X_i), \gamma h(X_i)] \quad (4)$$

值得注意的是，除尺度3外，每个尺度都设置2个CSAF模块。其中尺度5与尺度7不同，只需跨1个尺度，对尺度3和尺度5进行特征提取。

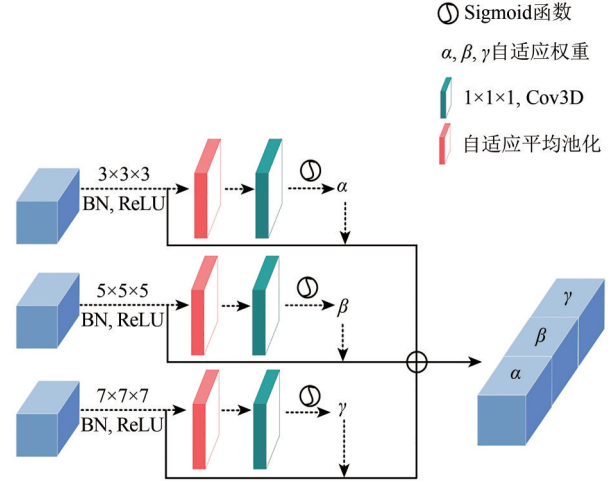


图2 CSAF模块结构

Fig. 2 The structure of CSAF module

2.2 联合池化特征融合

2.2.1 一阶二阶联合池化

经过多尺度自适应特征提取模块，输出不同尺度下得到的四维张量 $(H \times W \times C \times M)$ ，式中， H 表示空间的长， W 表示空间的宽， C 表示通道数， M 表示特征图数量。将第三维和第四维进行连接，可转化为三维张量 $(H \times W \times K)$ ，其中 $K = C \times M$ 。然后，设置一个二维卷积层提取不同特征图的互补信息，并进入一阶二阶联合池化特征融合。

二阶统计量通常比一阶统计量包含更丰富的信息，但直接替换一阶统计量可能会丢失其独有的信息。因此，采用一阶二阶联合池化，通过平均池化（Avg Pooling）获得一阶统计量；通过快速协方差池化（Fast Covariance Pooling）获得二阶统计量。矩阵平方根归一化在计算协方差池化中起关键作用，现有计算矩阵平方根的方法，在很大程度上依赖于特征分解EIG（Eigen decomposition）或奇异值分解SVD（Singular Value Decomposition），这些方法在GPU上的运算效率有待进一步提升。为了实现快速计算，本文采用了一种基于牛顿舒尔茨迭代的快速协方差池化，设计了具有循环嵌入有向图结构的元层来计算协方差矩阵的近似平方根。基于牛顿舒尔茨迭代的快速协方差池化，只涉及GPU友好的矩阵乘法，使网络在GPU上可以高效运行。

如图3所示，输入的特征图是大小为 $(H \times W \times K)$ 的三维张量。第1步将三维张量重塑为特征矩阵 $X \in \mathbb{R}^{K \times N}$ ，其中 $N = H \times W$ ，每一个列向量代表空

间中对应位置的通道信息。第2步利用特征矩阵 X 计算得到协方差矩阵 Σ ，计算公式如下：

$$\Sigma = X \bar{I} X^T \quad (5)$$

式中， $\bar{I} = \frac{1}{N} \left(I - \frac{1}{N} \mathbf{1} \right)$ ， I 和 $\mathbf{1}$ 分别代表 $N \times N$ 单位矩阵和元素都为1的矩阵。协方差矩阵 Σ 是对称矩阵，其每一个元素代表着某2个通道之间的相关性，同时包含空间信息和通道信息。

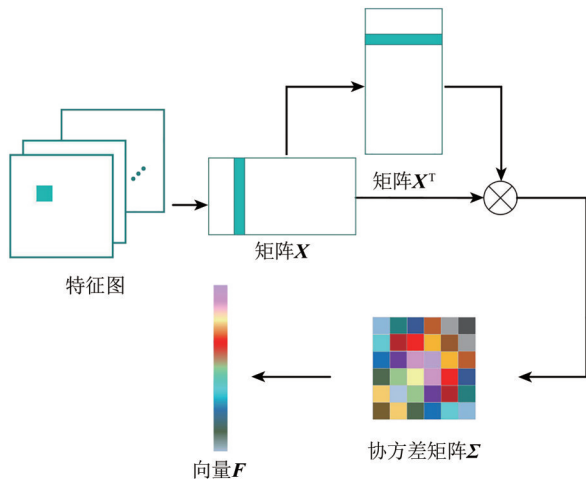


图3 快速协方差池化结构

Fig. 3 The structure of Fast Covariance Pooling

采用具有循环嵌入的有向图结构，计算协方差矩阵 Y 的近似平方根，包含预归一化、牛顿舒尔茨迭代和后补偿3个结构层。

第1层：预归一化。通过除以协方差矩阵的迹或F-范数，保证第2层牛顿舒尔茨迭代的收敛性。具体处理公式如下：

$$A = \frac{1}{\text{tr}(\Sigma)} \Sigma, \text{ 或 } A = \frac{1}{\|\Sigma\|_F} \Sigma \quad (6)$$

设 Σ 的特征值为 λ_i ，则有 $\text{tr}(\Sigma) = \sum_i \lambda_i$ 和 $\|\Sigma\|_F = \sqrt{\sum_i \lambda_i^2}$ 。可以看出 $\|\Sigma - I\|_2$ 等于 $\Sigma - I$ 的最大奇异值，在迹和F-范数的情况下分别为 $\|\Sigma - I\|_2 = 1 - \frac{\lambda_1}{\sum_i \lambda_i}$ ， $\|\Sigma - I\|_2 = 1 - \frac{\lambda_1}{\sqrt{\sum_i \lambda_i^2}}$ 。2种情况值均 < 1 ，满足后续收敛条件。

第2层：牛顿舒尔茨迭代。更新牛顿舒尔茨迭代中涉及的耦合矩阵方程，用于计算近似矩阵平方根。具体来说，为计算 A 的矩阵平方根 Y ，给定 $Y_0 = A$ 和 $Z_0 = I$ ，其中 $K=1, 2, \dots, N$ 。耦合迭代形式如下：

$$Y_k = Y_{k-1} p_{lm}(Z_{k-1} Y_{k-1}) q_{lm}(Z_{k-1} Y_{k-1})^{-1} \quad (7)$$

$$Z_k = p_{lm}(Z_{k-1} Y_{k-1}) q_{lm}(Z_{k-1} Y_{k-1})^{-1} Z_{k-1}$$

式中， p_{lm} 和 q_{lm} 是多项式， l 和 m 是非负整数。式 (7) 满足局部收敛，即：如果 $\|A - I\| < 1$ ，则 Y_k 和 Z_k 收敛到 Y 和 Y^{-1} 。当 $l=0, m=1$ 时，称为牛顿舒尔茨迭代，不涉及GPU不友好的矩阵逆的计算：

$$Y_k = \frac{1}{2} Y_{k-1} (3I - Z_{k-1} Y_{k-1}) \quad (8)$$

$$Z_k = \frac{1}{2} (3I - Z_{k-1} Y_{k-1}) Z_{k-1}$$

第3层：后补偿。由于预归一化处理会降低数据的数量级，因此设计后补偿，通过乘以协方差矩阵的迹或F-范数进行数据还原：

$$C = \sqrt{\text{tr}(\Sigma)} Y_N \text{ 或 } C = \sqrt{\|\Sigma\|_F} Y_N \quad (9)$$

通过上述结构输出的结果是一个对称矩阵，第3步对矩阵的上三角进行向量化，得到最终快速协方差池化的结果向量 F ，维数为 $K(K+1)/2$ 。

2.2.2 池化特征融合

不同尺度得到的池化特征需进行融合得到最终的特征，从而用于分类。如果直接采用串联连接或相加的方式融合，意味着网络对不同池化特征向量的每个维度同等关注。为此突出不同特征之间的差异性，本文设计了池化特征融合模块。

以快速协方差池化为例。首先，拼接3个尺度对应的快速协方差特征向量 $v_{cp} \in R^{K \times 1}$ ，得到特征向量 $V_{cp} \in R^{3K \times 1}$ 。接下来，为减少冗余信息，利用线性函数对特征向量 V_{cp} 降维，得到与融合前维数相同的融合协方差池化向量 $V'_{cp} \in R^{K \times 1}$ 。最后使用归一化指数函数 Softmax 计算每个维度的权重向量 θ ，用以强化特征。类似地，平均池化通过相同的操作得到融合平均池化向量 V'_{avg} 。据此，给出融合向量 E' 的形式：

$$E' = \theta \cdot [V'_{cp}, V'_{avg}] \quad (10)$$

式中， θ 表示池化特征向量对应的权重向量，能够反映每个元素的重要程度。由于池化特征融合模块位于模型的深层，为避免梯度消失问题，采用残差连接，将池化融合特征向量 E' 与原始特征融合。具体形式如下：

$$E = E' + [V'_{cp}, V'_{avg}] \quad (11)$$

最终将得到的融合向量 E 送入全连接层，得到分类结果。

3 结果与分析

3.1 实验数据

(1) Indian Pines (IP) 数据集。利用 AVIRIS 传感器于 1992 年 6 月获取的印第安纳州西北部印第安松树试验场的高光谱分辨率影像。影像大小为 145 像素×145 像素。传感器的空间分辨率为 20 m，

波长范围 400—2500 nm；共包含 220 个波段。剔除 104—108、150—163、220 这 20 个坏波段，余下 200 个波段用于分类。影像共标注地物类别 16 类，主要以森林、植物和农作物为主。影像伪彩色合成图像和类别标记图如图 4 所示。每类按比例随机选取 5% 的样本作为训练集，其余作为测试集，具体信息见表 1，其中训练集和测试集样本总数分别为 512 和 9737。

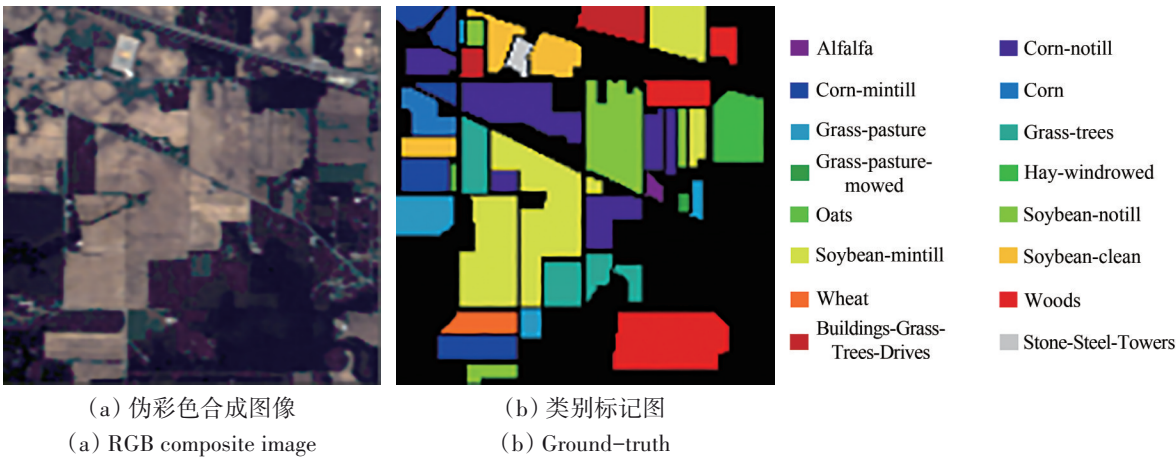


图 4 Indian Pines 高光谱图像数据
Fig. 4 Indian Pines Hyperspectral image

表 1 Indian Pines 数据训练集和测试集				
Table 1 Training and testing sets of Indian Pines data				
编号	地物类别	训练集 样本数	测试集 样本数	类别对应 颜色
1	Alfalfa	2	44	
2	Corn-notill	71	1357	
3	Corn-mintill	41	789	
4	Corn	12	225	
5	Grass-pasture	24	459	
6	Grass-trees	37	693	
7	Grass-pasture-mowed	1	27	
8	Hay-windrowed	24	454	
9	Oats	1	19	
10	Soybean-notill	49	923	
11	Soybean-mintill	123	2332	
12	Soybean-clean	30	563	
13	Wheat	10	195	
14	Woods	63	1202	
15	Buildings-Grass-Trees-Drives	19	367	
16	Stone-Steel-Towers	5	88	
总计		512	9737	

(2) Houston (HU) 数据集。利用 ITRES CASI-1500 传感器于 2013 年获取的休斯顿大学校园和邻近城市地区的高光谱影像。该影像大小为 349×1905 像素，传感器的空间分辨率为 2.5 m，波长范围 380—1050 nm；包含 144 个波段。影像共标注地物类别 15 类，主要以植被、住宅为主。影像伪彩色合成图像和类别标记图如图 5 所示。每类按比例随机选取 5% 的样本作为训练集，其余作为测试集，具体信息见表 2，其中训练集和测试集样本总数分别为 751 和 14278。

(3) Pavia University (PU) 数据集：利用 ROSIS 传感器于 2003 年获取的意大利北部帕维亚大学的高光谱影像。影像大小为 610×340 像素。传感器的空间分辨率为 1.3 m，波长范围 430—860 nm。剔除 12 个受噪声影响的波段，余下 103 个波段用于分类。影像共标注地物类别 9 类，主要以树、沥青为主。影像如图 6 所示。每类按比例随机选取 1% 的样本作为训练集，其余作为测试集，具体信息见表 3，其中训练集和测试集样本总数分别为 427 和 42349。

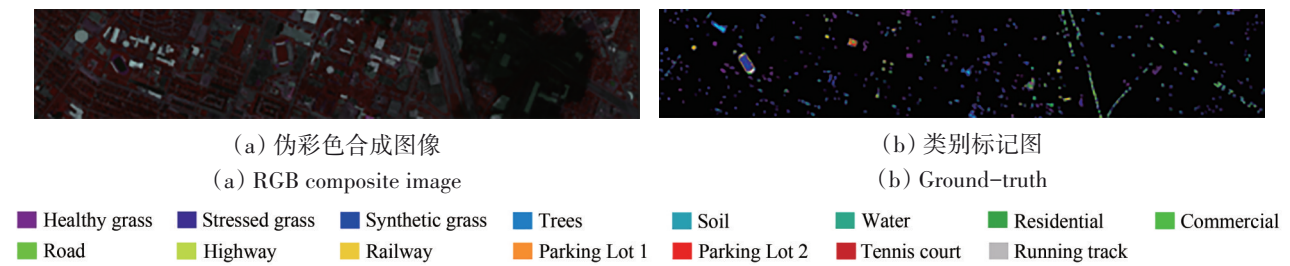


图5 Houston 高光谱图像数据

Fig. 5 Houston Hyperspectral image data

表2 Houston 数据训练集和测试集

Table 2 Training and testing sets of Houston data

编号	地物类别	训练集 样本数	测试集 样本数	类别对应 颜色
1	Healthy grass	63	1188	紫色
2	Stressed grass	63	1191	深蓝色
3	Synthetic grass	35	662	蓝色
4	Trees	62	1182	浅蓝色
5	Soil	62	1180	青色
6	Water	16	309	蓝绿色
7	Residential	63	1205	绿色
8	Commercial	62	1182	黄绿色
9	Road	63	1189	亮绿色
10	Highway	61	1166	黄绿色
11	Railway	62	1173	黄色
12	Parking Lot 1	62	1171	橙黄色
13	Parking Lot 2	23	446	橙红色
14	Tennis court	21	407	红色
15	Running track	33	627	灰色
总计		751	14278	

3.2 对比算法与实验设置

本文所提出的结合协方差池化和跨尺度特征提取的高光谱分类算法与下面采用的主流算法进行对比实验：

(1) SVM (Support Vector Machine)：支持向量机；

(2) 2D-CNN (2D-Convolutional Neural Networks)：二维卷积神经网络分类算法 (Makantasis 等, 2015)；

(3) HybirdSN (Hybrid Spectral-Net)：结合 3D-CNN 和 2D-CNN 提取空间和光谱信息 (Roy 等, 2020)；

(4) SSRN (Spectral-Spatial Residual Network)：光谱—空间残差网络 (Zhong 等, 2018)；

(5) A2S2KResNet (Attention Based Adaptive Spectral Spatial Kernel ResNet)：基于注意力机制和空间光谱核学习的残差网络 (Roy 等, 2020)；

(6) A-SPN (Attention-Based Second-Order Pooling Network)：基于注意力的二阶池化网络 (Xue 等, 2021)；

(7) MCNN-CP (Mixed Convolutions and Covariance Pooling)：协方差池化的混合 CNN 卷积 (Zheng 等, 2021)。

本文算法实验结果为随机选取训练样本重复运行 10 次的平均值。分类性能评价指标为总体分类准确率 OA (Overall Accuracy)，平均分类准确率 AA (Average Accuracy) 和 Kappa 系数 K 。

本文实验基于 Pytorch 深度学习框架，在配备 2.7 GHz Intel Xeon Gold 6258R CPU 和 2 块 NVIDIA GeForce RTX 2080 Ti GPU 的计算机上完成。为减少网络运算开销，首先通过 PCA 将高光谱数据降维至 30 维，然后投入网络。对于本文提出的方法，使用 Adam 优化器，初始学习率为 0.001，迭代次数为 50，批次训练集大小为 16。所有对比算法的参数设置与其原文保持一致。表 4 列出了各算法的具体参数设置，包括窗口大小、批次大小和学习率等。可见：由于网络特性不同，各算法参数有一定差异，其中窗口大小最大为 25 最小为 7；批次大小基本不固定；学习率常设为 0.001。

3.3 实验结果

(1) Indian Pines 数据实验结果。Indian Pines 数据不同算法的分类精度如表 5 所示。不同算法的分类结果如图 7 所示。可知只使用光谱特征的 SVM 算法和 2D-CNN 算法的分类图中有明显的椒盐噪声。分类图左上角 16 和 12 类部分容易出现错分，在此部分本文算法分类结果基本正确。由表 5 可知与其他对比算法相比，本文算法效果最佳，在不同

分类指标上有明显提升。其中，2—4类是大豆子类、10—12类是玉米子类，同一子类的类间相似度高不易区分。本文算法与对比算法相比具有更高的鉴别能力，在第3、4、11、12类中取得最优分类性能，在第2、10类中取得次优分类性能。第7、9类样本数少，其中第9类取得最好分类准确率，第7类取得次优分类准确率，说明本文算法在小样本分

类中也有不错效果。同样采用多尺度思想的MCNN-CP算法的OA、AA和K分别为96.033%、89.627%和95.473%，而本文算法得到的结果可以达到97.630%、95.517%和97.296%，总体分类准确率提高了1.6%。与MCNN-CP相比，本文所提出的跨尺度特征融合，具有较好的特征学习能力。

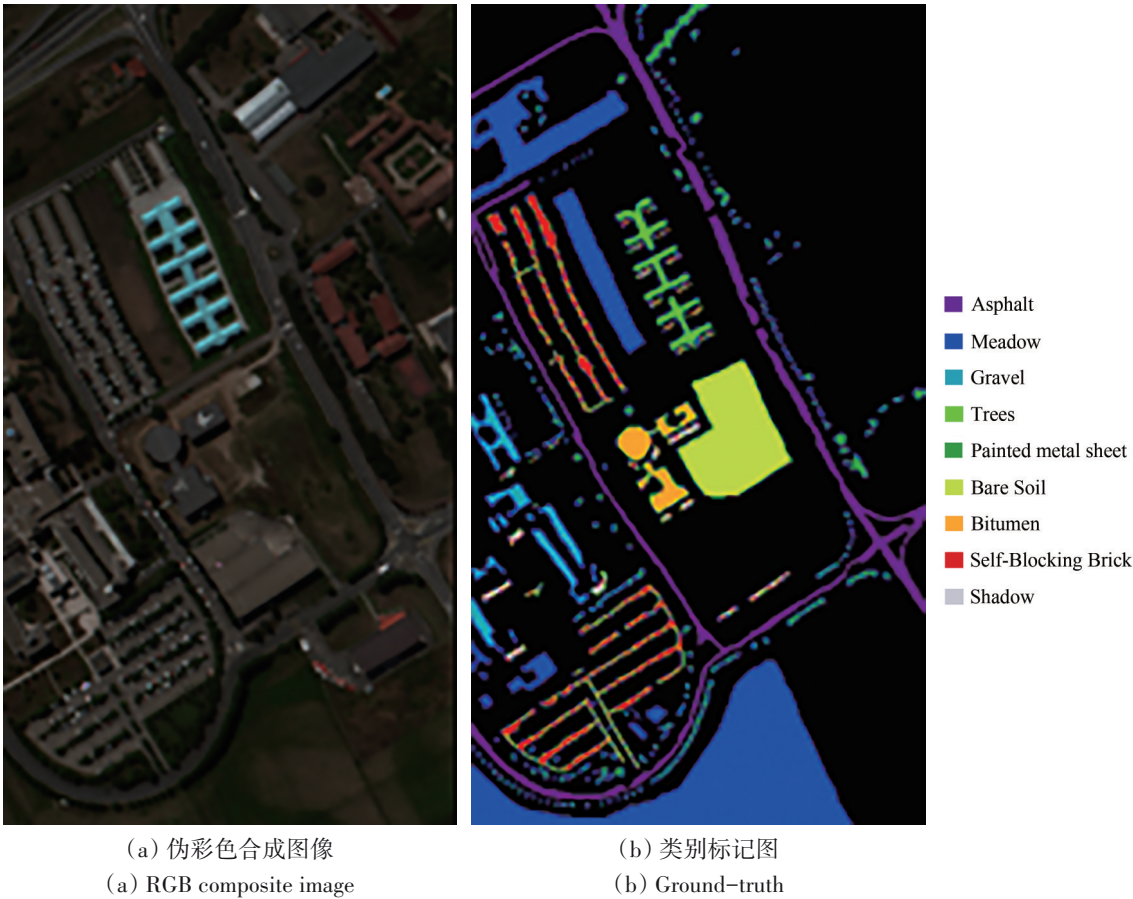


图6 Pavia University 高光谱图像数据
Fig. 6 Pavia University Hyperspectral image data

表3 Pavia University 数据训练集和测试集				
Table 3 Training and testing sets of Pavia University data				
编号	地物类别	训练集 样本数	测试集 样本数	类别对应 颜色
1	Asphalt	66	6565	■
2	Meadow	186	18463	■
3	Gravel	21	2078	■
4	Trees	31	3033	■
5	Painted metal sheet	13	1332	■
6	Bare Soil	50	4979	■
7	Bitumen	13	1317	■
8	Self-Blocking Brick	37	3645	■
9	Shadow	10	937	■
总计		427	42349	

表4 各算法参数设置			
Table 4 Parameter setting of each algorithm			
算法	窗口大小	批次大小	学习率
2D-CNN	11	32	0.0001
HybirdSN	25	128	0.0010
SSRN	7	16	0.0003
A2S2KResNet	11	32	0.0010
A-SPN	9	64	0.0100
MCNN-CP	11	256	0.0010
CP-CSFE	15	16	0.0010

(2) Houston 数据实验结果。Houston 数据集不同算法的分类精度见表6。可见使用5%数据训练，本文算法的OA、AA和K分别达到98.482%，98.473%

和98.359%，该结果比对比算法相比分类效果有所提升。同时，在合成草、土壤、水、贸易区、公路、铁路、停车场1和网球场8个地物类别中分类准确率高于99%。在第6—12类中达到所列算法里的最高分类精度。观察HU数据分类结果图图8，可见本文方法的分类结果更接近地面的真实值。

表5 Indian Pines数据不同算法的分类精度
Table 5 Classification results of different methods on the Indian Pines data

编号	SVM	2D-CNN	HybirdSN	SSRN	A2S2KResNet	A-SPN	MCNN-CP	CP-CSFE
1	12.381	95.588	88.646	94.116	78.181	92.272	88.759	90.979
2	56.261	50.488	93.852	89.296	96.536	76.565	98.234	97.630
3	43.610	68.800	92.616	90.581	96.831	87.842	95.330	97.474
4	31.653	63.158	91.203	93.022	97.511	79.644	99.087	99.285
5	69.204	79.135	90.915	96.036	96.644	89.411	94.615	98.315
6	89.581	86.312	93.931	97.214	99.682	97.084	98.040	98.504
7	37.037	95.459	90.120	53.478	54.814	100	91.523	98.233
8	98.571	91.684	94.928	97.002	100	100	94.741	98.439
9	5.263	99.009	90.081	0	26.315	97.777	98.095	99.047
10	55.832	45.488	90.327	90.443	94.106	82.510	92.667	93.517
11	71.913	68.448	92.511	96.281	97.427	89.562	96.762	97.528
12	43.725	43.551	94.918	96.310	94.280	71.134	96.026	97.016
13	85.938	91.549	95.115	99.160	90.769	99.692	96.132	99.244
14	94.423	93.455	96.912	98.377	99.217	97.254	95.416	99.552
15	55.376	74.468	96.622	96.200	95.720	90.680	96.787	99.564
16	98.182	88.406	92.576	91.709	93.409	88.181	94.614	93.179
总体分类准确率OA	64.421	68.183	93.303	94.108	96.578	87.899	96.033	97.630
平均分类准确率AA	64.156	59.479	86.404	86.201	91.755	89.975	89.627	95.517
Kappa系数K	59.900	63.768	92.343	93.287	96.095	86.156	95.473	97.296

注：粗体为最好结果。

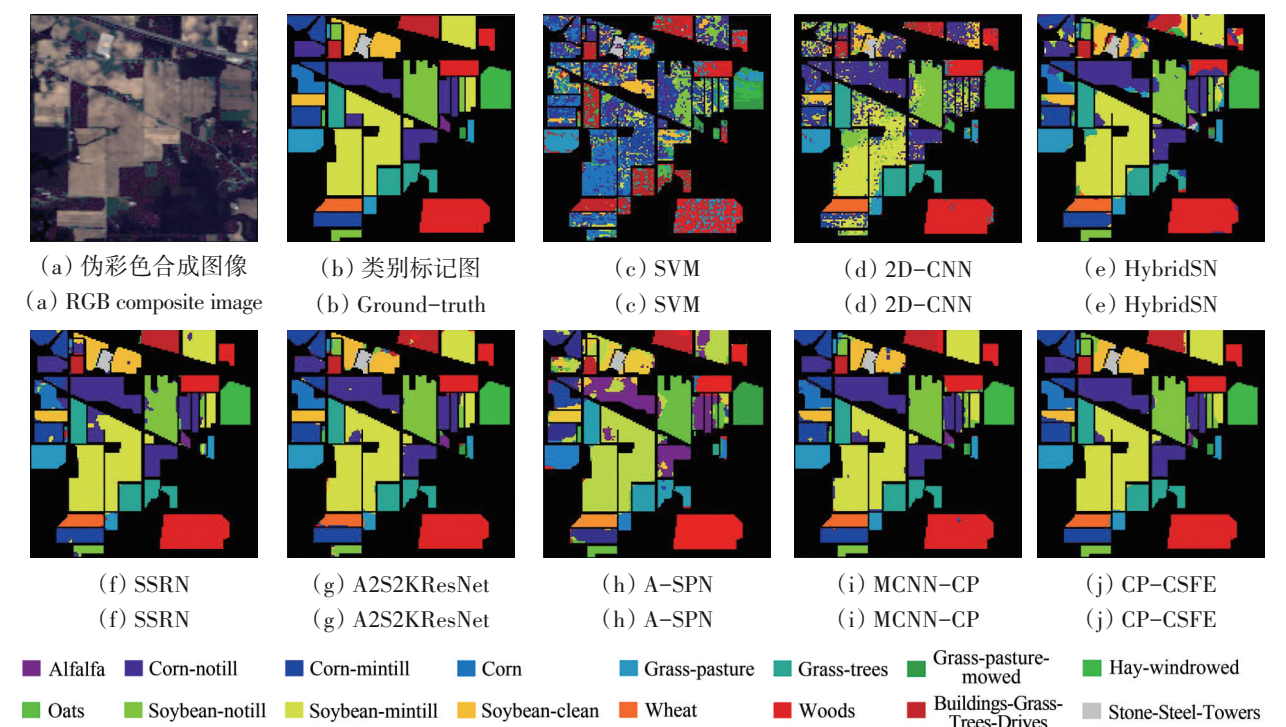


图7 IP数据分类结果(RGB composite image指伪彩色合成图像;Ground-truth指类别标记图)
Fig. 7 IP data classification maps

表6 Houston数据集不同算法的分类精度

Table 6 Classification results of different methods on the Houston data

编号	SVM	2D-CNN	HybirdSN	SSRN	A2S2KResNet	A-SPN	MCNN-CP	CP-CSFE
1	95.667	83.390	97.402	95.278	99.697	97.135	98.598	98.374
2	91.922	80.043	95.649	99.524	98.068	98.654	97.739	98.540
3	99.543	88.131	99.156	99.706	99.849	99.474	99.574	99.729
4	99.281	88.095	94.636	99.828	98.172	99.564	97.503	97.657
5	94.617	76.494	99.254	94.238	99.898	100	98.556	99.268
6	100	98.889	98.548	100	93.980	98.170	97.021	100
7	73.415	64.646	93.621	95.645	97.294	91.674	94.388	97.295
8	82.987	59.784	95.493	95.454	92.605	90.536	97.530	99.326
9	66.355	82.353	92.004	95.981	96.298	95.473	94.199	96.438
10	74.496	73.016	96.934	92.456	98.610	97.909	97.118	99.497
11	70.676	96.350	94.711	95.037	99.215	95.712	97.095	99.266
12	63.592	48.222	96.793	90.881	99.026	97.666	98.091	99.281
13	88.235	11.864	98.539	95.214	93.497	100	94.644	95.111
14	93.160	84.792	99.008	93.036	100	100	95.155	99.713
15	99.515	84.161	98.276	98.272	100	100	99.316	98.650
OA	83.211	60.408	96.114	95.557	97.757	96.939	97.113	98.482
AA	81.895	67.052	96.085	96.037	97.999	97.441	97.357	98.473
K	81.821	57.727	95.798	95.197	97.575	96.688	96.879	98.359

注：粗体为最好结果。

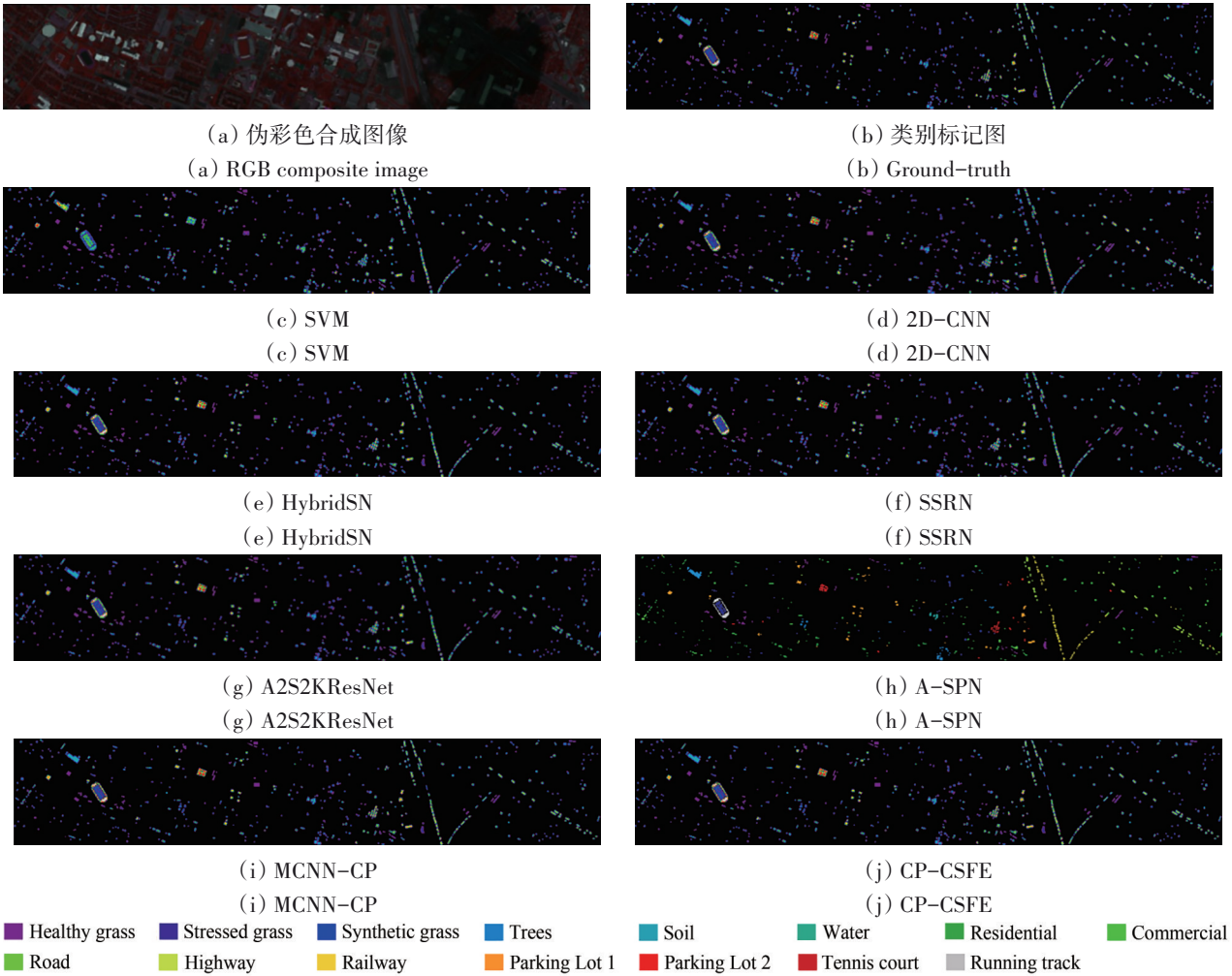


图8 HU数据分类结果

Fig. 8 HU data classification results

(3) Pavia University 数据实验结果。图 9 为不同算法的分类结果图, 从分类图上可以观察到, 只使用光谱信息的 SVM、2D-CNN 算法有明显的椒盐噪声。浅蓝色是第 3 类沙砾, 红色是第 8 类砖块, 这 2 类极易相互混淆错分, 本文算法混淆情况很少。Pavia University 数据集不同算法的分类精度测算结果见表 7。可知: 在使用 1% 标记样本进行

训练时, 本文算法得到的 OA、AA 和 K 结果可以达到 98.211%、96.211% 和 97.626%, 与其他算法相比具有更好的分类性能其中 2 类地物的分类 OA 高于 99%, 3 类地物的分类 OA 高于 97%。7 种方法对第 8 类是砖块的分类精度不高, 错分的情况比较多, 本文提出的方法取得了最好的结果, 与采用二阶统计量的 A-SPN 算法相比, 提高 5% 以上。

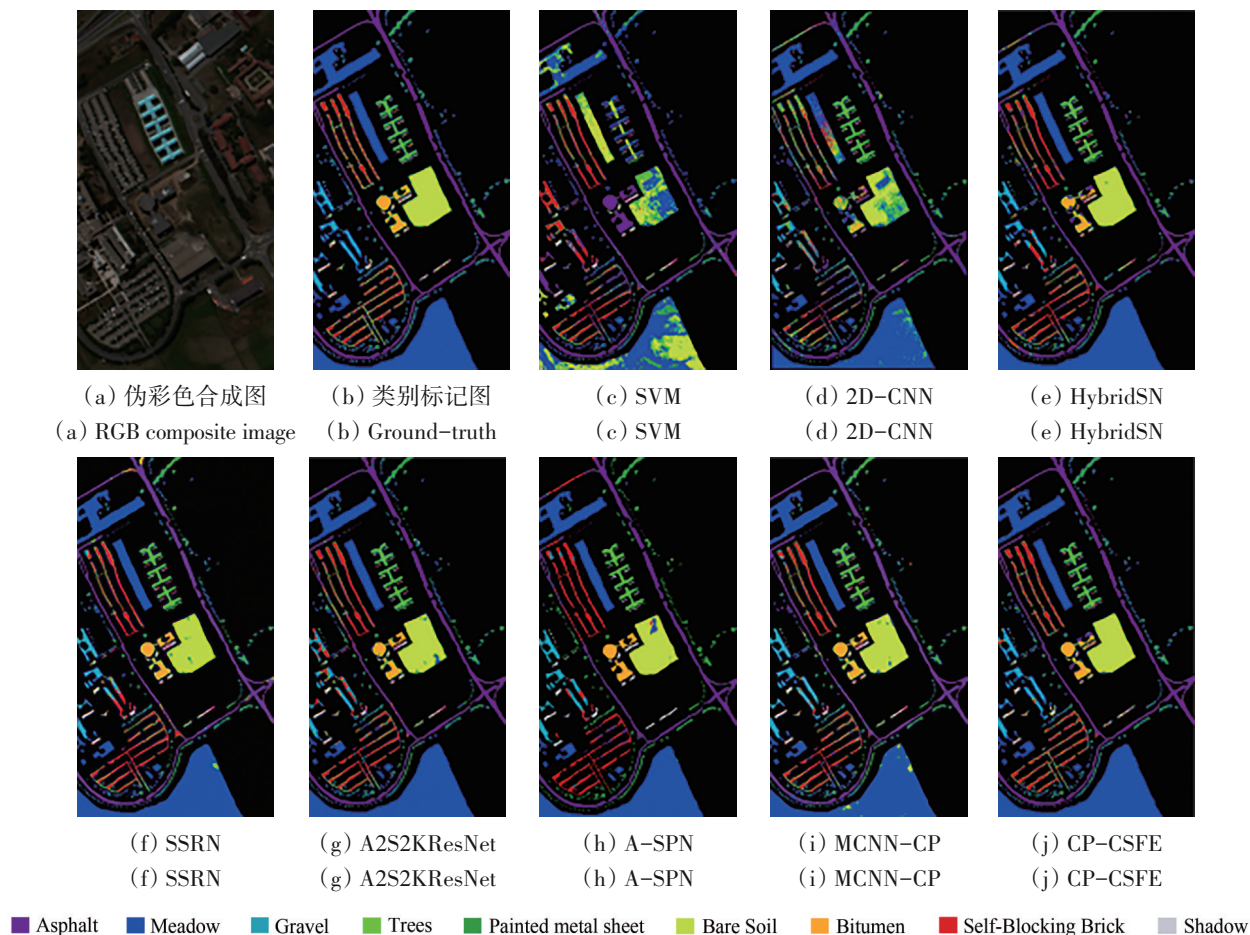


图9 PU数据分类结果

Fig. 9 HU data classification maps

3.4 参数分析及消融实验

3.4.1 参数敏感性分析

采用参数敏感性实验, 分别观察窗口大小、批次大小以及学习率对算法分类精度的影响。为分析窗口大小对 OA 的影响, 设置窗口大小范围为 $\{7 \times 7, 9 \times 9, 11 \times 11, 13 \times 13, 15 \times 15, 17 \times 17, 19 \times 19\}$, 结果见图 10。可见: OA 随窗口增大呈上升趋势, 窗口大小取 15 时效果最好。对于批次大小, 图中结果显示批次大小对 OA 影响不大, 批次大小选定为 16 时效果相对更好, 因此在实验过程中选定批次大小为 16。为分析学习率对算法的影响, 设置

学习率取值范围为 $\{0.0001, 0.001, 0.003, 0.01, 0.03\}$, 在 3 个数据集上 OA 均随学习率增加而减小。当学习率取 0.001 时达到最好分类效果, 因此在实验过程中, 设学习率为 0.001。

3.4.2 快速协方差池化

为验证从光谱特征图中提取的二阶统计量对分类精度提升有效, 将本文算法和只含平均池化的网络和只含快速协方差池化的网络进行比较, 实验结果如表 8。可见: 在 3 个数据集上, 只含平均池化的网络和只含快速协方差池化的网络的总体分类准确率均低于一阶二阶联合池化的算法。

结果表明快速协方差池化所提供的二阶统计量可能提供了更丰富的高阶信息，有助于分类准确率的提升；一阶池化所提供的一阶统计量也能提供特定特征。合理结合一阶二阶统计量能提升分类精度。

表 7 Pavia University数据集不同算法的分类精度
Table 7 Classification results of different methods on the Pavia University data

编号	SVM	2D-CNN	HybirdSN	SSRN	A2S2KResNet	A-SPN	MCNN-CP	CP-CSFE
1	76.618	67.762	91.995	97.638	96.569	93.346	94.315	96.825
2	82.176	92.013	98.133	99.201	99.937	99.868	97.735	99.478
3	65.140	60.384	89.449	96.443	86.188	83.888	89.898	98.770
4	96.750	87.677	94.468	99.706	93.979	93.405	91.028	97.383
5	98.278	99.626	96.825	99.984	99.985	100	99.924	97.938
6	93.407	79.520	98.395	95.290	93.191	95.492	97.760	99.687
7	68.909	88.968	87.232	97.824	91.921	97.752	94.777	93.845
8	66.253	51.846	78.531	86.332	85.574	89.903	92.655	95.482
9	99.787	99.020	80.119	99.241	99.530	93.788	91.134	94.240
OA	81.270	80.424	93.823	97.067	96.027	96.042	95.680	98.211
AA	67.800	70.319	87.921	96.851	95.617	94.160	93.915	96.211
K	73.974	74.009	91.780	96.114	94.692	94.740	94.261	97.626

注：粗体为最好结果。

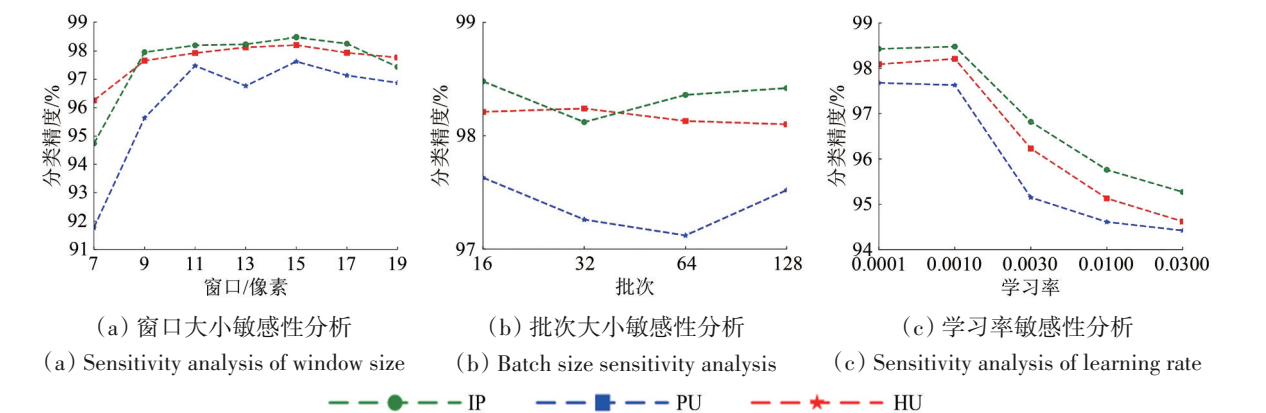


Fig. 10 Parameter sensitivity analysis of window size, batch and learning rate

表 8 有或无快速协方差池化的分类精度
Table 8 Overall accuracy values with/without fast covariance pooling

项目	IP	HU	PU
只含平均池化	96.778	97.839	97.570
只含快速协方差池化	96.789	98.039	97.764
一阶二阶联合池化	97.630	98.482	98.211

注：粗体为最好结果。

3.4.3 多尺度特征提取

为验证多尺度特征提取可以获得更具有鉴别能力的特征，从而提升分类精度，将多尺度特征

提取相关实验结果展示如表9。可见：只采用单尺度时，不存在跨尺度特征提取，联合池化部分只有单尺度下的一阶统计量和二阶统计量，故无需池化特征融合模块，直接将2种池化结果拼接即可；采用2种尺度时，CSAF模块只需使用一次，利用自适应参数进行跨尺度特征提取。后续的池化特征融合模块对2个尺度下的池化结果进行融合；在印第安松树试验场IP、休斯顿大学校园和临近城市地区HU和意大利北部帕维亚大学PU这3个数据集上，跨尺度特征提取的网络总体分类准确率最高。这说明多尺度特征提取模块能有效聚合3种尺度下不同互补信息，产生更具判别性特征。

表9 有或无多尺度特征提取的分类精度

Table 9 Overall accuracy values with/without multiscale feature extraction

%

尺度	IP	HU	PU
尺度3	96.556	97.918	97.72
尺度5	97.326	98.011	97.652
尺度7	97.244	97.851	97.578
尺度3和5	97.148	98.060	97.598
尺度3和7	97.400	98.015	97.786
尺度5和7	97.310	97.870	97.636
尺度3、5和7	97.630	98.482	98.211

注：粗体为最好结果。

表10 样本泛化性实验

Table 10 Sample generalization experiment

数据集	训练样本占比/%	SVM	2D-CNN	HybirdSN	SSRN	A2S2ResNet	A-SPN	MCNN-CP	CP-CSFE
IP	1	46.89	25.12	58.01	80.21	70.04	72.46	76.19	84.68
	5	64.42	68.18	93.30	94.11	96.58	87.90	96.03	97.63
	10	72.24	75.08	98.37	98.11	98.76	92.20	98.54	99.17
HU	1	74.97	31.97	80.96	90.78	88.19	83.26	85.03	91.14
	5	83.21	60.41	96.11	95.56	97.76	96.94	97.11	98.48
	10	85.12	72.12	99.11	98.81	99.31	98.83	99.04	99.33
PU	0.5	78.03	72.18	83.91	95.78	93.22	92.43	90.36	95.97
	1	81.27	80.42	93.82	97.07	96.03	96.04	95.68	98.21
	3	85.57	90.00	97.04	99.29	99.14	98.61	98.57	99.38

注：粗体为最好结果。

3.4.5 算法运算效率对比

为比较算法运算效率，各算法的网络参数量、每轮次训练时间、测试时间的结果的对比如表11所示。可见多尺度会带来更多参数，相比于其他算法计算时间更长。其中：MCNN-CP和CP-CSFE网络采用多尺度特征提取，相比于SSRN、A-SPN、

3.4.4 样本泛化性实验

针对3个数据集，观察不同的训练样本量对算法分类结果的影响，测试不同算法的泛化能力，结果见表10。可见：随着训练样本数量的增加，算法整体精度提高，本文算法在小样本情况下也取得较好的分类效果，达到最高分类精度；本文算法在IP数据集上，用1%数据做训练样本OA到达84.68%，使用10%数据做训练样本，则可以到达99.17%；PU数据空间分辨率高，仅使用0.5%的样本数据总体分类精度也能达到95.97%；对于HU数据集，本文算法的分类效果在1%、5%、10%的训练样本比例上都优于其他算法。

A2S2KResNet算法具有更多的参数量；从训练和测试时间来看，本文CP-CSFE算法训练时间比HybirdSN和A2S2KResNet稍慢，但明显优于A-SPN和MCNN-CP等算法。综合来看，本文算法在保证分类性能较优的情况下，并未显著增加计算时间，整体训练时长仍在可接受范围内。

表11 各算法运算效率对比

Table 11 Comparison of Calculation efficiency of each algorithm

数据集	算法名称	参数	训练 时长/s	测试 时长/s	数据集	算法名称	参数	训练 时长/s	测试 时长/s	数据集	算法名称	参数	训练 时长/s	测试 时长/s
IP	SVM		10.31	0.21	HU	SVM		59.06	0.32	PU	SVM		14.38	0.45
	2D-CNN	2.51M	89.29	4.34		2D-CNN	2.51M	4.86	6.26		2D-CNN	2.50M	20.07	17.26
	HybirdSN	4.97M	0.26	1.73		HybirdSN	4.97M	0.37	2.45		HybirdSN	4.96M	0.22	7.19
	SSRN	0.33M	7.79	38.65		SSRN	0.25M	1.87	80.10		SSRN	0.19M	0.94	99.20
	A2S2KResNet	0.10M	0.40	57.90		A2S2KResNet	0.10M	0.40	76.41		A2S2KResNet	0.10M	1.80	47.20
	A-SPN	0.61M	15.52	9.39		A-SPN	0.30M	17.13	222.33		A-SPN	0.09M	8.89	49.84
	MCNN-CP	1.58M	7.67	13.95		MCNN-CP	1.49M	9.46	17.80		MCNN-CP	1.31M	3.45	42.66
	CP-CSFE	2.58M	2.12	13.56		CP-CSFE	2.57M	3.11	19.77		CP-CSFE	2.56M	1.77	58.53

4 结 论

为了充分提取空间和光谱信息,以获得更好的分类性能,本文中提出一种用于HSI分类的结合协方差池化和跨尺度特征提取的算法,主要结论如下:(1)该模型跨越3个尺度进行空谱信息的选择和提取,调节不同尺度映射特征图的重要性,得到更具适应性互补特征;(2)引入快速协方差池化,提出联合池化和池化特征融合策略学习到更具判别性特征提高了分类性能;(3)本文提出的算法分类结果优于本文采用的其他7种主流对比算法。

通过实验发现,多尺度特征融合在提升分类性能的同时带来更多的网络参数,导致训练时间增加。深入研究应结合高光谱图像的数据特性,进一步优化多尺度组合方式,并且应考虑如何减少网络参数构建轻量化网络,在保证分类性能的同时,提高计算效率。

参考文献(References)

- Benediktsson J A, Palmason J A and Sveinsson J R. 2005. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3): 480-491 [DOI: 10.1109/TGRS.2004.842478]
- Camps-Valls G, Gomez-Chova L, Munoz-Mari J, Vila-Frances J and Calpe-Maravilla J. 2006. Composite kernels for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 3(1): 93-97 [DOI: 10.1109/LGRS.2005.857031]
- Carreira J, Caseiro R, Batista J and Sminchisescu C. 2015. Free-form region description with second-order pooling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(6): 1177-1189 [DOI: 10.1109/TPAMI.2014.2361137]
- Chen Y, Nasrabadi N M and Tran T D. 2011. Hyperspectral image classification using dictionary-based sparse representation. *IEEE Transactions on Geoscience and Remote Sensing*, 49(10): 3973-3985 [DOI: 10.1109/TGRS.2011.2129595]
- Chen Y S, Jiang H L, Li C Y, Jia X P and Ghamisi P. 2016. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10): 6232-6251 [DOI: 10.1109/TGRS.2016.2584107]
- Chen Y S, Lin Z H, Zhao X, Wang G and Gu Y F. 2014. Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6): 2094-2107 [DOI: 10.1109/JSTARS.2014.2329330]
- Demarez V. 1999. Seasonal variation of leaf chlorophyll content of a temperate forest. Inversion of the PROSPECT model. *International Journal of Remote Sensing*, 20(5): 879-894 [DOI: 10.1080/014311699212975]
- Du P J, Xia J S, Xue Z H, Tan K, Su H J and Bao R. 2016. Review of hyperspectral remote sensing image classification. *Journal of Remote Sensing*, 20(2): 236-256 (杜培军, 夏俊士, 薛朝辉, 谭琨, 苏红军, 鲍蕊. 2016. 高光谱遥感影像分类研究进展. *遥感学报*, 20(2): 236-256) [DOI: 10.11834/jrs.20165022]
- Gong Z Q, Zhong P, Yu Y, Hu W D and Li S T. 2019. A CNN with multiscale convolution and diversified metric for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(6): 3599-3618 [DOI: 10.1109/TGRS.2018.2886022]
- Ham J, Chen Y C, Crawford M M and Ghosh J. 2005. Investigation of the random forest framework for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3): 492-501 [DOI: 10.1109/TGRS.2004.842481]
- He N J, Fang L Y, Li S T, Plaza A and Plaza J. 2018. Remote sensing scene classification using multilayer stacked covariance pooling. *IEEE Transactions on Geoscience and Remote Sensing*, 56(12): 6899-6910 [DOI: 10.1109/TGRS.2018.2845668]
- He N J, Paoletti M E, Haut J M, Fang L Y, Li S T, Plaza A and Plaza J. 2019. Feature extraction with multiscale covariance maps for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2): 755-769 [DOI: 10.1109/TGRS.2018.2860464]
- Ionescu C, Vantzos O and Sminchisescu C. 2015. Matrix backpropagation for deep networks with structured layers//2015 IEEE International Conference on Computer Vision. Santiago: IEEE: 2965-2973 [DOI: 10.1109/ICCV.2015.339]
- Jia S, Shen L L and Li Q Q. 2015. Gabor feature-based collaborative representation for hyperspectral imagery classification. *IEEE Transactions on Geoscience and Remote Sensing*, 53(2): 1118-1129 [DOI: 10.1109/TGRS.2014.2334608]
- Lee H and Kwon H. 2017. Going deeper with contextual CNN for hyperspectral image classification. *IEEE Transactions on Image Processing*, 26(10): 4843-4855 [DOI: 10.1109/TIP.2017.2725580]
- Li P H, Xie J T, Wang Q L and Gao Z L. 2018. Towards faster training of global covariance pooling networks by iterative matrix square root normalization//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 947-955 [DOI: 10.1109/CVPR.2018.00105]
- Li P H, Xie J T, Wang Q L and Zuo W M. 2017a. Is second-order information helpful for large-scale visual recognition?//2017 IEEE International Conference on Computer Vision. Venice: IEEE: 2089-2097 [DOI: 10.1109/ICCV.2017.228]
- Li S M, Zhu X Y and Bao J. 2019. Hierarchical multi-scale convolutional neural networks for hyperspectral image classification. *Sensors*, 19(7): 1714 [DOI: 10.3390/s19071714]
- Li T, Zhang J P and Zhang Y. 2014. Classification of hyperspectral im-

- age based on deep belief networks//2014 IEEE International Conference on Image Processing. Paris: IEEE: 5132-5136 [DOI: 10.1109/ICIP.2014.7026039]
- Li W, Chen C, Su H J and Du Q. 2015. Local binary patterns and extreme learning machine for hyperspectral imagery classification. *IEEE Transactions on Geoscience and Remote Sensing*, 53(7): 3681-3693 [DOI: 10.1109/TGRS.2014.2381602]
- Li Y, Zhang H K and Shen Q. 2017b. Spectral - spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sensing*, 9(1): 67 [DOI: 10.3390/rs9010067]
- Lin T Y, RoyChowdhury A and Maji S. 2015. Bilinear CNN models for fine-grained visual recognition//2015 IEEE International Conference on Computer Vision. Santiago: IEEE: 1449-1457 [DOI: 10.1109/ICCV.2015.170]
- Ma L, Crawford M M and Tian J W. 2010. Local manifold learning based K -nearest-neighbor for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 48(11): 4099-4109 [DOI: 10.1109/TGRS.2010.2055876]
- Makantasis K, Karantzalos K, Doulamis A and Doulamis N. 2015. Deep supervised learning for hyperspectral data classification through convolutional neural networks. *IEEE International Geoscience and Remote Sensing Symposium*. Milan: IEEE: 4959-4962 [DOI: 10.1109/IGARSS.2015.7326945]
- Melgani F and Bruzzone L. 2004. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8): 1778-1790 [DOI: 10.1109/TGRS.2004.831865]
- Mura M D, Benediktsson J A, Waske B and Bruzzone L. 2010. Extended profiles with morphological attribute filters for the analysis of hyperspectral data. *International Journal of Remote Sensing*, 31(22): 5975-5991 [DOI: 10.1080/01431161.2010.512425]
- Murphy R J, Whelan B, Chlingaryan A and Sukkarieh S. 2019. Quantifying leaf-scale variations in water absorption in lettuce from hyperspectral imagery: a laboratory study with implications for measuring leaf water content in the context of precision agriculture. *Precision Agriculture*, 20(4): 767-787 [DOI: 10.1007/s11119-018-9610-5]
- Peng J T, Chen H, Zhou Y C and Li L Q. 2017. Ideal regularized composite kernel for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(4): 1563-1574 [DOI: 10.1109/JSTARS.2016.2621416]
- Peng J T, Sun W W and Du Q. 2019. Self-paced joint sparse representation for the classification of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2): 1183-1194 [DOI: 10.1109/TGRS.2018.2865102]
- Peng J T, Sun W W, Wei T H and Fan W Q. 2020. A modified correlation alignment algorithm for the domain adaptation of GF-5 hyperspectral image. *Journal of Remote Sensing (Chinese)*, 24(4): 417-426 (彭江涛, 孙伟伟, 魏天慧, 范文琦. 2020. 高分五号高光谱影像的关联对齐域适应与分类. *遥感学报*, 24(4): 417-426) [DOI: 10.11834/jrs.20209212]
- Roy S K, Krishna G, Dubey S R and Chaudhuri B B. 2020. HybridSN: exploring 3-D - 2-D CNN feature hierarchy for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 17(2): 277-281 [DOI: 10.1109/LGRS.2019.2918719]
- Roy S K, Manna S, Song T C and Bruzzone L. 2021. Attention-based adaptive spectral-spatial kernel ResNet for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(9): 7831-7843 [DOI: 10.1109/TGRS.2020.3043267]
- Shimoni M, Haelterman R and Perneel C. 2019. Hypersectral imaging for military and security applications: combining myriad processing and sensing techniques. *IEEE Geoscience and Remote Sensing Magazine*, 7(2): 101-117 [DOI: 10.1109/MGRS.2019.2902525]
- Song W W, Li S T, Fang L Y and Lu T. 2018. Hyperspectral image classification with deep feature fusion network. *IEEE Transactions on Geoscience and Remote Sensing*, 56(6): 3173-3184 [DOI: 10.1109/TGRS.2018.2794326]
- Sun W W, Zhang D F, Yang G and Li W Y. 2018. Band selection for hyperspectral imagery based on weighted probabilistic archetypal analysis. *Journal of Remote Sensing*, 22(1): 110-118 (孙伟伟, 张殿发, 杨刚, 李巍岳. 2018. 加权概率原型分析的高光谱影像波段选择. *遥感学报*, 22(1): 110-118) [DOI: 10.11834/jrs.20186446]
- Tao C, Pan H B, Li Y S and Zou Z R. 2015. Unsupervised spectral-spatial feature learning with stacked sparse autoencoder for hyperspectral imagery classification. *IEEE Geoscience and Remote Sensing Letters*, 12(12): 2438-2442 [DOI: 10.1109/LGRS.2015.2482520]
- Tong Q X, Zhang B and Zhang L F. 2016. Current progress of hyperspectral remote sensing in China. *Journal of Remote Sensing*, 20(5): 689-707 (童庆禧, 张兵, 张立福. 2016. 中国高光谱遥感的前沿进展. *遥感学报*, 20(5): 689-707) [DOI: 10.11834/jrs.20166264]
- Tuzel O, Porikli F, Meer P. 2006. Region covariance: A fast descriptor for detection and classification. *ECCV European Conference on Computer Vision*. 3952: 589-600. [DOI: 10.1007/11744047_45]
- Xie J, He N J, Fang L Y and Ghamisi P. 2021. Multiscale Densely-Connected Fusion Networks for Hyperspectral Images Classification. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(1): 246-259 [DOI: 10.1109/TCSVT.2020.2975566]
- Xue Z H, Zhang M X, Liu Y F and Du P J. 2021. Attention-based second-order pooling network for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(11): 9600-9615 [DOI: 10.1109/TGRS.2020.3048128]
- Zhang C J, Li G D, Du S H, Tan W Z and Gao F. 2019. Three-dimensional densely connected convolutional network for hyperspectral remote sensing image classification. *Journal of Applied Remote Sensing*, 13(1): 016519 [DOI: 10.1117/1.JRS.13.016519]
- Zhang H K, Li Y, Zhang Y Z and Shen Q. 2017. Spectral-spatial classification of hyperspectral imagery using a dual-channel convolutional neural network. *Remote Sensing Letters*, 8(5): 438-447 [DOI: 10.1080/2150704X.2017.1280200]
- Zhao W Z and Du S H. 2016. Spectral-spatial feature extraction for

- hyperspectral image classification: a dimension reduction and deep learning approach. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8): 4544-4554 [DOI: 10.1109/TGRS.2016.2543748]
- Zheng J W, Feng Y C, Bai C and Zhang J L. 2021. Hyperspectral image classification using mixed convolutions and covariance pooling. *IEEE Transactions on Geoscience and Remote Sensing*, 59(1): 522-534 [DOI: 10.1109/TGRS.2020.2995575]
- Zhong Z L, Li J, Luo Z M and Chapman M. 2018. Spectral-spatial residual network for hyperspectral image classification: a 3-D deep learning framework. *IEEE Transactions on Geoscience and Remote Sensing*, 56(2): 847-858 [DOI: 10.1109/TGRS.2017.2755542]

Hyperspectral classification algorithm based on covariance pooling and cross scale feature extraction

WANG Le¹, PENG Jiangtao^{1,3}, CHEN Na¹, SUN Weiwei^{2,3}

1.Hubei Key Laboratory of Applied Mathematics, Faculty of Mathematics and Statistics, Hubei University, Wuhan 430062, China;

2.Department of Geography and Spatial Information Techniques, Ningbo University, Ningbo 315211, China;

3.Industrial Research Institute of Remote Sensing, Ningbo University, Shiye Intelligent Technology Co., Ltd., Ningbo 315211, China

Abstract: The deep convolution neural network algorithm has achieved excellent performance in hyperspectral image classification. However, these deep learning algorithms generally use first-order pooling operation, which ignores the correlation between different spectral bands. Thus, obtaining high-order statistical discriminant features is difficult. In addition, using these algorithms to choose the optimal window size and capture different receptive field information is complicated. This paper proposes a hyperspectral classification method combining covariance pooling and cross-scale feature extraction to solve the aforementioned problems. This method aims to automatically extract the complementary and discriminative information of different scales and exploit the first- and second-order pooling features to improve the classification performance.

A covariance pooling and cross-scale feature extraction method is proposed for hyperspectral image classification. In this method, a cross-scale adaptive feature extraction module is designed. This module can automatically combine multiscale feature information and obtain complementary information of different visual fields, avoiding the scale selection problem. Furthermore, the first- and second-order statistics combined with spatial-spectral information are obtained using the joint pooling operation of average and fast covariance pooling. Finally, the first- and second-order pooled features are fused for classification.

A total of 5%, 5%, and 1% labeled samples were randomly selected from three public hyperspectral datasets, namely, Indian pines, Houston University, and Pavia University, respectively. The overall classification accuracy of the proposed algorithm reached 97.63%, 98.48%, and 98.21%, and the classification performance was better than the state-of-the-art deep learning methods.

Cross-scale feature extraction considers the complementary spatial-spectral information between different scales to obtain additional adaptive feature information. Combining fast covariance and average pooling, the discriminant features are obtained by pooling feature fusion to obtain superior classification results.

Key words: hyperspectral image classification, covariance pooling, multiscale, feature fusion, convolution neural networks

Supported by National Key Research and Development Program of China (No. 2020YFA0714200); Hubei Province Natural Science Foundation of China (No. 2021CFA087); National Natural Science Foundation of China (No. 42171351, 42122009); Zhejiang Province “Pioneering Soldier” and “Leading Goose” Research and Development Project (No. 2023C01027)