



计算机科学与探索

Journal of Frontiers of Computer Science and Technology

ISSN 1673-9418, CN 11-5602/TP

《计算机科学与探索》网络首发论文

题目：混合深度 CNN 联合注意力的高光谱图像分类
作者：王燕，吕艳萍
网络首发日期：2021-07-29
引用格式：王燕，吕艳萍. 混合深度 CNN 联合注意力的高光谱图像分类. 计算机科学与探索. <https://kns.cnki.net/kcms/detail/11.5602.TP.20210728.1358.002.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

混合深度 CNN 联合注意力的高光谱图像分类

王燕⁺, 吕艳萍

兰州理工大学 计算机与通信学院, 兰州 730050

+ 通信作者 E-mail: wangyan@lut.cn

摘要:深度学习中的卷积神经网络(convolutional neural networks, CNN)能充分利用计算机的计算能力,高效的提取遥感图像的特征,取得很好的成果,特别是在高光谱图像分类方面取得了很大的进展。为了在有限的高光谱样本上充分提取光谱和空间特征,提高高光谱图像分类的精度,提出了混合深度卷积联合注意力(hybrid deep CNN-attention, HDC-Attention)的模型。具体地,首先利用核主成分分析(kernel principal component analysis, KPCA)和小批量K均值(minibatch k-means, MBK-means)对高光谱图像进行组合降维,有效地消除数据冗余并保留主要信息量,使得降维后的数据具有最佳区分度。然后将降维后的数据输入HDC网络进行充分地光谱-空间特征提取。最后利用光谱-空间注意力(Spectral-Spatial Attention),重新分配光谱-空间特征的权重,增强有用的空谱特征,抑制无用的特征。提出的模型在三个公开数据集上进行了多次实验,在有限的标记样本下,三个数据集的OA, AA, Kappa分类指标均超过99%以上。

关键词:高光谱图像分类; KPCA; 卷积神经网络; 光谱-空间注意力机制; 深度学习

文献标志码: A **中图分类号:** TP751.1

Hybrid deep CNN joint attention for hyperspectral image classification

WANG Yan⁺, LYU Yanping

School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China

Abstract: Convolutional neural networks (CNNs) in deep learning can make full use of the computing power of computers to efficiently extract the features of remote sensing images. This has achieved good results, especially in the classification of hyperspectral images. In order to fully extract spectral and spatial features from limited hyperspectral samples and improve the accuracy of hyperspectral image classification, a hybrid deep CNN joint attention (HDC-Attention) model is proposed. Firstly, use kernel principal component analysis (KPCA) and mini batch K-means (MBK-means) to reduce the dimensionality of hyperspectral images. This effectively eliminates data redundancy and retains the main amount of information. The result is that the dimensionality-reduced data has the best discrimination. Then use the HDC network to extract the spectral-spatial features from the dimensionality-reduced data. Finally, the Spectral-Spatial Attention is used to redistribute the weights of spectral-spatial features. It enhances useful spatial-spectral features and suppresses useless features. The proposed model has been tested on three public data sets for many times. With a limited number of labeled samples, the OA, AA, and Kappa classification indicators of the three data sets all exceed 99%.

Key words: hyperspectral image classification; KPCA; convolutional neural network; spectral-spatial attention mechanism; deep learning

基金项目: 国家自然科学基金(61863025); 甘肃省重点研发计划-工业类(18YF1GA060)。

This work was supported by National Natural Science Foundation of China (61863025) and Key R & D Program of Gansu Province-Industrial (18YF1GA060).

高光谱图像 (hyperspectral image, HSI) 是由高光谱遥感传感器或者成像光谱仪获得的, 与传统的 RGB 图像相比, 高光谱图像包含二维空间信息, 在第三维中包含丰富的光谱信息^[1], 大量的光谱波段可以区分各种具有高细节的材料。此外, 利用图像的形状、纹理和几何结构等空间特征, 可以提高识别率, 非常有助于分类和识别任务。因此高光谱数据被广泛的应用到矿产资源开发^[2]、农业试验^[3]、异常检测^[4]和食品分析^[5]等领域。

尽管 HSI 中具有丰富的光谱信息, 但是也带来了新的问题。随着高光谱数据的不断增加, 导致了新的实际和理论问题。在实际问题方面, 高维数据与有限训练样本之间的不平衡, 导致分类结果随维数的增加而下降的现象^[6]。人工标记 HSI 的成本较高而导致的标签样本短缺问题。HSI 的空间布局比较复杂, 不同的光谱材料分类难度更大。理论方面, 在高维空间中, 为高光谱图像开发的传统算法可能不再适用。

传统的高光谱图像分类方法大多基于光谱信息, 例如主成分分析 (principal component analysis, PCA)^[7]、线性判别式分析 (linear discriminant analysis, LDA)^[8]和流形学习 (manifold learning)^[9]都是基于光谱信息的分类方法, 但是只依赖光谱特征进行分类会导致“椒盐现象”的发生, 高光谱图像中的像素并不是相互独立的, 忽略了相邻像素间存在的空间相关性。因此, 为进一步提高分类性能, 将空间特征纳入分类方法中, 例如基于手工制作的 Gabor 滤波^[10]、形态学属性轮廓^[11]、三维离散小波^[12]等方法被相继提出。虽然手工制作的空間特征有助于改善椒盐现象并在一定程度上提高了分类精确度。但是, 手工特征通常仅具备某种特定类型的空间结构信息, 无法呈现高光谱图像低级空间属性的多样性。之后提出了光谱-空间特征融合的方法, 如基于对象或像素分割、决策融合、特征融合的方法。在特征融合方法中, 基于深度学习的算法在 HSI 分类中应用广泛。与之前的分类模型相比, 深度学习最显著的特点是在端到端的分层框架中学习高

级特征。

典型的深度学习方法有自动编码器 (auto encoder, AE)^[13]、深度信念网络 (deep belief network, DBN)^[14]和卷积神经网络 (convolutional neural networks, CNN)^[15]。在这些深度网络模型中, CNN 在 HSI 的特征提取和分类中表现出更好的性能。YAO 等提出了聚类算法和 2D CNN 的深度学习模型^[16]。WEI 等使用 2D CNN 提取 HSI 的深层空间特征, 同时利用 1D CNN 提取深层光谱特征, 通过两个网络模型全连接层的连接完成空间特征和光谱特征融合^[17], 但是基于 1D 或 2D 的深度学习网络无法同时利用三维数据中的光谱和空间信息。它们大多是分别获取空间和光谱这两种不同类型的信息, 然后通过某种融合方法或简单堆叠的方式获取空谱联合特征。所以上述方法并没有考虑和利用不同频带之间的关联性。

在此基础上, 基于 3D CNN 的各种分类方法取得了不错的分类结果^[18]。3D CNN 可以联合提取光谱和空间特征以保持 HSI 立方体的相关性质, 其中光谱和空间信息被独立并联合地包含在 3D 结构中。ZHANG 等提出了一种深而宽的 3D CNN 模型 SSDANet^[6], 虽然取得了很高的分类精度, 但模型复杂度很大, 在 Indian Pines 数据集上的训练时间达到 25814.89s。ZHANG 等提出一种多通道网络^[19], 利用 3D 密集连接网络提取空谱特征, 密集连接性使模型更深入, 但是忽略了网络的宽度, 这将导致细节特征随着模型的加深而逐渐丢失。

近几年来, 注意力机制被引入 CNN 中。LU 等提出了一种基于三维通道和空间注意力的多通道光谱-空间残差网络 (CSMS-SSRN)^[20], 其中的三维通道和空间注意力机制更多的关注空间特征而忽略了光谱特征。SUN 等提出了一种空间注意力机制网络 (SSAN)^[21], 将光谱-空间网络 (SSN) 与 SSAN 相结合提取空谱特征。FANG 等引入光谱注意力机制来增强光谱特征的可区分性, 从而提高了训练模型的性能^[22]。上述基于注意力机制的方法仅利用光谱特征或空间特征, 忽略了 HSI 的特

殊结构。

尽管上述模型在 HSI 分类上取得了一定的结果，但是提取有效且具有辨识性的空谱特征仍然是一个相当大的挑战。由于 HSI 标签样本数量短缺，不同标记样本的不平衡也降低了 HSI 分类的准确性，同物异谱和异物同谱现象的普遍存在也增加了分类的难度。

针对上述问题提出了混合深度卷积联合注意力 (hybrid deep CNN-attention, HDC-Attention) 网络模型。具体地，该方法首先利用核主成分分析 (kernel principal component analysis, KPCA) [23] 和小批量 K 均值 (mini batch k-means, MBK-means) 来压缩 HSI 中的光谱特征，然后将处理后的数据输入 HDC 网络进行光谱-空间提取，再利用光谱-空间注意力 (Spectral-Spatial Attention) 进一步增强关键的光谱-空间特征。最后利用 Softmax 进行分类任务。由于该模型可最大限度的提取光谱和空间特征，所以在受限的样本下表现出较好的分类性能。

1 相关工作

1.1 高光谱图像的降维处理

高光谱图像具有高维度和高特征冗余的特点，维度的降低可以缓解维数灾难问题，在计算资源有限的情况下，也可以避免内存不足的问题。

利用 PCA 对高光谱图像进行降维，是将数据集投影在一个空间中，沿着每个正交向量的方差最大化，没有考虑图像的空间结构和特征，而且 PCA 适用于数据的线性降维，而高光谱数据由于其成像过程中如电磁波和大气的复杂作用、临近地物反射的电磁波的干扰、光谱仪的空间分辨率的限制等因素是非线性的。KPCA 在继承线性 PCA 优点的基础上通过引用核函数增强了高光谱数据的非线性信号处理能力，从而检索更高阶的统计信息，可以更好的处理遥感图像。

KPCA 是利用核方法对 PCA 的一种非线性扩展，其基本思想是：对于输入矩阵 X ，通过一个非线性

映射函数将 X 映射到高维甚至是无穷维的特征空间 F ，然后在特征空间 F 中进行 PCA 处理。

给定一组高光谱数据集 $X = \{X_1, X_2, \dots, X_N\}$ ， X 的每一列表示一个样本，其中 $\{X_i \in R^K, i = 1, \dots, N\}$ 。设一个非线性映射 ϕ 将 X 映射到特征空间 F (D 维)：

$$\phi(X): R^K \rightarrow R^D \quad (1)$$

将矩阵 X 映射到 F 中得到一个 $D \times K$ 的新矩阵 $\phi(X) = \{\phi(X_1), \phi(X_2), \dots, \phi(X_N)\}$ 。然后在 F 中对 $\phi(X)$ 进行降维。 F 的协方差矩阵为：

$$C_F = \frac{1}{N} [\phi(X_1), \dots, \phi(X_N)] \begin{bmatrix} \phi(X_1)^T \\ \dots \\ \phi(X_N)^T \end{bmatrix} \quad (2)$$

令 $X^T = [\phi(X_1), \dots, \phi(X_N)]$ ，得到

$$C_F = \frac{1}{N} X^T X \quad (3)$$

由于 $\phi(X)$ 可能是超高阶或者无限阶。将对角化 C_F 的操作转化为求解核函数 k 的特征问题：

$$k = XX^T = \begin{bmatrix} k(X_1, X_1) & \dots & k(X_1, X_N) \\ \vdots & \ddots & \vdots \\ k(X_N, X_1) & \dots & k(X_N, X_N) \end{bmatrix} \quad (4)$$

求解 k 的特征值和特征向量：

$$(XX^T)P = \lambda P \quad (5)$$

其中， P 为矩阵 k 的特征向量 (特征空间中的权重向量)， λ 为矩阵 k 的特征值。

将 (5) 式左右同时乘 X^T ，得到：

$$(X^T X)(X^T P) = \lambda (X^T P) \quad (6)$$

由于 $N \cdot C_F = X^T X$ 矩阵 k 和 C_F 的特征值都为 λ ，所以 C_F 的特征向量为 $X^T P$ 。将特征向量归一化：

$$V = \frac{1}{\|X^T P\|} X^T P = \frac{1}{\sqrt{\lambda}} X^T P \quad (7)$$

在上式中， λ 和 P 可以通过矩阵 k 求解， X^T 仍未知，但求解 X 在 V 上的投影即可解出，即

$$V^T \phi(X_i) = \frac{1}{\sqrt{\lambda}} P^T \begin{bmatrix} k(X_1, X_i) \\ \vdots \\ k(X_N, X_i) \end{bmatrix} \quad (8)$$

常用的核函数如多项式核(Polynomial kernel, Poly 核):

$$k(x, y) = (ax^T y + c)^d \quad (9)$$

径向基函数核(Radial Basis Function kernel, RBF 核):

$$k(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) = \exp(-\gamma \|x - y\|^2) \quad (10)$$

线性核(Linear kernel):

$$k(x, y) = x^T y + c \quad (11)$$

使用 KPCA 进行降维操作时,核函数的选择和参数设置是影响结果的重要部分。

MBK-means 聚类算法是 K means 算法的改进,其思想是随机从整体特征中做一个抽样,选取出一小部分数据来代替整体。XIU 等证明在进行大数据量的遥感影像分类时 MBK-means 算法比 K-means 算法更具优势^[24]。通过抽样不仅可以提高迭代的效率,正确率也可以达到标准。

KPCA 在保持高光谱数据的空间特征结构的同时,压缩光谱信息并减少了光谱冗余,MBK-means 通过抽样的方法将经过 KPCA 降维后的特征进行进一步的降维,同时也减少了运算时间。

1.2 3D CNN 和 2D CNN

基于 CNN 的深度学习模型之所以能够有如此优秀的性能,是因为不同卷积层所学习的图像特征之间有着非常清晰的层次结构。最前面卷积层的神经元主要被颜色激活,说明光谱特征是特征学习过程的基础。中间层学习的特征对应于图像中的角、边缘和纹理,而最后一层特征则是针对类别的,主要对应于对象和部分的对象。

实际上,自然图像和遥感图像的特征本质上是按层次组织的。特征学习是光谱-空间的学习过程,和 3D CNN 学习光谱-空间特征的过程相吻合,因此,以顺序方式处理光谱特征和空间特征是合理的。

光谱特征是高层次空间特征的基础,空间特征用于分类决策。3D CNN 在三个维度上同时工作,其中两个维度为空间维,剩下的一维为光谱维,高光谱图像也是三维的立方体,3D CNN 联合提取的光谱和空间特征被独立并联合地包含在 3D 结构中。所以 3D-CNN 更加适合提取高光谱图像的特征。

3D CNN 的卷积层利用三维卷积核对三维输入进行运算。 $v_{l,j}^{x,y,z}$ 表示为第 l 层第 j 个特征映射 (x, y, z) 位置的神经元,其计算公式如(12)所示。

$$v_{l,j}^{x,y,z} = f\left(\sum_m \sum_{h=0}^{H_l-1} \sum_{w=0}^{W_l-1} \sum_{r=0}^{R_l-1} k_{l,j,m}^{h,w,r} v_{(l-1),m}^{(x+h),(y+w),(z+r)} + b_{l,j}\right) \quad (12)$$

其中, m 表示第 $(l-1)$ 层特征图的索引,卷积核的高度、宽度和光谱维数分别用 H_l 、 W_l 和 R_l 表示。 $k_{l,j,m}^{h,w,r}$ 表示第 l 层第 m 个特征图上 (h, w, r) 位置的第 j 个卷积核的权重。此外,第 $(l-1)$ 层第 m 特征图上位置 $(x+h, y+w, z+r)$ 的神经元值表示为 $v_{(l-1),m}^{(x+h),(y+w),(z+r)}$ 。 $b_{l,j}$ 表示偏置。

利用 2D-CNN 可以提取高光谱遥感图像目标像素周围的局部空间信息。2D CNN 来提取空间特征,卷积层利用二维卷积核对二维输入进行运算。第 l 层第 j 个特征图上神经元 $v_{l,j}^{x,y}$ 在位置 (x, y) 的值可通过公式(13)计算。

$$v_{l,j}^{x,y} = f\left(\sum_m \sum_{h=0}^{H_l-1} \sum_{w=0}^{W_l-1} k_{l,j,m}^{h,w} v_{(l-1),m}^{(x+h),(y+w)} + b_{l,j}\right) \quad (13)$$

其中, m 为第 $(l-1)$ 层的特征图索引。卷积核的高度和宽度分别用 H_l 和 W_l 表示。 $k_{l,j,m}^{h,w}$ 表示第 l 层第 m 个特征图上 (h, w) 位置的第 j 个卷积核的权重。第 $(l-1)$ 层第 m 个特征图上神经元在位置 $(x+h, y+w)$ 上的值表示为 $v_{(l-1),m}^{(x+h),(y+w)}$ 。 $b_{l,j}$ 是偏置。

2 模型框架

2.1 整体框架

图 1 展示了用于 HSI 分类的 HDC-Attention 框架图,主要由三部分组成:KPCA 和 MBK-means 组合降维模块,HDC 特征提取模块,Spectral-Spatial Attention 模块。

本文以 Indian Pines 的数据集为例, 说明所设计模型的实现细节。原始高光谱数据立方体像素大小为 $145 \times 145 \times 200$, 其中 145 表示高度和宽度, 200 表示光谱维度。首先对原始高光谱数据沿光谱维进行 KPCA 降维。将光谱维数的数量从 200 减少到 15, 同时保持原始立方体完整的空间维度(宽度和高度), 再将降维后的特征通过 MBK-means 聚类算法聚为 20 簇, 用簇中心的平均光谱特征来标记不同的簇。这样, 高光谱图像中的光谱特征得到很大程度的压缩。KPCA 和 MBK-means 的组合方法减少了 HSI 的光谱冗余, 保留了主要特征, 为后续卷积神经网络进行特征提取打好了基础。

然后, 选择 $25 \times 25 \times 15$ 像素邻域作为输入样本。第一, 分别通过 3 个改进的 3D CNN 模块, 得到通道数为 32 的 $25 \times 25 \times 15$ 特征图。改进的 3D CNN 网

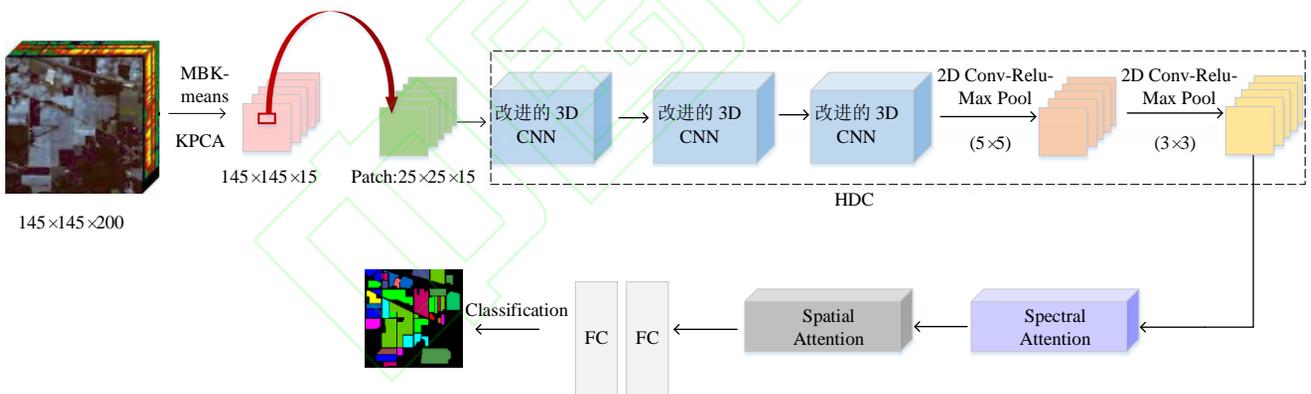


Fig.1. Network framework

图 1 网络框架

2.2 改进的 3D CNN 模块

如图 2 所示对输入特征先进行 BN 操作, 以加速网络的收敛。将输入特征每个通道分别卷积, 从而捕获每个通道的光谱和空间特征, 为了实现多尺度特征提取, 将输入特征输入到 $1 \times 1 \times 1$ 和 $3 \times 3 \times 7$ (可设置) 的双通道卷积核中, 每个通道有 16 个滤波器。不同大小的卷积核会带来不同的感受野, 既提取全局特征, 又提取局部细节特征。之后, 将两通道的特征进行级联 (concatenate), 最后再利用添加 ReLU 的 $1 \times 1 \times 1$ 3D CNN 增加网络的非线性, 使

网络可以依次增强光谱-空间特征, 产生具有代表性的特征, 增强模型的泛化能力。第二, 通过两个 2D CNN 进一步提取深度空间特征。通过有 64 个滤波器的 5×5 的 2D CNN, 生成 64 通道的 21×21 特征图。然后, 对生成的特征图进行 2×2 , 步长为 2 的最大池层的二次采样, 以降低网络的计算量, 得到 64 通道的 10×10 特征图。同样, 再通过有 64 个滤波器的 3×3 的 2D CNN, 生成 64 通道的 8×8 特征图。然后, 对生成的特征图进行 2×2 , 步长为 2 的最大池化处理, 得到 64 通道的 4×4 特征图。

再将上述特征图输入 Spectral-Spatial Attention 模块, 得到 64 通道的 2×2 特征图, 最后, 将上述特征图输入两层全连接层提取判别特征, Softmax 进行最终分类, 得到各个类别的分类精度。

得网络可以表达更加复杂的特征。在本文中, 3 个改进的 3D CNN 模块的可设置卷积大小分别为: $3 \times 3 \times 7$ 、 $3 \times 3 \times 5$ 、 $3 \times 3 \times 3$ 。

为了防止过拟合, 提高模型的泛化能力, 对于该模块中的所有卷积层使用 L2 正则化, 步长为 1 并且使用填充, 每个特征映射的空间大小在输入任何卷积层后都不会改变, 方便后续模块的叠加。相比较于单个 3D CNN, 该模块既提取了多样化特征, 又增强了模型的泛化性。

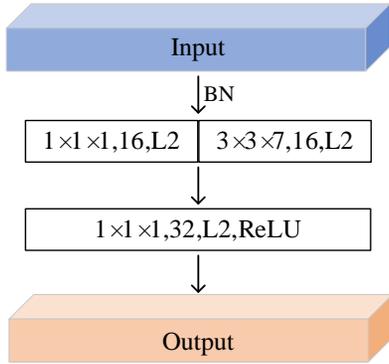


Fig.2 Improved 3D CNN module
图 2 改进的 3D CNN 模块

2.3 Spectral -Spatial Attention 模块

同一物体在不同波段的光谱响应可能有很大差异, 这表明不同波段的分辨能力是不同的。除此之外, HSI 的不同位置也有不同的语义信息。例如, 物体边缘通常比其他位置更有区别。光谱-空间注意力模块可以从 HSI 中自动学习不同波段和位置的重要性, 从而提升有用的光谱-空间信息, 抑制无用的信息。

光谱注意力模块 (Spectral Attention) 如图 3 所示。给定一个中间特征图 $F \in R^{C \times H \times W}$, 其中 C 、 H 和 W 分别表示 F 的通道数、高度和宽度。首先分别应用全局平均池化 (global average pooling, GAP) 和全局最大池化层 (global max pooling, GMP), 并进行级联为特征图 F_G 。GAP 对整个网络从结构上做正则化处理防止过拟合, 同时也是聚合其空间维度的过程, GMP 可以补充 GAP 的全局特征, 其计算如 (14) 所示。然后, 依次通过两个有 64 个滤波器, 长度为 5 的 1D CNN, 得到 64 通道的长度

为 m 的光谱注意力图 $Spa(F)$, 其计算公式如 (15) 所示。之后将光谱注意力图与中间特征图相乘, 得到沿空间维度扩展的特征图 F' , 其计算公式如 (16) 所示。最后添加两层全连接层来捕获光谱间的相关性, 提高了细化后的光谱特征图的区分能力。

$$F_G = \text{Conca}(F_{GAP}, F_{GMP}) \quad (14)$$

$$Spa(F) = \sigma(\text{Conv}(\delta(\text{Conv}(F_G)))) \quad (15)$$

$$F' = F \otimes Spa(F) \quad (16)$$

其中, σ 表示 Sigmoid 函数, δ 是 ReLU 函数。 F_{GAP} 为特征图 F 经 GAP 处理后的特征图, F_{GMP} 为特征图 F 经 GMP 处理后的特征图, \otimes 表示元素乘法。

空间注意力 (Spatial Attention) 模块如图 4 所示。使用 1×1 卷积将通道数减为 1 的特征图 F_C , 其计算公式如 (17) 所示。然后分别使用滤波器为 1, 5×5 、 3×3 的 2D CNN 导出空间注意力图 $Spa(F)$, 其计算公式如 (18) 所示。然后 $Spa(F)$ 的值沿通道维度扩展为 F' , 其计算公式如 (19) 所示。最后使用最大池化层使输出为固定大小。

$$F_C = \delta(\text{Conv}(F)) \quad (17)$$

$$Spa(F) = \sigma(\text{Conv}(\delta(\text{Conv}(F_C)))) \quad (18)$$

$$F' = F \otimes Spa(F) \quad (19)$$

光谱注意力模块更关注哪些波段是有意义的特征, 空间注意力模块更关注哪些位置的特征具有代表性。给定中间特征图, Spectral-Spatial Attention 模块按顺序推导出沿光谱和空间两个独立维度的注意力图, 然后将注意力图输入特征图重新分配光谱-空间特征的权重, 进一步增强了有用的光谱-空间特征。

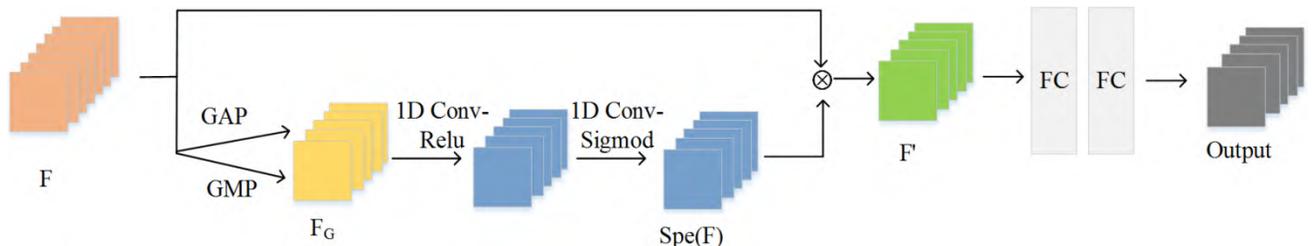


Fig.3 Spectral Attention Module
图 3 Spectral Attention 模块

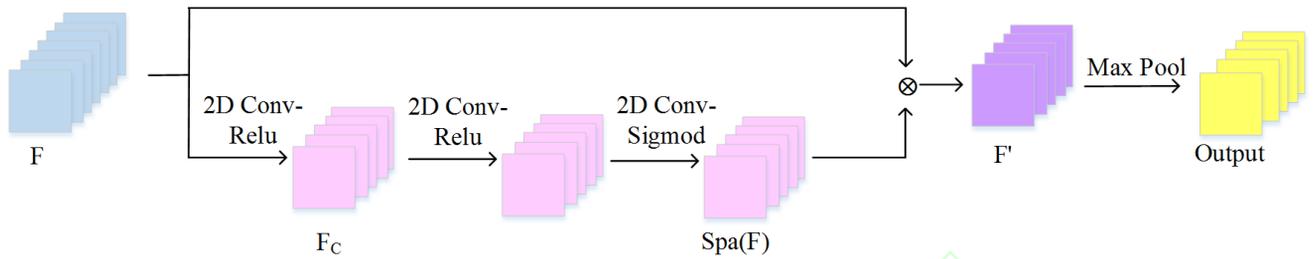


Fig.4 Spatial Attention Module
图 4 Spatial Attention 模块

3 实验及结果

在性能评估方面，对三个著名的 HSI 数据集 Indian Pines、Pavia University、Salinas scene 进行了大量的数据分类实验，使用了三个定量指标，包括平均准确度（AA）、总体准确度（OA）和 Kappa 系数（Kappa）。

所有的实验都是在 Intel(R) Xeon(R) Silver 4116 CPU@2.10GHz, 内存为 128GB 的 PC 机上运行，具体程序由 Pycharm2019 编写，在 Windows10 系统下基于 Python3.7.6 的 Tensorflow2.1 框架实现。

3.1 实验数据集介绍及实验细节描述

3.1.1 高光谱数据集

Indian Pines 是由 AVIRIS 传感器收集的印第安纳州西北部的印度松树试验场的图像，数据集的大小为 145×145 像素，剔除 20 个噪声波段，剩余的 200 个波段作为研究对象。空间分辨率约为 20 米，很容易产生混合像素，因此增加了分类的难度。它包含 16 个类别，主要是自然植被，包括超过 60% 的农业，超过 30% 的森林和其他多年生自然植被。图 5 (a) (b) 分别为数据集的彩色图像，地面基准图。

Pavia University 是由德国的机载反射光学光谱成像仪对意大利的帕维亚城附近所拍摄的高光谱数据。数据集的大小为 610×340 ，其中 12 个波段由于受噪声影响被剔除，因此一般使用的是剩下 103 个光谱波段所成的图像。该光谱所成图像的空间分辨率为 1.3m，这些像素中共包含 9 类地物，包括沥青道路、砖块、牧场等。图 6 (a) (b) 分别为数据集的彩色图像，地面基准图。

Salinas scene 是由 AVIRIS 传感器拍摄的加利福尼亚萨利纳斯山谷的图像，数据集的大小为 512×217 ，剔除 20 个噪声波段，对剩下的 204 个波段进行研究。空间分辨率为 3.7m。该数据包括花椰菜、耕地、芹菜等 16 类地物。图 7 (a) (b) 分别为数据集的彩色图像，地面基准图。

3.1.2 实验细节描述

在实验中，经过多次运行实验，将 batch size 设为 128，epoch 设为 100。学习率为 0.001，分别随机选取三个数据集的 20% 为训练数据量，80% 为测试数据量进行实验，并且重复实验 10 次，最后取这 10 次的平均值。

3.2 实验结果及分析

3.2.1 KPCA+MBK-means 的实验性能

表 1 表示数据集 Indian Pines、Pavia University、Salinas scene 在不同核函数: Poly 核、RBF 核、Linear 核和不同参数下的性能表。Poly 核善于提取样本的全局特性，具有很强的泛化能力，核函数的参数 d 取为 1-8 的正整数，本文选取 1, 2, 3, 5，从表中可以看出三个数据集上 OA、AA、Kappa 呈现先增大后减小的趋势，当 $d=2$ 时取得最大值（黑色字体标出），当 $d=5$ 时，三个数据集的 Kappa 降为 0，所以 d 的选取非常关键。RBF 核函数善于提取样本的局部信息，具有很强的学习能力，参数 γ 取值在 (1/50, 1/0.1) 范围内，本文选取 0.02, 1.25, 10。从表中可以看出，三个数据集上 OA、AA、Kappa 也呈现先增大后减小的趋势，但变化趋势较 Poly 核较小。

三个数据集在加入 MBK-means 方法所需要的

训练时间分别为：1367s、6430s、8059s，去掉 9070s、10508s，减少的运算时间分别是：559s、2449s，MBK-means 方法所需要的训练时间分别为：1926s、2640s。

Table 1 KPCA's different Kernel and parameter classification performance table on different data sets

表 1 不同数据集上 KPCA 的不同 Kernel 及参数分类性能表

		Poly (d=1,2,3,5)	RBF ($\gamma=0.02,1.25,10$)	Linear
Indian Pines	OA(%)	(98.31, 99.40 ,87.91,23.95)	(86.40,97.46,97.30)	98.63
	AA(%)	(98.52, 99.33 ,82.83,6.25)	(79.68,96.68,95.53)	98.62
	kappa \times 100	(98.22, 99.31 ,86.16,0)	(84.41,97.10,96.92)	98.44
Pavia University	OA(%)	(98.87, 99.93 ,90.57,37.98)	(89.34,98.32,96.52)	98.89
	AA(%)	(98.90, 99.95 ,86.47,17.09)	(82.43,97.73,97.31)	98.57
	kappa \times 100	(98.98, 99.97 ,88.49,0)	(86.68,98.90,97.34)	98.87
Salinas scene	OA(%)	(98.79, 99.98 ,91.23,35.32)	(88.24,98.21,96.89)	98.98
	AA(%)	(98.90, 99.97 ,87.81,18.25)	(85.23,97.21,96.86)	98.88
	kappa \times 100	(98.78, 99.98 ,89.15,0)	(87.34,98.32,96.91)	98.89

3.2.2 改进的 3D CNN 以及 HDC 的实验性能

表 2 表示不同数据集上改进的 3D CNN, HDC 的分类性能表。可以看出，仅通过 3D CNN 提取特征表现较差，但是通过改进 3D CNN 可以更好的提取到光谱-空间特征，在 Indian Pines 数据集上 OA, AA, Kappa 分别增加了 3.1%、4.1%、4.26%。因为 Indian Pines 数据集的噪声相对较大，多尺度特征也为数据去噪提供了很好的优化，使模型更

稳定，所以，多尺度特征在 HSI 的深度特征提取中起着重要的作用。

3 个改进的 3D CNN 模型的叠加融合了深度和尺度信息，特征提取效率提高。HDC 模块在后续添加了 2D CNN, 2D CNN 在 3D CNN 的基础上进一步学习了更多抽象级别的空间表示，同时降低了模型的复杂性，在 Indian Pines 数据集上 OA、AA、Kappa 分别增加了 0.87%、1.52%、1.29%。

Table 2 Improved 3D CNN, HDC classification performance table on different data sets

表 2 不同数据集上改进的 3D CNN, HDC 的分类性能表

		3D CNN	改进的 3D CNN	HDC
Indian Pines	OA(%)	94.39	96.49	97.36
	AA(%)	93.38	96.39	97.64
	kappa \times 100	92.93	96.19	97.48
Pavia University	OA(%)	96.82	98.57	99.10
	AA(%)	95.92	97.84	98.87
	kappa \times 100	95.78	98.87	99.13
Salinas scene	OA(%)	96.33	97.17	98.23
	AA(%)	96.32	98.11	98.92
	kappa \times 100	96.28	98.58	99.01

3.2.3 Spectral -Spatial Attention 的实验性能

表 3 表示不同数据集上 Spectral -Spatial Attention 的分类性能表。可以看出，在三个数据集上，注意力模块进一步提高了模型的精确度，体现了该模块的有效性。从表中可以看到，Indian Pines 数据集仅添加 Spectral Attention 模块，OA、AA、Kappa 分别增加 0.94%、1.13%、0.95%；仅添加 Spatial Attention 模块，OA、AA、Kappa 分

别增加 1.05%、0.67%、0.94%；添加 Spectral-Spatial Attention 模块，OA、AA、Kappa 分别增加 2.04%、1.69%、1.83%。这主要是因为 Spectral Attention 模块主要由 1D CNN 更多的关注光谱特征，Spatial Attention 模块主要由 2D CNN 更多的关注空间特征，两个注意力的组合进一步关注更典型的光谱-空间特征，从而进一步提高了分类的精确度。

Table 3 Classification performance table with Spectral -Spatial attention module on different data sets
表 3 不同数据集上 Spectral -Spatial Attention 的分类性能表

		Spectral	Spatial	Spectral-Spatial
Indian Pines	OA(%)	98.30	98.41	99.40
	AA(%)	98.57	98.31	99.33
	Kappa×100	98.43	98.42	99.31
Pavia University	OA(%)	99.23	99.43	99.97
	AA(%)	99.25	99.35	99.96
	Kappa×100	99.36	99.47	99.96
Salinas scene	OA(%)	99.48	99.39	99.98
	AA(%)	99.03	99.21	99.97
	kappa×100	99.37	99.41	99.98

3.2.4 不同方法下的实验性能分析

为了验证所提出模型 (HDC-Attention) 的有效性, 将 HDC-Attention 方法与其他方法进行比较, 比较方法包括 HybridSN^[25]、SSRN^[26]、Cluster CNN^[16]、Fast 3D CNN^[27]。这五种方法都是基于深度学习的分类方法。表 4、5、6 分别表示了 Indian Pines、Pavia University、Salinas scene 数据集上的不同分类方法以及每一类的分类结果, 每类中将最高精度用黑体标出。从表中可以看出, HDC-Attention 在三个数据集上的分类精度普遍高

于其他方法, 特别是 Pavia University 数据集, 有六类地物分类完全正确。Pavia University 和 Salinas scene 数据集的分类精度均高于 Indian Pines, 这是因为这两个数据集本身的数据清晰度高于 Indian Pines 数据集, 所以分类难度相较于 Indian Pines 数据集也低。在训练时间方面, Fast 3D CNN 模型更有优势, 在三个数据集上的训练时间均为最少, HDC-Attention 方法的训练时间较长, 因为 HDC 模块保留了更为丰富的光谱-空间特征, 所以需要更长的计算时间。

Table 4 Comparison table of classification accuracy of each category of Indian Pines under different methods
表 4 不同方法下 Indian Pines 每类分类精度对比表

#	Class	SSRN	Fast 3D CNN	Cluster CNN	HybridSN	HDC-Attention
1	Alfalfa	98.24	100.00	100.00	97.30	100.00
2	Corn-notill	98.30	98.11	97.02	96.85	97.98
3	Corn-mintill	97.39	92.14	98.72	99.70	100.00
4	Corn	98.46	98.21	98.89	100.00	100.00
5	Grass-pasture	99.21	100.00	98.33	99.74	97.40
6	Grass-trees	98.92	100.00	99.01	99.83	99.14
7	Grass-pasture-mowed	100.00	100.00	100.00	100.00	100.00
8	Hay-windrowed	99.25	88.00	87.50	100.00	100.00
9	Oats	97.69	100.00	98.43	93.75	100.00
10	Soybean-notill	98.50	96.42	98.87	98.71	100.00
11	Soybean-mintill	98.36	100.00	99.23	99.39	99.84
12	Soybean-clean	97.37	97.32	97.06	96.63	99.16
13	Wheat	98.29	97.60	100.00	99.39	100.00
14	Woods	98.73	100.00	99.47	100.00	99.90
15	Buildings-Grass-Trees-Drives	96.79	97.70	100.00	99.68	100.00
16	Stone-Steel-Towers	95.49	100.00	95.94	97.30	95.94
	OA(%)	98.00	98.00	98.77	98.83	99.40
	AA(%)	98.35	97.69	98.83	98.98	99.33
	Kappa×100	97.71	97.72	98.66	98.64	99.31
	Train Time(s)	2031	113	199	171	1367

Table 5 Comparison table of classification accuracy of each category of Pavia University under different methods
表 5 不同方法下 Pavia University 每类分类精度对比表

#	Class	SSRN	Fast 3D CNN	Cluster CNN	HybridSN	HDC-Attention
1	Asphalt	99.89	97.74	99.57	100.00	100.00
2	Meadows	99.43	99.75	98.98	100.00	100.00
3	Gravel	98.67	93.93	100.00	99.94	100.00
4	Tree	99.74	96.98	98.91	99.31	99.61
5	Painted metal sheets	100.00	100.00	99.91	100.00	100.00
6	Bare Soil	99.71	98.83	100.00	99.89	99.40
7	Bitumen	98.16	92.11	98.08	100.00	100.00
8	Self-blocking bricks	98.02	95.01	99.06	99.90	99.94
9	Shadows	100.00	99.87	98.59	99.98	100.00
	OA(%)	98.69	98.07	98.77	99.37	99.93
	AA(%)	98.79	98.13	98.50	99.48	99.95
	Kappa \times 100	99.27	97.16	98.70	99.29	99.97
	Train Time(s)	6521	512	851	237	6430

Table 6 Comparison table of classification accuracy of each category of Salinas scene under different methods
表 6 不同方法下 Salinas scene 每类分类精度对比表

#	Class	SSRN	Fast 3D CNN	Cluster CNN	HybridSN	HDC-Attention
1	Bro- coli_green_weeds_1	98.54	95.79	100.00	98.57	100.00
2	Bro- coli_green_weeds_2	98.44	97.93	96.46	97.98	100.00
3	Fallow	97.47	97.01	99.04	99.75	100.00
4	Fallow_rough_plow	98.21	96.37	97.76	100.00	99.85
5	Fallow_smooth	98.71	100.00	98.58	99.69	100.00
6	Stubble	98.89	99.30	98.96	98.83	100.00
7	Celery	100.00	100.00	100.00	100.00	100.00
8	Grapes_untrained	99.45	99.25	100.00	100.00	100.00
9	Soil_vinyard_develop	97.54	100.00	98.10	98.75	99.96
10	Corn_senesced_green_ weeds	97.79	98.72	98.58	98.49	100.00
11	Lettuce_romaine_4wk	98.30	98.59	99.47	99.92	100.00
12	Lettuce_romaine_5wk	97.12	97.06	96.45	97.95	100.00
13	Lettuce_romaine_6wk	98.78	97.60	100.00	99.39	99.88
14	Lettuce_romaine_7wk	98.63	100.00	98.47	100.00	99.97
15	Build- ings-Grass-Trees-Vin- yard_untrained	96.99	97.50	100.00	99.61	100.00
16	Vinyard_vertical_trellis	94.00	93.22	97.59	98.50	100.00
	OA(%)	98.87	98.62	98.47	98.87	99.79
	AA(%)	97.98	97.59	98.29	98.98	99.94
	Kappa \times 100	97.59	97.97	98.57	98.97	99.96
	Train Time(s)	7989	760	1201	1020	8059

图 5-7 的 (c)、(d)、(e)、(f)、(g) 给出了不同方法在三个数据集上的分类图谱，从图中可以看出，在 Indian Pines 数据集上，SSRN、Fast 3D CNN、Cluster CNN 方法最终分类结果图中含有一些斑点，并且存在小区域内错分的情况，HybridSN 方法的结果相对较高一些，已没有大量的斑点，而 HDC-Attention 模型斑点非常少。

Indian Pines 数据集中类别 2 的“Corn-notill”和类别 4 的“Corn”一些像素被错误地分类为彼此，因为这两个土地覆盖区域有相似的光谱-空间特征，可以观察到通过提出的模型获得的分类图是最平滑和最清晰的。特别是与其他模型相比，由于椒盐噪声最少，分类结果更接近实际情况。原因有以下两点：首先，混合深度网络能够提取有

效且具有辨识性的空谱特征，这可以使模型更好地识别不同类型的地物特征，甚至是具有相似光谱-空间信息的土地覆盖物。其次，Spectral-Spatial Attention 模块增强了有用的空谱特征，这有利于更好的分类。可以看到 HDC-Attention 方法在所有数据集上比其他方法具有更少的误分类，特别是

对于一些相似的类。

总的来说，与其它四种方法相比，HDC-Attention 方法取得了比较好的结果，分类结果图含有非常少量的斑点，且在同类的小区域内相对平滑，好几类地物几乎完全正确分类。

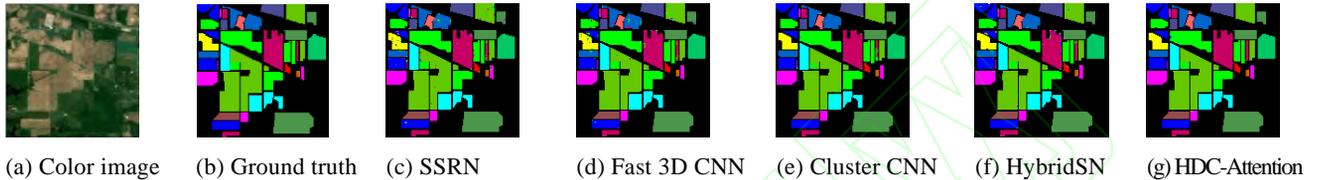


Fig.5 The classification maps of Indian Pines using different methods

图 5 不同方法在 Indian Pines 上的分类图

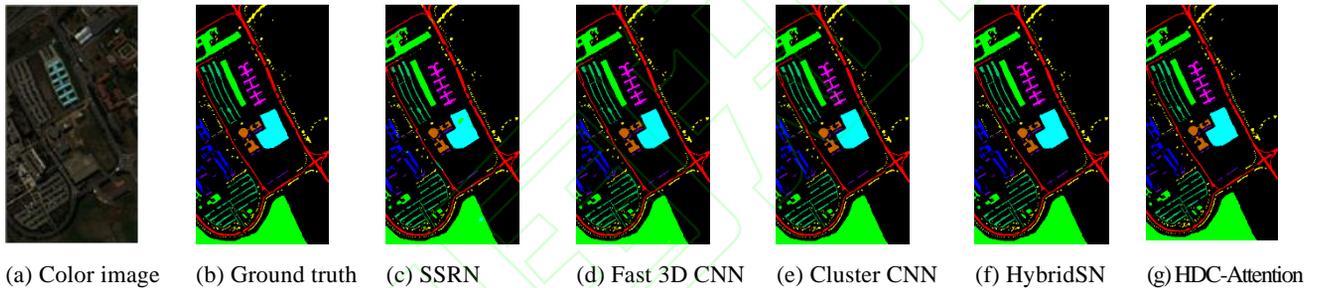


Fig.6 The classification maps of Pavia University using different methods

图 6 不同方法在 Pavia University 上的分类图

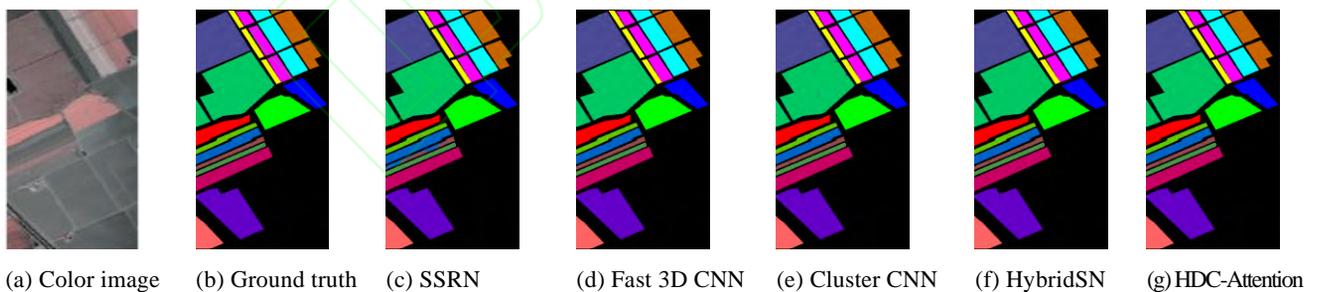


Fig.7 The classification maps of Salinas scene using different methods

图 7 不同方法在 Salinas scene 上的分类图

4 结束语

本文提出的 HDC-Attention 分类方法，首先利用 KPCA 和 MBK-means 对高光谱图像进行组合降维，然后将降维后的数据输入 HDC 网络，不仅进行充分的光谱-空间特征提取，还加强了空间特征。最后利用 Spectral-Spatial Attention 模块，在 HDC 模块的基础上，增强有用的光谱-空间特征，抑制无用的特征进行分类任务。

本文首先做了 KPCA 和 MBK-means 的组合降维，证明了 KPCA 降维的有效性，MBK-means 进一步降维，减少了计算时间。然后利用基于改进的 3D CNN 模块构建了 HDC 模块，充分证实了 HDC 能够有效提取光谱和空间的特征，此外还做了加入 Spectral-spatial Attention 模块的对比实验，证明了该模块能够进一步增强有用的空谱特征。所提出的

方法不仅在有限的样本下表现出了较好的分类性能，而且降低了模型复杂度。

虽然所提出的方法和其他方法相比，分类的性能较好，但是还存在不足之处。之后将进一步优化模型继续提高模型效率和鲁棒性，用更少的训练样本来达到更高的分类精度，继续探索优化 KPCA 的参数问题。

参考文献:

- [1] Imani M, Ghassemian H. An overview on spectral and spatial information fusion for hyperspectral image classification: Current trends and challenges[J]. *Information Fusion*, 2020, 59:59-83.
- [2] Naoto Y, Jonathan C, Karl S. Potential of Resolution-Enhanced Hyperspectral Data for Mineral Mapping Using Simulated EnMAP and Sentinel-2 Images[J]. *Remote Sensing*, 2016, 8(3):172.
- [3] Liang L, Di L, Zhang L, et al. Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method[J]. *Remote Sensing of Environment*, 2015, 165: 123-134.
- [4] TAN H F, LUO T W, YANG G, et al. Research on background suppression methods in hyperspectral anomaly detection[J]. *Optoelectronics. Laser*, 2016, 27(2): 177-181.
谭海峰,罗天文,杨桃等.高光谱异常检测中背景抑制方法研究[J].*光电子. 激光*,2016,27(2):177-181.
- [5] Caporaso N, Whitworth M B, Grebby S, et al. Non-destructive analysis of sucrose, caffeine and trigonelline on single green coffee beans by hyperspectral imaging[J]. *Food Research International*, 2018, 106:193-203.
- [6] Zhang X, Wang Y, Zhang N, et al. SSDANet: Spectral-spatial Three-dimensional Convolutional Neural Network for Hyperspectral Image Classification[J]. *IEEE Access*, 2020, 8: 127167-127180.
- [7] Lai H, Deng J, Wen S. Application of ToF-SIMS and PCA to study interaction mechanism of dodecylamine and smithsonite[J]. *Applied Surface Science*, 2019, 496: 143698.
- [8] Hajer N, Raoudha B D, Ikram Amous B A. Efficient cloud service discovery approach based on LDA topic modeling[J].*Journal of Systems and Software*,2018,146:233-248.
- [9] Zhi-Min L I, Jie Z, Hong H, et al. Semi-supervised bundle manifold learning for hyperspectral image classification[J]. *optics and precision engineering*, 2015, 23(5):1434-1442.
- [10] Lin, He, Jun, 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.
- [11] Hong D, Wu X, Ghamisi P, et al. Invariant Attribute Profiles: A Spatial-Frequency Joint Feature Extractor for Hyperspectral Image Classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(6): 3791- 3808.
- [12] Cao X, Xu L, Meng D, et al. Integration of 3-dimensional discrete wavelet transform and Markov random field for hyperspectral image classification[J]. *Neurocomputing*, 2017, 226(Feb.22):90-100.
- [13] Okan Bilge Özdemir A, Gedik B E, Yasemin Yardımcı Çetin C. Hyperspectral classification using stacked auto-encoders with deep learning[C]//*Workshop on Hyperspectral Image & Signal Processing: Evolution in Remote Sensing*. IEEE, 2017.
- [14] Zhao X, Chen Y S, Jia X P. Spectral-Spatial Classification of Hyperspectral Data Based on Deep Belief Network[J]. *IEEE journal of selected topics in applied earth observations and remote sensing*, 2015, 8(6):2381-2392.
- [15] Hu W H, Wei Y Y, Li Z. et al. Deep Convolutional Neural Networks for Hyperspectral Image Classification[J]. *Journal of Sensors*, 2015, 2015:1-12.
- [16] Yao W, Lian C, Bruzzone L. ClusterCNN: Clustering-Based Feature Learning for Hyperspectral Image Classification[J]. *IEEE Geoscience and Remote Sensing Letters*, 2020: 1-5.
- [17] WEI X P, YU X C, TAN X, et al. Hyperspectral image classification based on CNN and 3D Gabor filter[J]. *Journal of Computer Aided Design and Graphics*, 2020, 32(1): 90-98.
魏祥坡,余旭初,谭熊等.CNN 和三维 Gabor 滤波器的高光谱图像分类[J].*计算机辅助设计与图形学学报*, 2020, 32(1): 90-98.
- [18] Hamida A B, Benoit A, Lambert P et al. 3-D Deep Learning Approach for Remote Sensing Image Classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(8):4420-4434.
- [19] Zhang C J, Li G D, Du S H. et al. Multi-Scale Dense Networks for Hyperspectral Remote Sensing Image Classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*,2019,57(11):9201-9222.
- [20] Lu Z, Xu B, Sun L, et al. 3D Channel and Spatial Attention Based Multi-Scale Spatial Spectral Residual Network for Hyperspectral Image Classification[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, PP(99):1-1.
- [21] Sun H, Zheng X, Lu X, et al. Spectral-Spatial Attention Network for Hyperspectral Image Classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(5): 3232-3245.
- [22] Fang B, Li Y, Zhang H, et al. Hyperspectral Images Classification Based on Dense Convolutional Networks with Spectral-Wise Attention Mechanism[J]. *Remote Sensing*, 2019, 11(2):159.

- [23]Lahdhiri H, Taouali O. Reduced Rank KPCA based on GLRT chart for sensor fault detection in nonlinear chemical process[J]. Measurement,2021,169.
- [24]XIU Y C, YANG W J. Application of Mini Batch K-means Algorithm in Remote Sensing Image Classification[J]. Journal of Ludong University (Natural Science Edition), 2017, 33(4):359-363.
修瑛昌,杨文静.Mini Batch K-means 算法在遥感影像分类中的应用[J].鲁东大学学报(自然科学版), 2017, 33(4): 359-363.
- [25] Roy S K, Krishna G, Dubey S R, et al. HybridSN: Exploring 3D-2D CNN Feature Hierarchy for Hyperspectral Image Classification[J]. IEEE Geoscience and Remote Sensing Letters, 2020, 17(2):277-281.
- [26]Zhong Z L, Li J, Luo Z M, et al. Spectral-Spatial Residual Network for Hyperspectral Image Classification: A 3-D Deep Learning Framework[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(2):847-858.
- [27]Muhammad A.A Fast 3D CNN for Hyperspectral Image Classification[J]. IEEE Geoscience and Remote Sensing Letters, 2020.



王燕 (1971-), 女, 甘肃泾川, 硕士, 教授, CCF 会员, 主要研究方向为模式识别与人工智能。
WANG Yan, born in 1971, M.S. professor. CCF member. Her research interests include Pattern recognition and Artificial intelligence.



吕艳萍 (1995-), 女, 甘肃榆中, 硕士, 主要研究方向为模式识别与人工智能。
LYU Yanping, born in 1995, M.S. Her research interests include Pattern recognition and Artificial intelligence.