

基于快速双分支密集连接网络和双注意力机制的高光谱图像分类

杨国亮, 龚家仁*, 习浩, 李世聪, 邹俊峰

江西理工大学电气工程与自动化学院, 江西 赣州 341000

摘要 近年来, 基于深度学习的高光谱图像(HSI)分类研究引起了各领域的广泛关注。HSI光谱波段数多、信息冗余度高、计算复杂, 出现训练样本不足的问题, 容易导致模型训练过拟合, 影响分类精度。为了提高分类精度并减少训练时间, 提出一种基于三维卷积神经网络(3D-CNN)并结合双分支双注意力机制的快速密集连接网络, 用于HSI的分类。首先利用主成分分析(PCA)对原始数据进行降维, 减少冗余信息, 然后采用双分支密集连接结构, 并结合快速傅里叶变换(FFT)的双分支高效通道注意力(ECA)机制, 同时增加了一个FFT层, 既保证了模型分类精度, 也加快了模型的训练速度。在多个公共高光谱数据集上对方法进行实验验证, 性能评估指标采用总体分类精度(OA), 平均分类精度(AA)和Kappa系数。实验结果表明, 所提方法在提高分类精度的同时, 显著减少了训练时间和测试时间。

关键词 图像处理; 高光谱图像分类; 主成分分析; 密集连接; 双注意力机制; 快速傅里叶变换

中图分类号 TP751.2

文献标志码 A

doi: 10.3788/LOP202259.0810002

Hyperspectral Image Classification Based on Fast Double Branch Dense Connection Network and Double Attention Mechanism

Yang Guoliang, Gong Jiaren*, Xi Hao, Li Shicong, Zou Junfeng

School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou, Jiangxi 341000, China

Abstract In recent years, the classification of hyperspectral images (HSI) based on deep learning has attracted extensive attention in various fields. The number of HSI spectral bands is large and information redundancy is high, which increases calculation complexity. However, a lack of training samples can easily lead to overfitting of model training and reduce classification accuracy. To improve the classification accuracy and reduce the training time, this study proposes a fast dense connection network based on a three-dimensional convolutional neural network (3D-CNN) with double branch and double attention mechanism for HSI classification. In this study, the original data is dimensionless using principal component analysis (PCA) to reduce redundant information. Then, a double efficient channel attention (ECA) mechanism with a double branch dense connection structure and a fast Fourier transform (FFT) is used. It not only ensures the model's classification accuracy but also increases the model's training speed. We conducted experiments on multiple public hyperspectral datasets, and the performance evaluation indices are overall classification accuracy (OA), average classification accuracy (AA), and Kappa coefficient. The experimental results show that the proposed method can improve the classification accuracy and significantly reduce the training time and testing time.

Key words image processing; hyperspectral image classification; principal component analysis; dense connection; dual attention mechanism; fast Fourier transform

收稿日期: 2021-07-02; 修回日期: 2021-08-19; 录用日期: 2021-08-25

基金项目: 国家自然科学基金(51365017)

通信作者: *964331424@qq.com

1 引言

高光谱图像(HSI)广泛应用于军事、农业、林业、地质、海洋等领域,在各领域发挥着至关重要的作用,具有广阔的发展前景。HSI包含丰富的空间特征和光谱特征,但信息冗余度高、光谱维度高、空间分辨率低,导致数据难以处理,各类地物难以区分;而且HSI训练样本不足,模型训练容易过拟合,导致分类精度低。

为了解决上述问题,国内外研究者提出主要以传统机器学习和深度学习为代表的HSI分类方法。Melgani等^[1]首次将支持向量机(SVM)应用于HSI分类,可以有效提取HSI光谱维度的高维特征,但是该方法没有充分利用空间维度的空间特征,导致分类效果不佳。基于稀疏表示的HSI分类也被广泛应用,Zhang等^[2]采用同步正交匹配追踪技术求解非局部加权联合稀疏模型,利用了相邻像素之间的空间结构特征,提高了分类精度,但计算成本高。董安国等^[3]利用了谱聚类方法和稀疏模型确定像元的类,利用了HSI的光谱空间信息,考虑了噪声和边界对分类效果的影响。Li等^[4]只用整个训练集中一组特定的局部自适应样本来估计每个测试像素,协同表示向量,相比传统稀疏表示的算法,减少了训练时间。以上算法都只提取到HSI的浅层特征,而基于深度学习的HSI分类方法可以提取更加深层的特征,Chen等^[5]首次将深度学习的概念引入到HSI分类中,提出了一种联合光谱-空间特征的深度学习框架,采用单层自动编码器和多层堆叠自动编码器分别学习HSI的浅层特征和深层特征,展现了深度学习在HSI分类中的巨大潜力。Hu等^[6]利用深度卷积神经网络直接在光谱维度对HSI进行分类,但忽略了丰富的空间信息。黄鸿等^[7]利用多层深度学习模型对高光谱数据进行逐层学习,并联合空谱特征对HSI进行分类。Yue等^[8]将结合分层学习的深度卷积神经网络引入到HSI分类中,并结合逻辑回归构建分类器,通过训练分类器自动学习深层光谱和空间特征,得到更好的分类效果。基于三维卷积神经网络(3D-CNN)的深度学习算法^[9-11]具有1D-CNN的光谱特征提取能力,也具有2D-CNN的空间特征提取能力,可以同时提取空间光谱特征。ResNet^[12]和DenseNet^[13]的提出有效解决了深度神经网络的退化问题和梯度消失问题。Zhong等^[14]结合ResNet提出一种光谱空间残差网络

(SSRN),残差结构有助于梯度的反向传播,有效防止梯度爆炸和网络退化,网络中的空间和光谱残差块可以深度学习和判别丰富的空间特征和光谱特征。Wang等^[15]结合DenseNet的思想,提出一种端到端的快速密集光谱空间卷积神经网络(FDSSC),使用不同大小的卷积核分别提取空间特征和光谱特征,采用密集连接方式充分学习HSI的深度特征。Ma等^[16]提出一种双分支多注意力机制网络(DBMA),网络分为空间分支和光谱分支,并对两个分支采用两种不同的注意力机制,提取更具区分性和代表性的特征,减小了空间特征和光谱特征之间的干扰。Li等^[17]提出一种端到端的双分支双注意力机制网络(DBDA),采用了密集连接结构和Mish激活函数^[18]。Mish激活函数的平滑连续性提高了模型的特征表达和信息流通能力,同时提高了模型的优化和泛化性能,提高了分类精度,但DBDA训练和测试时间长,时间计算成本高。

为了减少训练和测试时间的同时提高分类精度,本文提出一种基于双分支和双注意力机制的快速密集网络(FDBDA)。网络的整体结构由空间分支和光谱分支组成。双分支结构中均采用3D-CNN,并在3D-CNN中用Mish激活函数代替ReLU激活函数,提高了特征表达和信息流通能力,同时也提高了模型的泛化性和鲁棒性。先对原始HSI进行预处理,采用主成分分析(PCA)对原始HSI进行降维,减少信息冗余,降低计算复杂度,然后对数据进行标准化。在光谱分支中,为了进一步提取HSI的显著特征,对PCA降维后的数据再进行下采样,在保证精度的同时减少冗余信息,经密集连接块后,采用一种更轻量的高效通道注意力(ECA)机制来提取HSI更深层的特征。受到FNet^[19]的启发,在引入注意力机制的同时对特征图进行快速傅里叶变换(FFT),并增加一个FFT层来替换一个卷积层,加快了计算速度。在几个公共数据集上对所提方法进行了实验,实验结果表明,所提方法在分类精度、训练时间、测试时间上都表现出优越的性能。

2 理论知识

2.1 数据预处理

HSI的信息高度冗余,处理时容易导致维数灾难,计算耗时长同时严重影响分类精度,因此采用PCA先对HSI进行特征降维,针对不同数据集选

取不同的主成分分量。为了便于对 HSI 数据进行分析,还需要对降维后的数据进行标准化。

假设原始三维 HSI 大小为 $H \times W \times L$, H, W, L 分别表示三维 HSI 的高度,宽度,波段数。可将三维 HSI 数据表示为向量 $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N]^T \in \mathbf{R}^{N \times L}$, N 表示像素个数, $N = H \times W$ 。

经过 PCA 降维后的 HSI 数据波段维度减为 B , 因此降维后的 HSI 数据可以表示为 $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N]^T \in \mathbf{R}^{N \times B}$ 。对降维后的数据进行标准化,标准化公式为

$$\tilde{X}_n = \frac{(X_n - X_{\text{mean}})}{X_{\text{std}}}, n = 1, 2, \dots, N, \quad (1)$$

式中: X_{mean} 表示 HSI 所有像素的均值; X_{std} 表示 HSI 所有像素的标准差。标准化后的数据为 $\tilde{\mathbf{X}} = [\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_N]^T \in \mathbf{R}^{N \times B}$ 。

选择以目标像素为中心,大小为 $P \times P \times B$ 的 HSI 小立方体数据作为网络的输入,输入数据可以表示为 $\tilde{\mathbf{X}} = [\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_N]^T \in \mathbf{R}^{M \times B}$, 其中 $M = P \times P$ 。

2.2 3D-CNN

3D-CNN 具备 1D-CNN 提取每个像素光谱特征的功能,也具备 2D-CNN 提取每个像素局部空间特征的功能。3D 卷积沿着光谱方向对三个光谱维度进行卷积,生成两个特征图,示意图如图 1 所示。3D-CNN 第 i 层中的第 j 个特征图的神经元在 (x, y, z) 处的值 v_{ij}^{xyz} 可以表示为

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

式中: m 对连接到当前第 j 个特征图的第 $i-1$ 层的特征图进行索引; ω_{ijm}^{pqr} 表示在 (p, q, r) 处连接到第 m 个特征图的权重; P_i 和 Q_i 表示空间卷积核的高度和宽度, R_i 表示光谱卷积核的长度; b_{ij} 是第 i 层中第 j 个特征图的偏差; $g(\cdot)$ 表示激活函数。

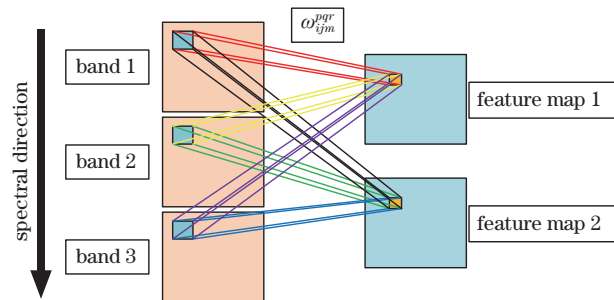


图 1 3D 卷积沿光谱方向进行卷积的示意图

Fig. 1 3D convolution schematic along the spectral direction

2.3 密集连接结构

ResNet 最早提出残差结构,目的是解决深度网络的退化问题,计算公式为

$$\mathbf{x}_l = H_{l-1}(\mathbf{x}_{l-1}) + \mathbf{x}_{l-1}, l \in \mathbf{N}^+, \quad (3)$$

式中: H 表示隐藏层,包含卷积层、激活层和批量归一化层; \mathbf{x}_{l-1} 和 \mathbf{x}_l 分别表示第 $l-1$ 个残差块的输入和输出。ResNet 结构如图 2 所示。

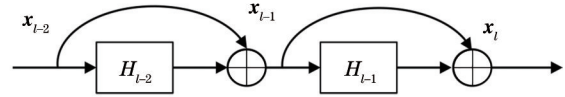


图 2 ResNet 结构示意图

Fig. 2 Schematic of ResNet structure

DenseNet 在 ResNet 的基础上进行了改进,缓解了梯度消失,减少了网络参数量,将第 $l-1$ 个残差块之前的每一个残差块的输出作为第 $l-1$ 个残差块的输入,充分利用了每一层的特征,计算公式为

$$\mathbf{x}_l = H_{l-1}([\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{l-1}]), l \in \mathbf{N}^+. \quad (4)$$

DenseNet 采用密集连接方式,结构如图 3 所示。

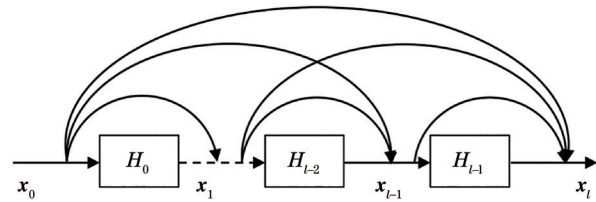


图 3 DenseNet 结构示意图

Fig. 3 Schematic of DenseNet structure

所采用的密集连接结构如图 4 所示。

选择以目标像素为中心,大小为 $P \times P \times B$ 的 HSI 小立方体数据作为密集连接网络的输入,初始通道数为 C_0 ,密集网络层数为 a ,卷积核个数为 k ,密集连接网络通道数满足的线性关系为

$$C_a = C_0 + a \times k, a \in \mathbf{N}^+. \quad (5)$$

2.4 Mish 激活函数

在图像分类问题上,加入激活函数的目的是将非线性特征引入到神经网络模型中,解决线性不可分的问题。采用 Mish 激活函数代替 ReLU 激活函数,作为一种自正则非单调的平滑连续激活函数,其无上界特性,可以有效避免梯度饱和问题;有下界特性,可以提高正则化性效果;同时保留了小部分负权值,可以有效防止梯度信息丢失,提高特征表达和信息流通能力;而平滑连续性可以提高模型的泛化性和鲁棒性。Mish 激活函数的数学表达式为

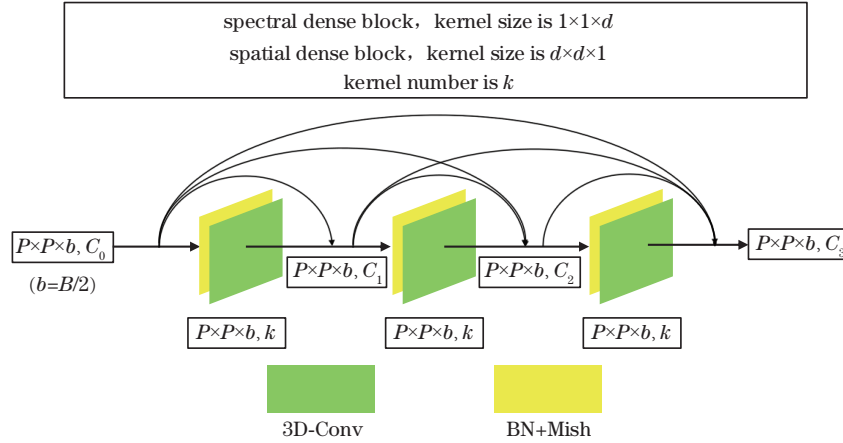


图 4 所提方法中密集连接结构示意图

Fig. 4 Schematic of dense connection structure in the proposed method

$$f(x) = x \tanh[\text{softplus}(x)] = x \tanh\{\ln[1 + \exp(x)]\}. \quad (6)$$

Mish 激活函数的值域为 $[\sim 0.31, +\infty)$, 其一次导数表达式为

$$f'(x) = \frac{\exp(x)\omega}{\delta^2}, \quad (7)$$

其中 $\omega = 4(x+1) + 4\exp(2x) + \exp(3x) + \exp(x)(4x+6)$, $\delta = 2\exp(x) + \exp(2x) + 2$.

2.5 ECA 机制

注意力机制可以表达更深层的特征, 为了提升训练速度并保证精度, 采用一种轻量级 ECA 模块, 如图 5 所示。ECA 先通过全局平均池化(GAP)获得集中的特征, 采用快速 1-D 卷积生成通道共享权重, 通过通道维度的非线性映射自适应确定卷积核大小 k , 通过带状矩阵来学习通道注意力。

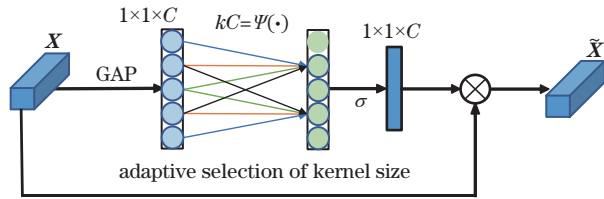


图 5 ECA 机制

Fig. 5 ECA mechanism

采用快速 1-D 卷积生成通道共享权重:

$$\omega = \sigma[C1D_k(y)], y \in \mathbf{R}^C, \quad (8)$$

式中: $C1D(\cdot)$ 表示 1-D 卷积; σ 表示 sigmoid 激活函数。

通道数 C 和卷积核大小 k 之间存在一种非线性映射, 可以表示为

$$C = 2^{(\gamma \times k - q)}, \quad (9)$$

式中: γ 和 q 表示常数项系数。

其中卷积核大小 k 由非线性映射自适应选择:

$$k = \phi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{q}{\gamma} \right\rfloor_{\text{odd}}, \quad (10)$$

式中: $\lfloor \cdot \rfloor_{\text{odd}}$ 表示最接近的奇数; γ 和 q 分别设置为 2 和 1。

2.6 快速傅里叶变换

FNet 用一个傅里叶子层替换自注意力层, 加快了 transform 的编码器速度, 且精度损失较小。受到 FNet 的启发, 为了提高速度的同时减小精度的损失, 引入注意力机制的同时对特征图进行 FFT, 并增加一个 FFT 子层, 因为傅里叶变换比卷积和注意力机制的计算更快, 这样既加快了模型的训练速度, 也增强了特征提取能力和稳定性, 最后对两个输出特征图进行融合, 通过实验证明, 这是简单有效的。

给定一个序列 $\{x_n\}, n \in [0, N-1]$, 离散傅里叶变换(DFT)可定义为

$$X(a) = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N}na}, 0 \leq a \leq N-1, \quad (11)$$

式中: $X(a)$ 表示经过 DFT 后的数据。由 (11) 式可知, DFT 的时间复杂度为 $O(N^2)$, FFT 是离散傅里叶变换的一种快速算法, 其通过蝶形运算递归求得结果, 减少了 DFT 乘法次数, 时间复杂度减小到 $O(N \log N)$ 。

3 网络框架

采用 3D-CNN 框架, 结合空间分支和光谱分支的双分支密集连接结构, 并分别引入空间注意力模块和光谱注意力模块, 所提网络结构如图 6 所示。

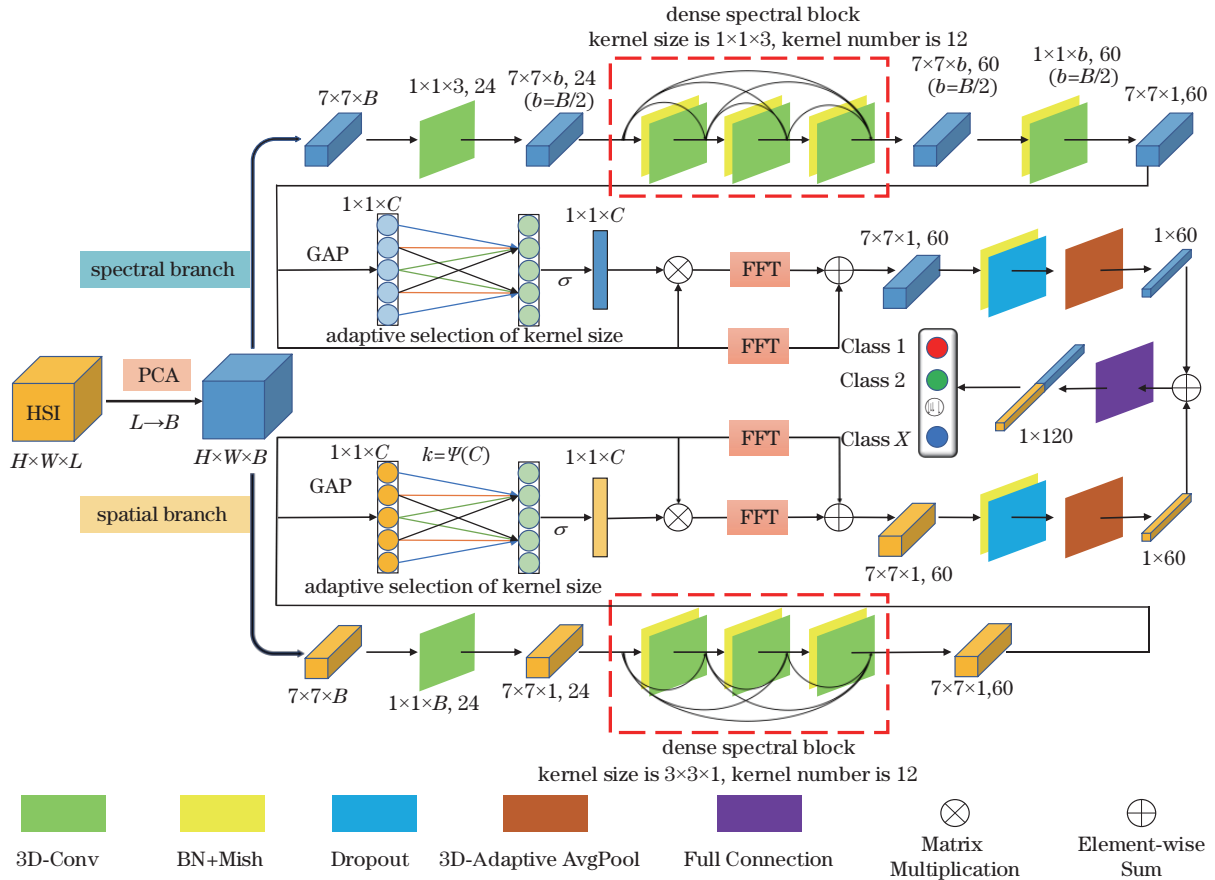


图 6 所提网络结构示意图

Fig. 6 Schematic of the proposed network structure

3.1 光谱分支结构

光谱分支结构先采用 $1 \times 1 \times 7$ 卷积核将 PCA 降维后的 HSI 的光谱维度再次进行特征提取, 进一步选择更加有价值的特征维度, 卷积核个数设为 24, 步长设置为 (1, 1, 2), 得到 b ($b = B/2$) 个光谱特征维度的光谱特征图。光谱密集连接块卷积层采用 $1 \times 1 \times 3$ 卷积核, 卷积核数量设置为 12, 步长设置为 (1, 1, 1), 并加入批量归一化 (BN) 层和 Mish 激活函数。 $1 \times 1 \times 3$ 卷积核不会提取任何空间特征, 因为它没有考虑空间相邻像素间的相关性, 但是它可以很好地保留空间特征并提取光谱特征。然后采用 $1 \times 1 \times b$ 卷积核对经过光谱密集连接块的输出特征图进行卷积, 生成大小为 $7 \times 7 \times 1$ 、通道数为 60 的光谱特征图。光谱密集连接块输出的特征图结合光谱 ECA 机制和 FFT 的光谱注意力块, 同时也对光谱密集连接块输出的特征图进行 FFT, 再对两者输出的特征图进行融合。为了克服过拟合, 加入了 BN 层, Mish 激活层和 dropout 层来增强数值稳定性, 参数 p_1 设置为 0.5。最后进

行自适应平均池化, 得到大小为 1×60 的光谱特征向量。

3.2 空间分支结构

空间分支结构通过 24 个大小为 $1 \times 1 \times B$ 的卷积核, 对尺寸为 $7 \times 7 \times B$ 的输入数据进行降维, 降维后得到 24 个大小为 $7 \times 7 \times 1$ 的空间特征图。空间密集连接块卷积层采用 $3 \times 3 \times 1$ 卷积核, 卷积核个数设置为 12, 步长设置为 (1, 1, 1), 并加入 BN 层和采用 Mish 激活函数的激活层, 生成 60 个大小为 $7 \times 7 \times 1$ 的空间特征图。和光谱分支一样, 空间分支也加入了结合空间 ECA 机制和 FFT 的空间注意力块, 同时也对空间密集连接块输出的特征图进行 FFT, 最后进行特征融合。空间分支也加入了 BN 层, Mish 激活层和 dropout 层防止过拟合, 参数 p_1 设置为 0.5。最后进行自适应平均池化, 得到大小为 1×60 的空间特征向量。

3.3 光谱分支和空间分支特征融合

为了更好地地区分光谱域特征和空间域特征, 所提方法最后采用 contact 操作对光谱分支和空间分

支输出的特征向量进行特征融合,直接对两个特征向量进行 concat 操作,特征向量维数变为 120,最后通过一个全连接层和一个线性层完成地物的分类。

4 实验分析与讨论

采用三个公共高光谱数据对所提算法进行了实验验证,通过分析比较各算法实际检测输出的分类图像,OA,AA,Kappa 系数,训练时间和测试时间,来检验所提算法的有效性。本实验所有算法程序使用 Python3.7 搭载 PyTorch 1.8.0 深度学习框架实现,实验设备为 64 位 Windows 10 服务器。配置信息如下:CPU 为 Inter(R) Core(TM) CPU I9-10940X 3.30 GHz, GPU 为 16G NVIDIA Quadro RTX 5000,内存为 256 GB。

4.1 数据集介绍

采用三个公共 HSI 数据集,分别是 Indiana Pines (IN), The University of Pavia (UP), Botswana (BS) 高光谱数据集。

IN 数据集由机载可见红外成像光谱仪 (AVIRIS) 传感器在 Indiana 西北部采集,大小为 $145 \times 145 \times 200$,包含 145×145 个像素,16 个地物类别,去除吸水区域剩余 200 个波长范围为 $0.4 \sim 2.5 \mu\text{m}$ 的波段,空间分辨率为 20 m/pixel。

UP 数据集由反射光学系统成像光谱仪 (ROSIS-

3) 在 Italy 北部的 Pavia University 采集,大小为 $610 \times 340 \times 103$,包含 610×340 像素,9 个地物类别,去除受噪声影响的波段,剩余 103 个波长范围从 $0.43 \sim 0.86 \mu\text{m}$ 的波段,空间分辨率为 1.3 m/pixel。

BS 数据集由 NASA 的 EO-1 号卫星在 Botswana 的 Okavango 三角洲上空采集,大小为 $1476 \times 256 \times 145$,包含 1476×256 个像素,14 个地物类别,去除未校准、吸水区域和噪声波段,剩余 145 个波长范围为 $0.4 \sim 2.5 \mu\text{m}$ 的波段,空间分辨率为 30 m/pixel。

4.2 数据集划分

基于深度学习的 HSI 分类的性能一直局限于训练样本的不足,有监督学习需要大量样本进行训练,训练样本越多,分类精度则相对越高,但与此同时也会增加训练和测试的时间成本。为了验证在训练样本缺乏的情况下仍可以保证所提算法的有效性,并减少训练时间,在保证分类精度的情况下,将各个数据集的训练样本的划分比例设置为最小,而对于不同数据集划分比例的选择,需要根据数据集的具体大小来决定。对于较大的数据集,划分相对较少的训练样本即可获得较高的分类精度,而对于较小的数据集,则需要划分相对更多的训练样本才可获得较高的分类精度。实验中各数据集的划分情况如表 1 所示。

表 1 各数据集划分情况
Table 1 The ratio of each data set

Dataset	Class number	Training/Validation/Testing set ratio	Total number	Train	Validation	Test
IN	16	3%/3%/97%	10249	307	307	9635
UP	9	0.5%/0.5%/99.5%	42776	210	210	42356
BS	14	2.5%/2.5%/97.5%	3148	76	76	2996

4.3 参数设置

选择 SVM、CDCNN、SSRN、FDSSC、DBMA、DBDA 这几个算法作为对比,SVM 将原始立方体数据作为输入数据,其他算法将目标像素及其相邻像素组成的小立方体作为输入数据。各算法为了保证分类精度的同时尽量提高训练速度,采用的原始输入数据大小有所差异,可以用更短的训练时间得到对大小合适的输入数据更好的分类性能。所提算法在各数据集上进行了实验,得到了模型在不同输入数据大小 ($3 \times 3 \times L \sim 13 \times 13 \times L$) 情况下的分类总体精度,实验结果如图 7 所示。由图 7 可知,继续增大模型输入数据大小的同时,模型的分精度的增长幅度趋于平缓,甚至开始下降,而且大大

增加了训练和测试的时间成本,因此综合考虑分类精度和时间成本,最终选择 $7 \times 7 \times L$ 大小的小立方体数据作为所提模型的输入数据。各算法的数据输入尺寸详细信息如表 2 所示。

算法优化器选择 Adam;为了增强模型的泛化能力,损失函数采用标签平滑的交叉熵损失函数;采用动态学习率调整策略,学习率衰减策略采用余弦退火;为了减少训练时间,采用早期停止策略,早期停止参数设置为 20 个 epoch,如果验证集的精度在 20 个 epoch 内不再上升,那么模型将自动终止训练。

4.4 分析与讨论

在 IN,UP,BS 高光谱数据集上进行实验,采用训练时间,测试时间,OA,AA 和 Kappa 系数这几个

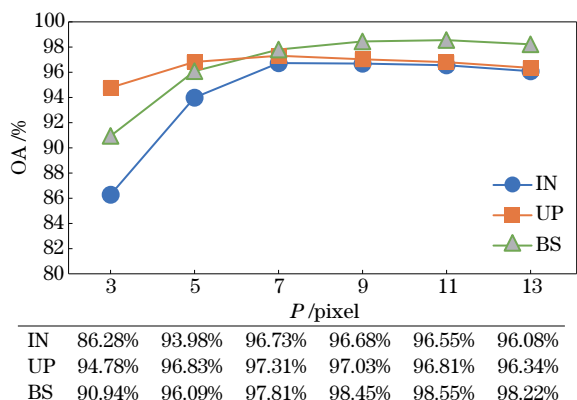


图 7 所提模型在各数据集中采用不同输入数据大小时的总体分类精度

Fig. 7 Overall classification accuracy of the proposed model with different input data sizes in each dataset

表 2 各算法参数

Table 2 Parameters of each algorithm

Algorithm	Basic framework	Input size	Batch size	Learning rate
SVM	SVM+RBF	$H \times W \times L$		
CDCNN	2D-CNN+ResNet	$5 \times 5 \times L$		
SSRN	3D-CNN+ResNet	$7 \times 7 \times L$		
FDSSC	3D-CNN+DenseNet	$9 \times 9 \times L$	64	0.0005
DBMA	3D-CNN+DenseNet+Attention	$7 \times 7 \times L$		
DBDA	3D-CNN+DenseNet+Attention	$9 \times 9 \times L$		
Proposed algorithm	3D-CNN+DenseNet+Attention+FFT	$7 \times 7 \times L$		

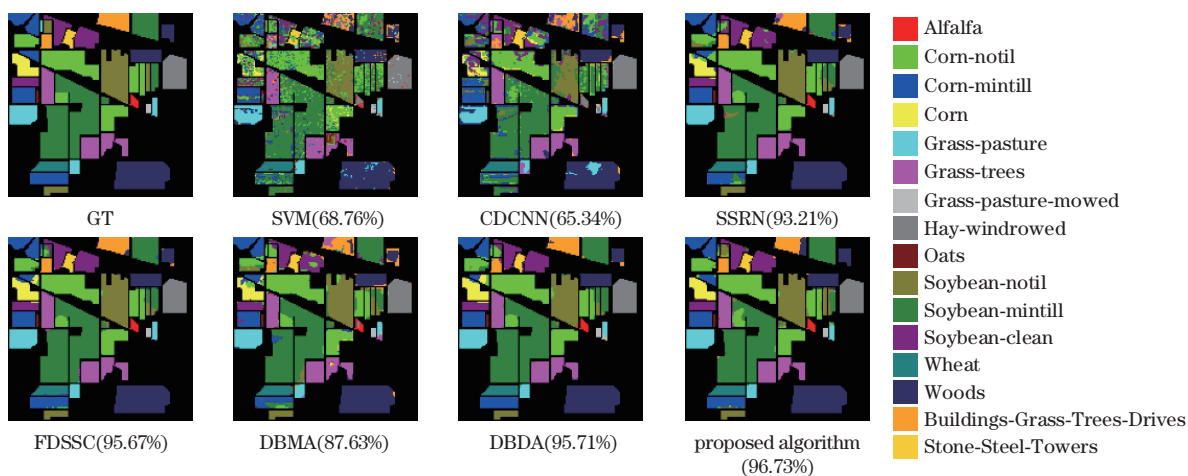


图 8 各算法在 IN 数据集上的分类图

Fig. 8 Classification maps of each algorithm on IN dataset

表 3 为各算法在 IN 数据集上选取 3% 训练样本后的分类结果。从表 3 可知,所提算法在 IN 数据集的五类地物上获得了最佳精度,分别是第 4, 10, 11, 12, 15 类, OA, AA 和 Kappa 系数都高于其他算法,分别达到了 96.73%, 96.18% 和 96.27%。训练时

性能指标进行验证,并在表中用粗体显示最佳性能指标。对各算法和所提算法在各数据集上的实验结果进行分析讨论,来验证所提算法的有效性。

















4.4.1 各算法在 IN 数据集上的分类结果和分类图

图 8 为各算法在 IN 数据集上的分类结果。和其他算法相比,所提算法在 IN 数据集上的总体分类效果较好,总体分类精度达到最高,为 96.73%,可以很好地识别出各类地物,背景干净;对少数地物识别效果不佳,如第 7 类和第 9 类地物的分类精度分别只有 89.85% 和 84.45%。其次是 DBDA 算法和 FDSSC 算法,总体分类精度分别达到 95.71% 和 95.67%。SVM 和 CDCNN 算法总体上分类效果较差,总体分类精度分别只有 68.76% 和 65.34%。

间和测试时间最短的是 CDCNN 算法,分别只需要 7.79 s 和 0.59 s,但是 CDCNN 算法的分类精度和 Kappa 系数太低,分类效果太差。训练时间最长的是 DBDA 算法,分别需要 97.23 s 和 5.45 s。所提算法的训练速度和测试速度排在第三,训练时间和测

表 3 各算法在 IN 数据集上选取 3% 训练样本后的分类结果

Table 3 Classification results of each algorithm selecting the 3% training samples from IN dataset

Class	Color	SVM	CDCNN	SSRN	FDSSC	DMBA	DBDA	Proposed algorithm
1		36.62	29.59	91.99	98.96	82.81	97.54	97.74
2		55.49	60.24	91.96	95.64	85.76	93.68	92.75
3		62.33	42.55	90.01	93.90	85.83	94.49	91.82
4		42.54	46.65	93.35	94.69	85.84	95.67	99.49
5		85.05	88.73	99.11	98.74	94.56	99.51	99.27
6		83.32	81.49	97.46	97.58	95.73	98.73	98.42
7		59.87	30.40	87.79	79.10	66.09	68.86	89.85
8		89.67	85.70	99.31	99.45	98.29	100.00	99.96
9		39.28	58.56	88.32	87.88	37.59	79.92	84.45
10		62.32	55.97	82.97	91.77	81.55	93.49	97.31
11		64.73	65.56	94.22	96.42	87.40	96.24	98.17
12		50.55	36.33	92.85	93.69	82.68	92.84	96.91
13		86.74	82.15	99.11	99.38	97.06	98.89	97.77
14		88.67	79.01	96.90	97.59	95.90	97.98	97.58
15		61.82	68.74	94.77	94.32	79.95	95.77	99.02
16		98.66	89.95	96.88	98.87	95.74	97.52	98.32
OA / %		68.76	65.34	93.21	95.67	87.63	95.71	96.73
AA / %		66.73	62.60	93.56	94.87	84.55	93.82	96.18
Kappa / %		63.98	60.01	92.26	95.06	85.88	95.11	96.27
Train_time / s		14.28	7.79	29.95	54.03	33.09	97.23	25.26
Test_time / s		1.42	0.59	2.58	3.08	3.65	5.45	1.69

试时间分别只需要 25.26 s 和 1.69 s, 具有很高的可行性。

4.4.2 各算法在 UP 数据集上的分类结果和分类图

图 9 为各算法在 UP 数据集上的分类结果。从图 9 可知: SVM 和 CDCNN 算法在 UP 数据集上的分类效果较差; DBDA 算法, FDSSC 算法和所提算法分类效果都较好, OA 都达到了 95% 以上。

表 4 为各算法在 UP 数据集上选取 0.5% 训练样本后的分类结果。从表 4 可知, UP 数据集共有 9 类地物, 所提算法在 UP 数据集上的分类表现优越, 9 类地物中有 7 类地物的分类精度达到最高, 而且 OA, AA 和 Kappa 系数也达到了最高, 分别为 97.31%, 97.12% 和 96.43%, 训练时间和测试时间排在第三, 分别为 10.93 s 和 4.75 s。虽然 DBDA 算法分类精度高, 但是训练时间和测试时间最长, 分别为 24.01 s 和 13.30 s。训练时间和测试时间最短的算法分别是 SVM 和 CDCNN, 分别只需要 3.75 s

和 2.57 s, 但两者的分类效果都较差。

4.4.3 各算法在 BS 数据集上的分类结果和分类图

图 10 为各算法在 BS 数据集上的分类结果。由图 10 可知, 所提算法的分类效果在 BS 数据集上表现优越, 可以很好地区分各类地物, OA 达到 97.81%; SVM 和 CDCNN 算法分类效果一般; 其他算法分类效果都较好, 总体分类精度都在 95% 左右。

表 5 为各算法在 BS 数据集上选取 2.5% 训练样本后的分类结果。从表 5 可知, BS 数据集共有 14 类地物, 所提算法在 BS 数据集上的分类表现优越, 14 类地物中有 9 类地物的分类精度达到最高, 第 1, 2, 3, 7, 11, 14 类地物的分类精度达到了 100%, 第 4, 9, 10 类地物的分类精度分别达到了 99.14%, 98.13%, 99.96%; 而且 OA, AA 和 Kappa 系数也达到了最高, 分别达到了 97.81%, 98.14% 和 97.63%; 训练时间和测试时间分别为 13.00 s 和 0.69 s, 排在第四。其次是 DBDA 算法的 OA, AA

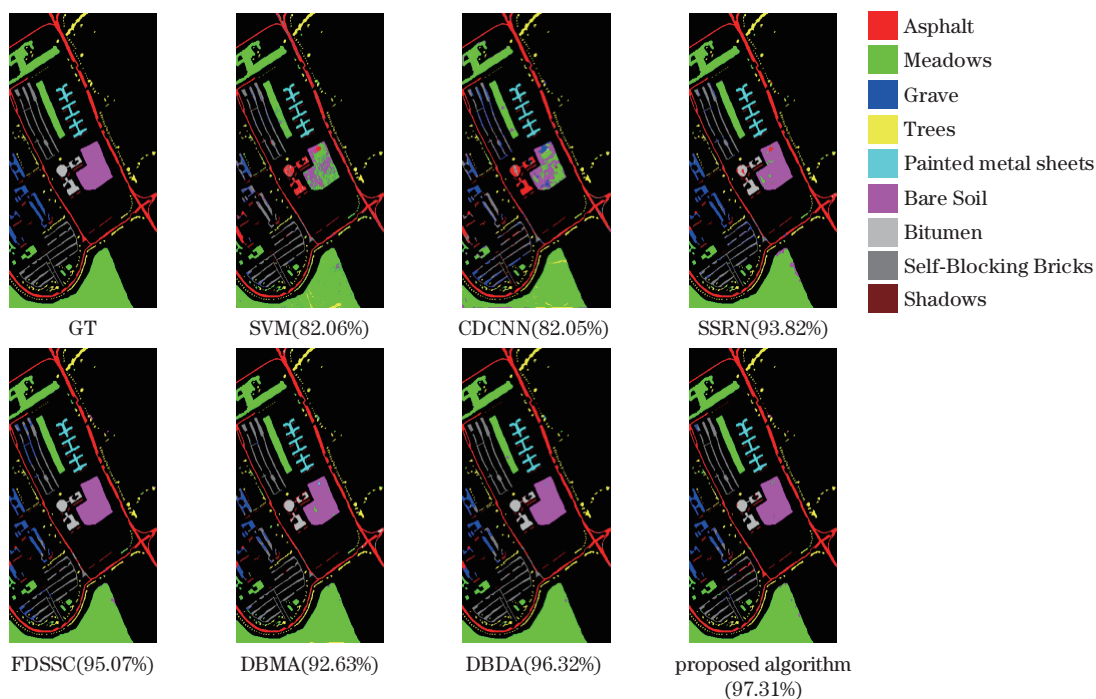


图9 各算法在UP数据集上的分类结果图

Fig. 9 Classification maps of each algorithm on UP dataset

表4 各算法在UP数据集上选取0.5%训练样本后的分类结果

Table 4 Classification results of each algorithm selecting the 0.5% training samples from UP dataset

Class	Color	SVM	CDCNN	SSRN	FDSSC	DMBA	DBDA	Proposed algorithm
1	■	81.26	80.54	93.17	91.74	93.13	93.20	95.75
2	■	84.53	87.37	96.94	99.06	96.63	99.02	99.33
3	■	56.56	26.31	75.11	85.39	76.90	96.41	97.51
4	■	94.34	95.11	99.15	99.42	96.91	98.17	99.56
5	■	95.38	96.98	99.51	99.29	99.21	98.57	99.67
6	■	80.66	73.86	95.13	97.20	94.73	98.13	98.81
7	■	49.13	46.70	93.22	94.62	84.29	98.56	98.69
8	■	71.16	63.52	91.35	84.91	83.84	85.78	85.62
9	■	99.94	92.10	99.26	98.88	97.30	98.47	99.10
OA / %		82.06	82.05	93.82	95.07	92.63	96.32	97.31
AA / %		79.22	73.61	93.65	94.50	91.44	96.25	97.12
Kappa / %		75.44	75.53	91.78	93.45	90.19	95.11	96.43
Train_time / s		3.75	7.39	14.07	22.34	16.88	24.01	10.93
Test_time / s		3.25	2.57	6.35	7.62	9.21	13.30	4.75

和 Kappa 系数最高,但训练时间和测试时间都是最长的,分别为 21.57 s 和 1.36 s。训练时间最短的是 SVM 算法,只需要 1.69 s。CDCNN 算法测试时间最短,只需要 0.21 s,但分类精度太低。

综上所述,各算法对 HSI 的分类都存在一些相同问题,例如对于一些光谱特征相近的地物,容易错误分类,对于一些地物的边缘特征,也难以区分。所提

算法在各个数据集上都表现出优越的分类性能,在各数据集上的 OA, AA 和 Kappa 指数都达到了最高,在训练时间和测试时间上也表现出很高的可行性。

4.4.4 模型参数量和计算力分析

模型参数量 (Parameters) 可以衡量一个模型的大小,模型计算力 (FLOPs) 表示模型在实际计算过程中的计算次数,可以衡量一个模型的复杂度,两

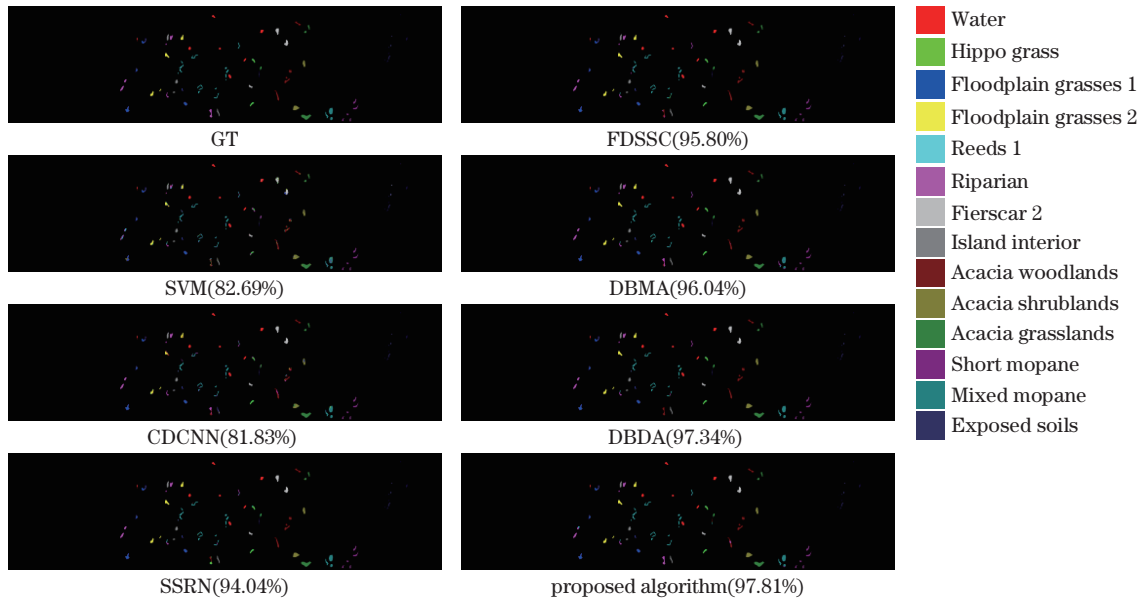


图 10 各算法在 BS 数据集上的分类结果图

Fig. 10 Classification maps of each algorithm on BS dataset

表 5 各算法在 BS 数据集上选取 2.5% 训练样本后的分类结果

Table 5 Classification results of each algorithm selecting the 2.5% training samples from BS dataset

Class	Color	SVM	CDCNN	SSRN	FDSSC	DMBA	DBDA	Proposed algorithm
1		99.09	91.82	98.11	98.59	99.46	98.97	100.00
2		83.72	60.46	93.79	98.57	93.90	99.90	100.00
3		85.58	90.46	96.56	99.86	98.81	100.00	100.00
4		80.98	89.82	96.26	93.63	91.02	94.78	99.14
5		78.45	83.25	87.70	87.85	91.81	89.61	89.54
6		60.32	67.48	86.80	89.02	90.83	95.07	91.04
7		92.87	96.83	99.65	99.96	98.44	100.00	100.00
8		87.24	81.41	96.13	97.73	96.78	99.95	98.32
9		75.25	67.70	95.06	97.47	95.84	96.85	98.13
10		71.77	83.64	90.95	96.09	95.88	97.28	99.96
11		92.14	81.01	99.97	99.66	99.46	99.35	100.00
12		91.95	80.70	99.22	98.55	96.75	99.94	98.37
13		82.91	95.59	93.43	94.91	99.26	98.01	99.50
14		94.95	78.62	100.00	99.78	99.66	99.67	100.00
OA / %		82.69	81.83	94.04	95.80	96.04	97.34	97.81
AA / %		84.09	82.06	95.26	96.55	96.28	97.81	98.14
Kappa / %		81.25	80.29	93.54	95.45	95.71	97.12	97.63
Train_time / s		1.69	6.46	10.95	13.64	13.75	21.57	13.00
Test_time / s		0.22	0.21	0.64	0.77	0.93	1.36	0.69

者反映了模型的可行性。通过实验计算了所提算法和其他对比的深度学习模型在各个数据集上的参数量和计算力。由表 6 可知,所提深度学习模型

参数量达到了最小,计算力仅次于 CDCNN 算法,但是所提算法的分类精度远高于 CDCNN 算法,因此表明所提算法可行性高,模型易于部署。

表 6 各深度学习模型的参数量和计算力

Table 6 Number of parameters and computing force of each deep learning model

	Dataset	CDCNN	SSRN	FDSSC	DBMA	DBDA	Proposed algorithm
IN	Parameters/ 10^6	1.06	0.36	1.23	0.61	0.38	0.18
	FLOPs/ 10^9	0.23	6.08	11.16	9.39	6.76	1.17
UP	Parameters/ 10^6	0.63	0.22	0.65	0.32	0.20	0.09
	FLOPs/ 10^9	0.12	3.11	5.77	5.61	4.34	0.54
BS	Parameters/ 10^6	0.82	0.28	0.90	0.44	0.28	0.24
	FLOPs/ 10^9	0.17	4.41	8.13	8.02	6.18	1.55

5 结 论

提出一种以 3D-CNN 为框架, 基于双分支快速密集连接和双注意力机制的 HSI 分类算法。采用 PCA 和 3D 卷积来减小高光谱数据的冗余信息并提取显著特征; 在 3D-CNN 中采用 Mish 激活函数, 提高了模型的泛化性和鲁棒性; 同时引入轻量 ECA 机制和 FFT 的注意力模块, 并增加一个 FFT 层来替换一个卷积层, 和注意力模块形成密集连接, 保证精度的同时也加快了速度, 提高了模型稳定性。在三个公共高光谱数据集上对不同算法进行了实验验证, 实验结果表明了所提算法在 HSI 分类中的优越性和可行性。

参 考 文 献

- [1] 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.
- [2] Zhang H Y, Li J Y, Huang Y C, et al. A nonlocal weighted joint sparse representation classification method for hyperspectral imagery[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7(6): 2056-2065.
- [3] Dong A G, Li J X, Zhang B, et al. Hyperspectral image classification algorithm based on spectral clustering and sparse representation[J]. Acta Optica Sinica, 2017, 37(8): 0828005.
董安国, 李佳逊, 张蓓, 等. 基于谱聚类和稀疏表示的高光谱图像分类算法[J]. 光学学报, 2017, 37(8): 0828005.
- [4] Li J Y, Zhang H Y, Huang Y C, et al. Hyperspectral image classification by nonlocal joint collaborative representation with a locally adaptive dictionary[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(6): 3707-3719.
- [5] Chen Y S, Zhao X, 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.

- [6] Hu W, Huang Y Y, Wei L, et al. Deep convolutional neural networks for hyperspectral image classification[J]. Journal of Sensors, 2015, 2015: 1-12.
- [7] Huang H, He K, Zheng X L, et al. Spatial-spectral feature extraction of hyperspectral image based on deep learning[J]. Laser & Optoelectronics Progress, 2017, 54(10): 101001.
黄鸿, 何凯, 郑新磊, 等. 基于深度学习的高光谱图像空-谱联合特征提取[J]. 激光与光电子学进展, 2017, 54(10): 101001.
- [8] Yue J, Zhao W Z, Mao S J, et al. Spectral-spatial classification of hyperspectral images using deep convolutional neural networks[J]. Remote Sensing Letters, 2015, 6(6): 468-477.
- [9] Chen Y S, 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.
- [10] Li Y, Zhang H K, Shen Q. Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network[J]. Remote Sensing, 2017, 9(1): 67.
- [11] Luo Y N, Zou J, Yao C F, et al. HSI-CNN: a novel convolution neural network for hyperspectral image [C]//2018 International Conference on Audio, Language and Image Processing (ICALIP), July 16-17, 2018, Shanghai, China. New York: IEEE Press, 2018: 464-469.
- [12] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas,

- NV, USA. New York: IEEE Press, 2016: 770-778.
- [13] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [14] 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.
- [15] Wang W J, Dou S G, Jiang Z M, et al. A fast dense spectral-spatial convolution network framework for hyperspectral images classification[J]. Remote Sensing, 2018, 10(7): 1068.
- [16] Ma W P, Yang Q F, Wu Y, et al. Double-branch multi-attention mechanism network for hyperspectral image classification[J]. Remote Sensing, 2019, 11(11): 1307.
- [17] Li R, Zheng S Y, Duan C X, et al. Classification of hyperspectral image based on double-branch dual-attention mechanism network[J]. Remote Sensing, 2020, 12(3): 582.
- [18] Mishra D. Mish: a self regularized non-monotonic neural activation function[EB/OL]. (2019-08-23) [2021-05-01]. <https://arxiv.org/abs/1908.08681v3>.
- [19] Lee-Thorp J, Ainslie J, Eckstein I, et al. FNet: mixing tokens with Fourier transforms[EB/OL]. (2021-04-09) [2021-06-01]. <https://arxiv.org/abs/2105.03824>.