

基于轻量级全连接张量映射网络的高光谱 图像分类方法

林知心, 郑玉棒*, 马天宇, 王蕊, 李恒超

(西南交通大学信息科学与技术学院, 四川成都 611756)

摘要: 近年来, 基于卷积神经网络的深度学习模型已经在高光谱图像分类领域取得优异表现. 然而, 模型性能的提升通常依赖于更深、更宽的网络结构, 导致参数量和计算量增长, 从而限制了模型在机载或星载载荷中的实际部署. 为此, 本文提出基于轻量级全连接张量映射网络的高光谱图像分类方法. 根据全连接张量网络分解的映射思想以及高光谱图像“图谱合一”的结构特点, 本文设计两种张量映射卷积单元, 通过使用多个具有全连接结构的小尺寸卷积核代替原始卷积核, 降低了卷积层的时间和空间复杂度. 此外, 基于新单元构建残差双分支张量模块. 双分支结构共享同一组权重参数, 并采用通道分割操作减少特征通道数, 提升特征提取过程的实时性. 本文所提模型通过使用新单元和新模块充分挖掘高光谱图像的局部空谱信息和全局光谱信息, 有效提高了分类性能并减少硬件资源消耗. 在三个常用高光谱图像数据集上的实验结果表明, 所提模型相较于其他现有工作具有更高的分类性能以及更低的参数量和计算量.

关键词: 高光谱图像分类; 模型压缩; 全连接张量网络分解; 卷积神经网络; 张量神经网络; 轻量卷积模块

基金项目: 国家自然科学基金 (No.62271418)

中图分类号: TP751.1

文献标识码: A

文章编号: 0372-2112(2024)10-3541-11

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20240477

Lightweight Fully-Connected Tensorial Mapping Network for Hyperspectral Image Classification

LIN Zhi-xin, ZHENG Yu-bang*, MA Tian-yu, WANG Rui, LI Heng-chao

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu, Sichuan 611756, China)

Abstract: In recent years, convolutional neural networks have demonstrated outstanding performance in HSIC (Hyperspectral Image Classification). However, the improvement of model performance involves adopting deeper and broader network architectures, leading to an increased number of parameters and operations, thus hindering deployment in airborne or on-board devices. To this end, this paper introduces a HSIC method based on the LiteFCTMN (Lightweight Fully-Connected Tensorial Mapping Network). We design two convolutional units based on the mapping way of FCTN (Fully-Connected Tensor Network) decomposition and the structural characteristics of HSIs. By mapping the original convolution kernel to multiple small-sized convolution kernels with fully-connected structures, the complexity of the novel units is reduced while their expressiveness is improved. In addition, the RDT (Residual Double-Branch Tensorial) module is constructed using the designed units. In this module, two branches share the same weights, and a channel split operation is employed to reduce the number of feature channels, thereby reducing complexity. The proposed model strategically leverages both local spatial-spectral information from RDT and global spectral information from the new units, resulting in enhanced classification performance and reduced hardware consumption. Experimental results on three widely used HSI datasets demonstrate that the proposed model achieves superior classification performance and lower complexity compared to the state-of-the-art works.

Key words: hyperspectral image classification; model compression; fully-connected tensor network decomposition; convolutional neural network; tensorial neural network; lightweight convolutional module

Foundation Item(s): National Natural Science Foundation of China (No.62271418)

1 引言

高光谱图像具有波段数目多、光谱分辨率高、信息丰富等特点,在目标检测和精确分类中占据得天独厚的优势,已被广泛应用于精准农业^[1]、生物医学^[2]和城市规划^[3]等多个方面^[4,5].高光谱图像分类作为其中的关键步骤,通过提取不同地物目标特有的判别性空谱特征来鉴别类别属性,具有重要的学术研究价值和实际应用需求.

近年来,卷积神经网络(Convolutional Neural Network, CNN)已经成为高光谱图像分类领域重要的技术手段.但是,这类模型通常需要大规模的训练参数和计算量,增加了存储和计算资源需求,导致模型的实际应用范围受限.为解决上述问题,在文献[6]中,作者提出了基于DenseNet网络^[7]的轻量级模型LiteDenseNet.该模型采用双分支密集结构提取高光谱图像的特征信息并通过引入分组卷积^[8]降低了CNN模型的计算成本和参数规模.针对文献[6]中使用分组卷积导致通道间信息相关性丢失的问题,Cui等^[9]通过移除激活函数和批归一化层(Batch Normalization, BN)^[10]构建了改进的三维深度可分离卷积结构来代替分组卷积,从而搭建复杂度更低的LiteDepthwiseNet模型.虽然上述工作可以一定程度降低模型复杂度,但是,这类设计紧凑结构的压缩方法非常依赖于专家经验并且模型的整体压缩率有限^[11].

相比之下,基于张量分解的模型压缩方法具有线性代数理论作为支撑.其根据低秩近似的思想,用多个小尺寸张量代替标准全连接层和卷积层的权重,降低压缩网络的分类精度损失^[12].在一些研究中,为了深入探索数据的低秩拓扑结构,研究者先将全连接层的权重矩阵重构为高阶张量,再将其分解为因子张量,并定义该方法为张量化法.依托该方法,Hu等^[13]基于张量链(Tensor Train, TT)分解^[14]设计了卷积长短时记忆单元的张量扩展结构提取高光谱图像的空谱特征.Ma等^[15]设计了基于张量环(Tensor Ring, TR)分解^[16]的卷积层并构建张量神经网络,实现了低复杂度的高光谱图像分类.关注到全连接张量网络(Fully-Connected Tensor Network, FCTN)^[17]分解解决了TT和TR分解中因子张量相关性度量能力有限和排列顺序敏感等问题,文献[18]首次将FCTN分解应用于模型压缩领域并在高光谱图像分类任务中取得了出色的分类性能和压缩效果.然而,张量化方法在实施过程中需要引入额外的计算步骤恢复原始卷积核,导致模型的整体计算量增长.

另一种基于张量分解的压缩方法(映射法)通过借鉴网络权重的数据结构,将各个因子张量直接视为不同层次的卷积核权重,避免了额外的计算步骤,能够同时降低模型的参数量和计算量^[19-21].文献[19]基于塔克(tucker)分解的映射思想将标准卷积重新设计为瓶颈卷积模块,文献[20]利用正则多元(Canonical Polyadic, CP)分解将标准卷积重构为可分离瓶颈卷积模块,文献[21]则是借助TT分解进一步拆分了瓶颈卷积模块.受此启发,本文使用表达能力更强的FCTN分解,从张量映射角度出发.首先,根据高光谱图像“图谱合一”的数据结构特点,设计两种基于FCTN分解的三维映射卷积单元,使用多个具有全连接结构的小尺寸因子张量代替原始卷积核张量,将单层高复杂度的卷积运算转化为多层低复杂度的卷积运算.进一步,融合新单元和点卷积设计三维残差双分支张量模块,其中,双分支结构共享同一组权重参数,并采用通道分割操作减少特征通道数以实现低复杂度的空谱特征提取.最后,构建轻量级全连接张量映射网络(Lightweight Fully-Connected Tensorial Mapping Network, LiteFCTMN),通过融合残差双分支张量模块提取的局部空谱特征和新单元提取的全局光谱特征,模型具有优异的高光谱图像分类性能.

2 理论基础

2.1 张量基础

张量是矩阵向任意维度的推广,张量图是张量的图形表示,用于可视化张量和张量运算.图1为张量图表示,如图所示,从节点延伸出的边的个数即为张量的阶数,每条边代表一个维度.

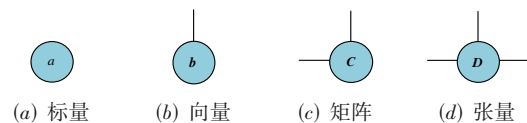


图1 张量图表示

张量运算是向量或矩阵运算的高维扩展.图2为两种基本张量运算的张量图表示,其中,运算类型由连接节点的线的形状确定,实线表示张量缩并运算,虚线表示张量卷积运算.

2.2 全连接张量网络分解

如图3(c)所示,FCTN分解将高阶张量分解为任意两个因子张量相互连接的形式.给定一个 d 阶张量 $\mathbf{X} \in \mathbb{R}^{L_1 \times L_2 \times \dots \times L_d}$,其FCTN分解的公式为

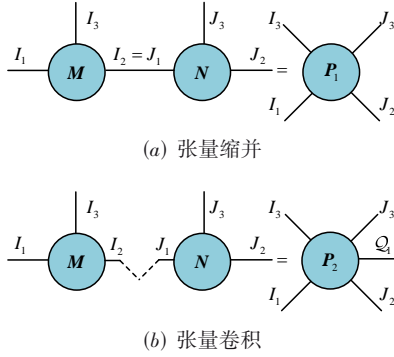


图2 张量运算

$$\begin{aligned}
 X(l_1, \dots, l_d) = & \sum_{r_{1,2}=1}^{R_{1,2}} \dots \sum_{r_{1,d}=1}^{R_{1,d}} \sum_{r_{2,3}=1}^{R_{2,3}} \dots \sum_{r_{d-1,d}=1}^{R_{d-1,d}} \\
 & \{ \mathbf{G}_1(l_1, r_{1,2}, \dots, r_{1,d}) \mathbf{G}_2(r_{1,2}, l_2, \dots, r_{2,d}) \dots \\
 & \mathbf{G}_k(r_{1,k}, \dots, l_k, \dots, r_{k,d}) \dots \mathbf{G}_d(r_{1,d}, \dots, r_{d-1,d}, l_d) \}
 \end{aligned} \quad (1)$$

其中, $\mathbf{G}_k \in \mathbb{R}^{R_{1,k} \times \dots \times R_{k-1,k} \times L_k \times R_{k,k+1} \times \dots \times R_{k,d}}$, $k \in [1, d]$, R_{k_1, k_2} ($1 \leq k_1 < k_2 \leq d, k_1, k_2 \in \mathbb{N}^+$) 是 FCTN 分解的秩. 通过全连接结构, FCTN 分解解决了 TT 分解和 TR 分解无法充分刻画高阶张量多维度之间相关性的问题. 同时, 这种分解方法具有转置不变性, 有利于提高分解形式的灵活性^[17].

3 轻量级全连接张量映射网络高光谱图像分类模型

3.1 基于 FCTN 分解的三维映射卷积层

3.1.1 FCTNMConv3D4 设计

一般情况下, 标准三维卷积层 (Conv3D) 使用大小为 $L \times L \times T \times I \times O$ 的五阶卷积核张量 \mathbf{K}_{3D} 将四阶输入张量 $\mathbf{X} \in \mathbb{R}^{H \times W \times D \times I}$ 映射至四阶输出张量 $\mathbf{Y} \in \mathbb{R}^{H' \times W' \times D' \times O}$, 其数学表达式为

$$\begin{aligned}
 & \mathbf{Y}(h', w', d', o) \\
 = & \sum_{l_1, l_2=1}^L \sum_{t=1}^T \sum_{i=1}^I \{ \mathbf{K}_{3D}(l_1, l_2, t, i, o) \mathbf{X}(h, w, d, i) \}
 \end{aligned} \quad (2)$$

其中, L 和 T 分别表示 \mathbf{K}_{3D} 的空间和光谱维度, I 和 O 是输入和输出通道数, H, W 和 D 为输入张量的高度, 宽度和长度, H', W' 和 D' 表示输出张量的高度, 宽度以及长度.

借助 FCTN 分解的映射思想, 将式 (2) 的运算过程转化为多步骤小尺寸卷积计算, 降低标准卷积层的复杂度. 对 \mathbf{K}_{3D} 直接进行 FCTN 分解, 可以得到五个具有全连接拓扑结构的小尺寸卷积核张量. 此时考虑图像的空间相关性, 将空间维度保留在同一个卷积核张量中, \mathbf{K}_{3D} 的 FCTN 分解为

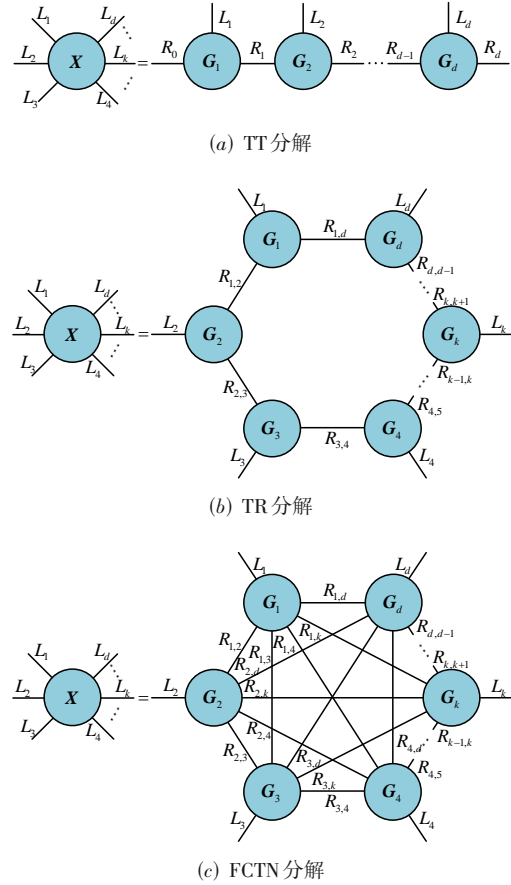


图3 张量分解

$$\begin{aligned}
 \mathbf{K}_{3D}(l_1, l_2, t, i, o) = & \sum_{r_{1,2}=1}^{R_{1,2}} \sum_{r_{1,3}=1}^{R_{1,3}} \sum_{r_{1,4}=1}^{R_{1,4}} \sum_{r_{2,3}=1}^{R_{2,3}} \sum_{r_{2,4}=1}^{R_{2,4}} \sum_{r_{3,4}=1}^{R_{3,4}} \\
 & \{ \mathbf{S}_1(l_1, l_2, r_{1,2}, r_{1,3}, r_{1,4}) \mathbf{S}_2(r_{1,2}, t, r_{2,3}, r_{2,4}) \\
 & \mathbf{G}_1(r_{1,3}, r_{2,3}, i, r_{3,4}) \mathbf{G}_2(r_{1,4}, r_{2,4}, r_{3,4}, o) \}
 \end{aligned} \quad (3)$$

其中, $\mathbf{S}_1 \in \mathbb{R}^{L \times L \times R_{1,2} \times R_{1,3} \times R_{1,4}}$ 和 $\mathbf{S}_2 \in \mathbb{R}^{R_{1,2} \times T \times R_{2,3} \times R_{2,4}}$ 是 \mathbf{K}_{3D} 分解得到的空间因子张量和光谱因子张量, $\mathbf{G}_1 \in \mathbb{R}^{R_{1,3} \times R_{2,3} \times I \times R_{3,4}}$ 和 $\mathbf{G}_2 \in \mathbb{R}^{R_{1,4} \times R_{2,4} \times R_{3,4} \times O}$ 分别为输入、输出通道因子张量. 此时, \mathbf{S}_1 同时保留了 \mathbf{K}_{3D} 的空间维度 $L \times L$, 有利于在特征提取过程中保护高光谱图像的空间结构.

将式 (3) 代入式 (2), 使用具有全连接结构的小尺寸卷积核对原始卷积核进行近似, 可以得到基于 FCTN 分解的三维映射卷积层:

$$\begin{aligned}
 \mathbf{Y}(h', w', d', o) = & \sum_{l_1, l_2=1}^L \sum_{t=1}^T \sum_{i=1}^I \sum_{r_{1,2}=1}^{R_{1,2}} \sum_{r_{1,3}=1}^{R_{1,3}} \sum_{r_{1,4}=1}^{R_{1,4}} \sum_{r_{2,3}=1}^{R_{2,3}} \sum_{r_{2,4}=1}^{R_{2,4}} \sum_{r_{3,4}=1}^{R_{3,4}} \\
 & \{ \mathbf{S}_1(l_1, l_2, r_{1,2}, r_{1,3}, r_{1,4}) \mathbf{S}_2(r_{1,2}, t, r_{2,3}, r_{2,4}) \\
 & \mathbf{G}_1(r_{1,3}, r_{2,3}, i, r_{3,4}) \mathbf{G}_2(r_{1,4}, r_{2,4}, r_{3,4}, o) \mathbf{X}(h, w, d, i) \}
 \end{aligned} \quad (4)$$

将上述单元命名为 FCTNMConv3D4, 其张量图表示如图 4(a) 所示. 相较于 Conv3D, 新单元的参数量压缩比 C_{3D4} 和计算量压缩比 E_{3D4} 为

$$C_{3D4} = \frac{L^2 TIO}{R^3(I+L^2+T+O)} \quad (5)$$

$$E_{3D4} = \frac{L^2 TIOH'W'D'}{R^3(IHWD+L^2H'W'DR^2+TH'W'DR^2+OH'W'D')} \quad (6)$$

其中, 式(5)的分子为 Conv3D 的参数量, 分母为 FCTNMConv3D4 单元的参数量; 式(6)的分子是 Conv3D 的计算量, 分母中, R^3IHWD 为 X 与 G_1 的计算量, $R^5L^2H'W'D$ 和 $R^5TH'W'D'$ 表示 S_1 与 S_2 的计算量, $R^3OH'W'D'$ 为卷积核 G_2 的计算量.

3.1.2 FCTNMConv3D3 设计

在 FCTNMConv3D4 的基础上, 考虑到高光谱图像具有“图谱合一”的数据结构特点. 为了充分挖掘高光谱图像内在信息, 在对 K_{3D} 进行 FCTN 分解时, 将其空间维度和光谱维度保留在同一个小尺寸卷积核中(即合并 FCTNMConv3D4 的 S_1 和 S_2 因子), 以提取高光谱图像的空谱联合特征. 此时, 五阶卷积核张量 K_{3D} 的分解表达式为

$$K_{3D}(l_1, l_2, t, i, o) = \sum_{r_{1,2}=1}^{R_{1,2}} \sum_{r_{1,3}=1}^{R_{1,3}} \sum_{r_{2,3}=1}^{R_{2,3}} \{S(l_1, l_2, t, r_{1,2}, r_{1,3}) G_1(r_{1,2}, i, r_{2,3}) G_2(r_{1,3}, r_{2,3}, o)\} \quad (7)$$

其中, $S \in \mathbb{R}^{L \times L \times T \times R_{1,2} \times R_{1,3}}$ 为 K_{3D} 的空谱因子张量.

将式(7)代入式(2), 使用包含空谱因子张量在内的三个具有全连接结构的小尺寸卷积核替代原始卷积核, 即可得到另一种基于 FCTN 分解的三维映射卷积层, 其数学表达式为

$$Y(h', w', d', o) = \sum_{l_1, l_2=1}^L \sum_{t=1}^T \sum_{i=1}^I \sum_{r_{1,2}=1}^{R_{1,2}} \sum_{r_{1,3}=1}^{R_{1,3}} \sum_{r_{2,3}=1}^{R_{2,3}} \{S(l_1, l_2, t, r_{1,2}, r_{1,3}) G_1(r_{1,2}, i, r_{2,3}) G_2(r_{1,3}, r_{2,3}, o) X(h, w, d, i)\} \quad (8)$$

将上述结构命名为 FCTNMConv3D3 单元. 与 Conv3D 相比, 该单元的参数量压缩比 C_{3D3} 和计算量压缩比 E_{3D3} 计算为

$$C_{3D3} = \frac{L^2 TIO}{R^2(I+L^2T+O)} \quad (9)$$

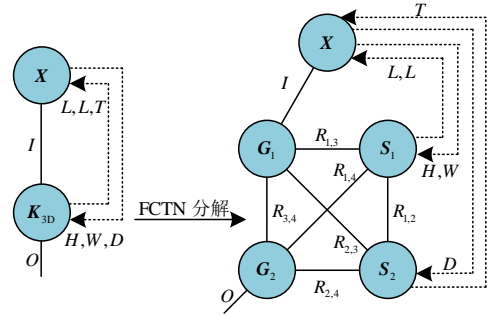
$$E_{3D3} = \frac{L^2 TIOH'W'D'}{R^2(IHWD+L^2TH'W'DR+OH'W'D')} \quad (10)$$

为验证 FCTNMConv 单元的高效性, 取 $L=T=3, I=48, O=12, H=H'=W=W'=9, D=D'=97, R=2$ 比较新单元和 Conv3D 的空间复杂度和计算复杂度. 如表 1 所

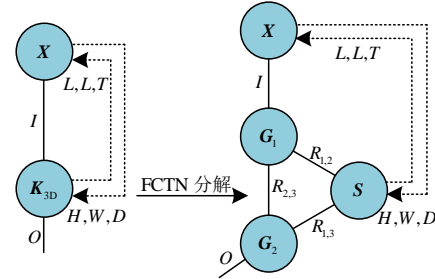
示, FCTNMConv 单元较 Conv3D 显著减小了空间和计算复杂度.

表 1 Conv3D 和 FCTNMConv 的空间和计算复杂度比较

单元	空间复杂度	计算复杂度
Conv3D	15 552	122 192 064
FCTNMConv3D4	576	6 788 448
FCTNMConv3D3	348	3 582 792



(a) FCTNMConv3D4 单元



(b) FCTNMConv3D3 单元

图 4 FCTNMConv 单元

3.2 三维残差双分支张量模块

本节设计的三维残差双分支张量 (Residual Double-branch Tensorial, RDT) 模块受 LiteDenseNet 和 LiteDepthwiseNet 启发, 主要由一个残差结构和一个双分支结构组成, 用于提取高光谱图像的空谱联合特征. LiteDenseNet 的三维双通道密集层使用分组卷积降低模型的复杂度, LiteDepthwiseNet 利用三维深度可分离卷积对该结构进行优化. 基于此, 新设计的 RDT 模块充分发挥 FCTNMConv3D3 单元对高光谱图像的空谱特征提取能力 (单个 FCTNMConv3D3 单元具有多层低复杂度的卷积运算), 减少上分支的特征提取单元, 使上下分支结构一致, 同时共享两分支的权重参数, 进一步压缩模型的时间和空间复杂度.

RDT 模块首先使用通道分割操作将输入数据 $X \in \mathbb{R}^{H \times W \times D \times I}$ 沿通道维度分割为 $X_1 \in \mathbb{R}^{H \times W \times D \times (I/2)}$ 和 $X_2 \in \mathbb{R}^{H \times W \times D \times (I/2)}$, 分别作为输入馈入两个分支. 每个分支由一个点卷积层和一个 FCTNMConv3D3 单元组成. 同时, 双分支结构共享同一组权重参数, 即两个分支是

输入不同,权重相同,结构相同的子网络.此时,双分支的输出特征图分别为

$$\mathbf{B}_1 = \text{FCTNMConv3D3}(\text{PWConv}(\mathbf{X}_1)) \quad (11)$$

$$\mathbf{B}_2 = \text{FCTNMConv3D3}(\text{PWConv}(\mathbf{X}_2)) \quad (12)$$

其中, $\mathbf{B}_1 \in \mathbb{R}^{H \times W \times D \times O}$, $\mathbf{B}_2 \in \mathbb{R}^{H \times W \times D \times O}$ 且 $O=I/2$ 为双分支的输出结果, $\text{PWConv}(\cdot)$ 表示三维点卷积运算.之后,将 \mathbf{B}_1 和 \mathbf{B}_2 沿通道维度进行拼接,得到大小为 $H \times W \times D \times 2O$ 的特征图并使用残差结构连接输入数据以促进信息的前后向传播,即

$$\mathbf{Y} = [\mathbf{B}_1; \mathbf{B}_2] + \mathbf{X} \quad (13)$$

其中, $\mathbf{Y} \in \mathbb{R}^{H \times W \times D \times 2O}$ 为 RDT 模块的输出特征图, $[\cdot; \cdot]$ 表示通道拼接操作. RDT 模块利用通道分割和权重共享,实现了低复杂度的空谱特征提取.

3.3 轻量级全连接张量映射网络

本节利用 RDT 模块和 FCTNMConv3D4 单元构建一个新的轻量级模型 LiteFCTMN 用于提取高光谱图像的判别性特征并完成分类.模型的整体网络结构如图 5 所示,主要包括数据预处理,空谱特征提取以及高光谱图像分类三个部分.

在数据预处理阶段,以不同场景的每个像素点为中心,截取其邻域数据,构建张量形式的全波段数据结构作为模型的输入.以 Indian Pines 数据集为例,模型的输入样本大小为 $(9 \times 9 \times 200, 1)$,其中,括号内的前两个数字表示输入数据的空间维度大小,第三个数字表示输入数据的光谱维度大小,最后一个数字代表特征通道数.

在空谱特征提取过程中,经过预处理后的数据将依次通过一个 Conv3D 单元,一个 RDT 模块和一个 FCTNMConv3D4 单元. Conv3D 单元的大小为 $(1 \times 1 \times 7 \times 1 \times$

24),括号内的数字依次对应为卷积核的高度,宽度,长度以及输入和输出特征通道数.完成初级语义信息提取之后,大小为 $(9 \times 9 \times 97, 24)$ 的特征图将传入 RDT 模块学习高光谱图像的局部空谱信息.在 RDT 模块中,采用通道分割操作得到两个大小为 $(9 \times 9 \times 97, 12)$ 的特征,将其作为输入分别馈入两个分支.每个分支先通过三维点卷积层将特征通道数扩展至 48,丰富特征表达,再使用 FCTNMConv3D3 单元完成特征提取.单元内,第一个卷积层的大小为 $(1 \times 1 \times 1 \times 48 \times 4)$,第二个卷积层提取局部空谱联合特征,大小为 $(3 \times 3 \times 3 \times 2 \times 2)$,第三个卷积层通过点卷积运算将特征通道数恢复为 12.通过拼接两个分支的输出结果,并与 Conv3D 的输出进行残差连接,得到大小为 $(9 \times 9 \times 97, 24)$ 的 RDT 模块输出特征图.最后,将该特征图传入 FCTNMConv3D4 单元学习高光谱图像的全局光谱信息.在单元中,三维点卷积层先将特征通道数缩减至 8,并通过 $(3 \times 3 \times 1 \times 2 \times 4)$ 和 $(1 \times 1 \times 97 \times 4 \times 2)$ 的 2 个标准三维卷积层学习高光谱图像空间和全局光谱信息,再使用点卷积层将特征通道数增加至 60,得到大小为 $(9 \times 9 \times 1, 60)$ 的输出特征图.由于 FCTNMConv3D 单元使用多个具有全连接结构的小尺寸卷积核对原始卷积核张量进行近似,通过使用新设计的单元,所提 LiteFCTMN 模型能够以更低的复杂度提取更丰富的判别性特征.

最后,空谱特征提取阶段的输出特征图将依次通过一个全局平均池化(Global Average Pooling, GAP)层,一个全连接(Fully Connected, FC)层和一个 Softmax 层完成高光谱图像分类任务.模型的详细参数如表 2 所示,该模型在每个卷积层之后都加入 BN 层和 ReLU 激活函数.算法 1 给出了 LiteFCTMN 模型完整的构建流程.

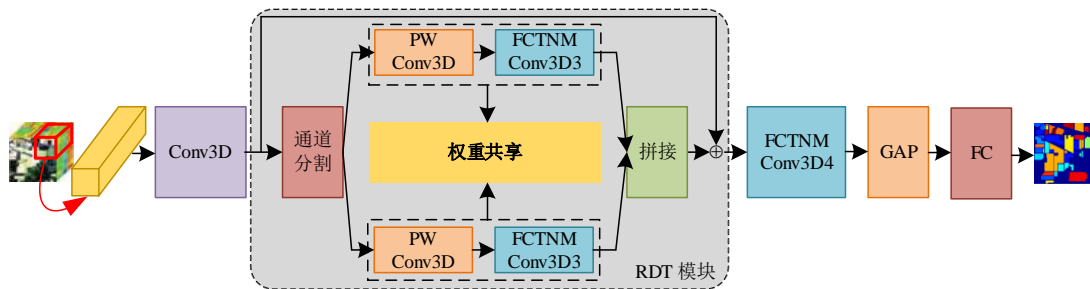


图 5 LiteFCTMN 模型

4 实验与分析

4.1 数据集

本文实验使用不同空间分辨率,不同像素数,不同波段数和不同地物类别数的 Indian Pines^[22]、Kennedy Space Center^[22]和 Houston^[23]数据集完成.每个数据集

被分为训练集和测试集,其中,Indian Pines 随机选取 5% 的样本作为训练数据,Kennedy Space Center 选取 3% 的样本作为训练集,其余样本用于测试,Houston 则使用预设的训练样本和测试样本.表 3 列出了三个数据集的详细训练样本个数和测试样本个数.

表 2 LiteFCTMN 模型在 Indian Pines 数据集的参数设置

层名称	卷积核尺寸	输出尺寸
输入层	—	(9×9×200, 1)
Conv3D	(1×1×7)	(9×9×97, 24)
通道分割	—	(9×9×97, 12)
PW Conv3D	(1×1×1)	(9×9×97, 48)
FCTNMConv3D3	(1×1×1)	(9×9×97, 4)
	(3×3×3)	(9×9×97, 4)
	(1×1×1)	(9×9×97, 12)
PWConv3D	(1×1×1)	(9×9×97, 48)
FCTNMConv3D3	(1×1×1)	(9×9×97, 4)
	(3×3×3)	(9×9×97, 4)
	(1×1×1)	(9×9×97, 12)
拼接	—	(9×9×97, 24)
残差连接	—	(9×9×97, 24)
FCTNMConv3D4	(1×1×1)	(9×9×97, 8)
	(3×3×1)	(9×9×97, 16)
	(1×1×97)	(9×9×1, 8)
	(1×1×1)	(9×9×1, 60)
GAP	—	(1×1×1, 60)
FC	—	(16)

算法 1 LiteFCTMN 构建流程

输入:

- (1) 高光谱图像训练集 X_t , 训练标签 Y_t , 测试集 X_v 以及测试标签 Y_v .
- (2) LiteFCTMN 模型参数 θ
- (3) 训练轮次 τ , 学习率 η

输出: 高光谱图像测试集样本预测类别

1: 随机初始化参数 θ

2: for t in range (0, τ) do

3: 通过标准三维卷积计算 X_t 的初级语义信息

4: 通过式(11)-(113)计算空谱联合特征

5: 通过式(4)计算全局光谱特征

6: 计算交叉熵损失 L_{cls}

7: 通过随机梯度下降更新

8: end for

9: 将测试集样本 X_v 输入训练好的模型完成分类

4.2 实验设置

为验证本文所提模型的性能以及复杂度, 采用 SVM^[24]、3DCNN^[25]、SSRN^[26]、DBDA^[27]、LiteDenseNet^[6]和 LiteDepthwiseNet^[9]六种高光谱图像分类算法作为对比模型. 采用的性能评价指标包括总体准确度 (Overall Accuracy, OA)、平均准确度 (Average Accuracy, AA) 和 Kappa (κ) 系数^[23]. 实验数据为 10 次随机实验的平均结果, 以消除随机选择训练样本造成的偏差. 实验的硬件环境为英特尔酷睿 i5-10600KF CPU 和 NVIDIA GeForce RTX 3070 GPU. LiteFCTMN 模型的代

码采用 Python 3.7.0 和 PyTorch 1.11.0 开发.

在所提模型的训练过程中, 综合考虑模型性能以及复杂度, 将 FCTNMConv3D 单元的秩设置为 2, 局部窗口大小设置为 9×9. 使用预热和余弦退火组合的学习率策略, 其中, 预热迭代次数为 50 轮次, 学习率峰值设置为 0.005. 模型训练迭代次数固定为 500 轮次, 批次大小为 16, 同时使用 Adam 优化器和交叉熵损失函数.

4.3 性能分析

根据上述实验环境以及参数设置, 表 4 列出了 LiteFCTMN 和其他六种高光谱图像分类模型在 Indian Pines、Kennedy Space Center 和 Houston 数据集下的分类性能. 由表可知, 所提模型在 OA、AA 和 κ 系数上均能取得最佳分类性能.

由于 SVM 模型破坏了高光谱图像张量形式的结构, 无法提取输入图像的空谱联合特征, 该模型在三个高光谱图像数据集中均表现出较低的性能. 与其他模型相比, 3DCNN 模型在 Kennedy Space Center 数据集中表现出最差的分类效果, 主要原因是该模型参数量过多, 且 Kennedy Space Center 数据集的训练样本数较少, 导致模型出现过拟合现象. DBDA 模型将 SSRN 模型连续的光谱特征学习阶段和空间特征学习阶段设计为双分支结构, 并加入注意力机制, 一定程度上提高了模型的性能. LiteDenseNet 和 LiteDepthwiseNet 模型的输入数据相同, 网络结构相似, 其分类性能也较为接近. 所提 LiteFCTMN 模型在三种不同数据集中均获得最高的分类性能. 主要原因在于新设计的 FCTNMConv3D 单元将单层卷积运算扩展为多层, 同时, 小尺寸卷积核之间的全连接拓扑结构能够有效降低卷积核解耦造成的信息损失, 有利于提高新单元特征提取能力, 也有助于模型学习更高级的语义信息, 从而提高了分类性能.

相较于 LiteDenseNet 模型, 所提模型在 Indian Pines 数据集中, OA 提升 0.94%, 在 Kennedy Space Center 数据集中提升 0.65%, 在 Houston 数据集中提升 1.02%. 与 LiteDepthwiseNet 模型相比, 所提模型在三个数据集中分别获得 0.56%、0.62% 以及 1.38% 的性能增益.

图 6 至图 8 为 LiteFCTMN 和其他对比模型在三个数据集下的分类结果图和原始场景的真实地物类别图. 从图中可以看出, LiteFCTMN 模型具有最平滑的分类结果图, 也与 Indian Pines、Kennedy Space Center 和 Houston 数据集的真实地物类别图最相似. 其余对比模型能够大致区分每个场景的不同类别, 但会出现错误分类的现象, 如 Indian Pines 数据集的 Soybean-notill 类, Kennedy Space Center 数据集的 Slash pine 和 Oak/Broad-leaf 类以及 Houston 数据集的 Highway 类.

表 3 三个数据集训练样本和测试样本个数

Indian Pines					Kennedy Space Center				Houston			
序号	颜色	类别	训练	测试	颜色	类别	训练	测试	颜色	类别	训练	测试
1		Alfalfa	5	41		Srub	23	738		Health grass	198	1 053
2		Corn-notill	71	1 357		CP swamp	7	236		Stressed grass	190	1 064
3		Corn-mintill	41	789		CP hammock	8	248		Synthetic grass	192	505
4		Corn	12	225		Slash pine	8	244		Trees	188	1 056
5		Grass-pasture	24	459		Oak/Broadleaf	5	156		Soil	186	1 056
6		Grass-trees	37	693		Hardwood	7	222		Water	182	143
7		Grass-pasture-mowed	5	23		Swamp	3	102		Residential	196	1 072
8		Hay-windrowed	24	454		Graminoid	13	418		Commercial	191	1 053
9		Oats	5	15		Spartina marsh	16	504		Road	193	1 059
10		Soybean-notill	49	923		Cattail marshl	49	923		Highway	191	1 036
11		Soybean-mintill	109	2 346		Salt marsh	12	392		Railway	181	1 054
12		Soybean-clean	30	563		Mud flats	13	406		Parking Lot 1	192	1 041
13		Wheat	12	193		Water	15	488		Parking Lot 2	184	285
14		Woods	63	1 202						Tennis court	181	247
15		Buildings-Grass-Trees-Drives	19	369						Running track	187	473
16		Stone-Steel-Towers	6	87								
总计			512	9 737			158	5 053			2 832	12 197

表 4 不同模型在三个数据集上的分类性能

单位:%

模型	Indian Pines			Kennedy Space Center			Houston		
	OA	AA	κ	OA	AA	κ	OA	AA	κ
SVM	79.66	79.09	76.68	80.44	71.08	78.13	70.85	68.35	68.43
3DCNN	95.69	95.53	95.09	76.70	67.79	73.94	83.51	83.38	82.12
SSRN	94.50	95.35	93.73	94.32	90.12	93.67	84.09	85.70	82.77
DBDA	94.37	95.06	93.58	94.74	91.30	94.15	84.93	87.21	83.66
LiteDenseNet	95.57	95.89	94.95	96.27	93.14	95.84	85.47	87.68	84.27
LiteDepthwiseNet	95.95	96.67	95.38	96.30	93.49	95.89	85.11	87.07	83.89
LiteFCTMN	96.51	96.93	96.02	96.92	94.34	96.58	86.49	88.02	85.35

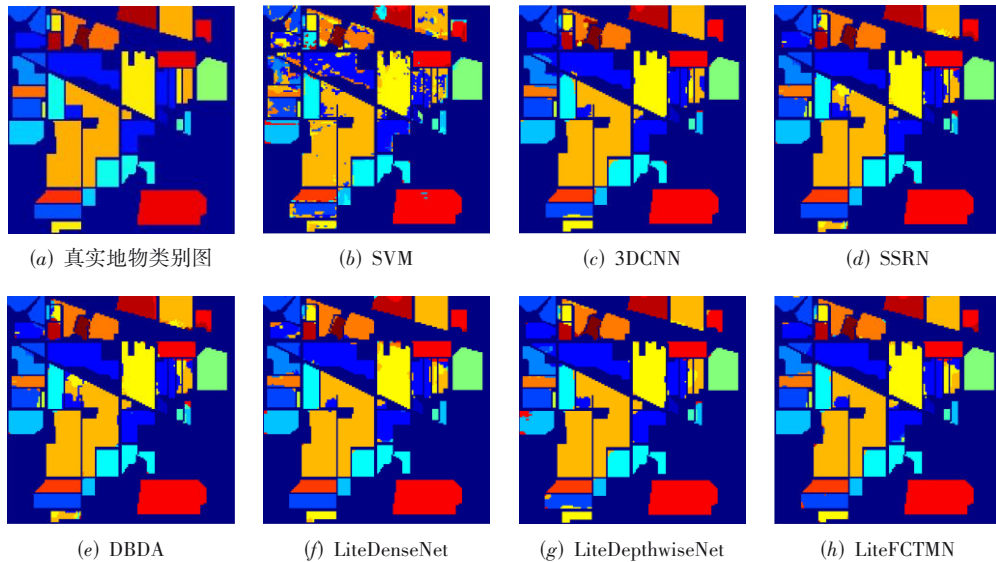


图 6 不同模型在 Indian Pines 数据集上分类结果图

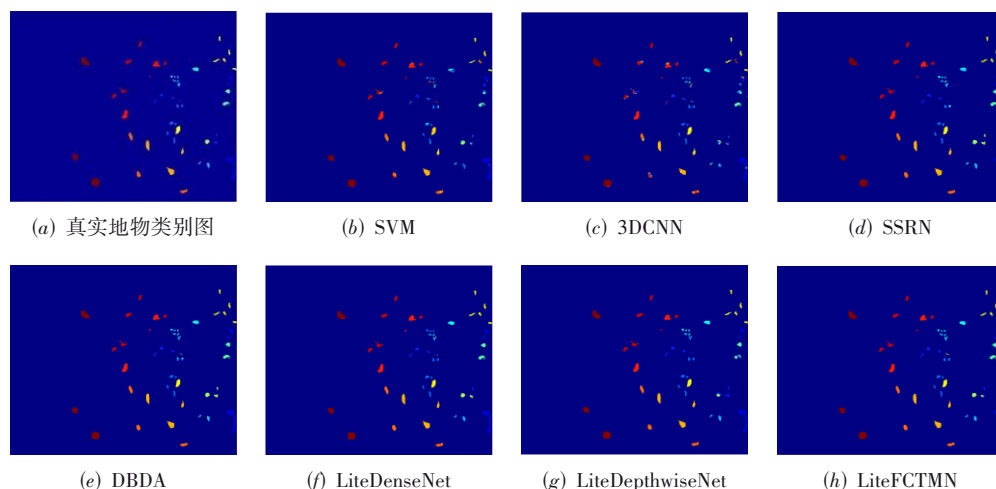


图7 不同模型在 Kennedy Space Center 数据集上分类结果图

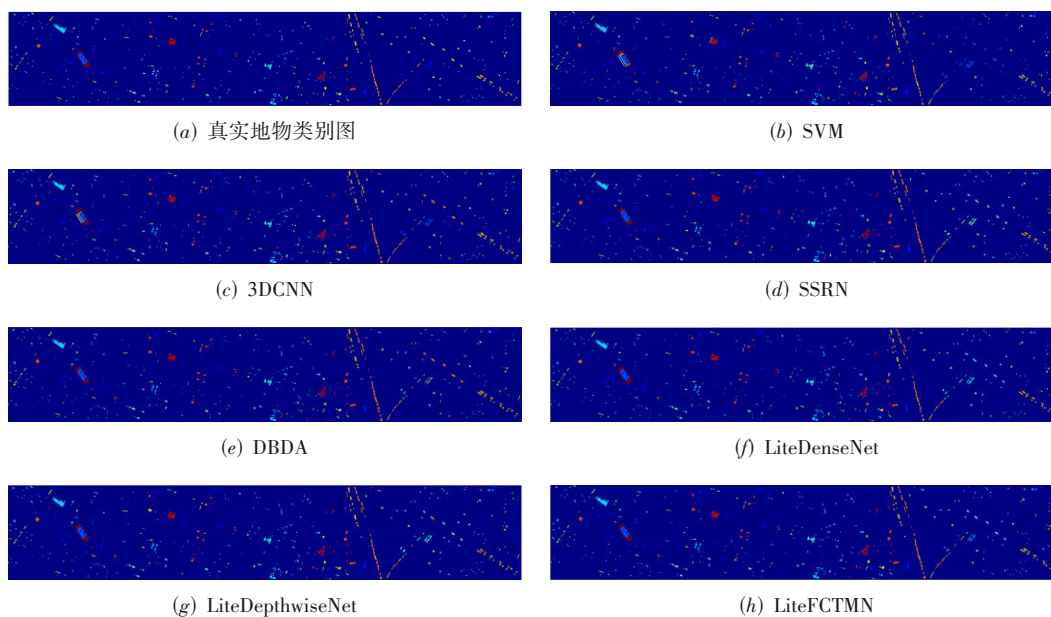


图8 不同模型在 Houston 数据集上分类结果图

4.4 性能分析复杂度比较

本节研究 LiteFCTMN 和对比模型的复杂度. 表 5 列出了不同模型在 Indian Pines、Kennedy Space Center 和 Houston 数据集上的参数量 (Parameters, Params), 单位为千 (kilo, k) 以及浮点运算数 (Floating Point Operations, FLOPs), 单位为兆 (Million, M), 并以此评估模型复杂度.

在所有模型中, 由于 3DCNN 模型输入数据较大, 该模型复杂度最高. DBDA 模型较 SSRN 模型增加了注意力机制, 导致模型的计算复杂度上升. LiteDepthwiseNet 模型将 LiteDenseNet 模型中分组卷积替换为深度可分离卷积, 有效减小了模型复杂度. 而所提 LiteFCTMN 模型使用新设计的 FCTNMConv3D 单元代替标准三维卷

积层, 用多个小尺寸卷积核近似原始的卷积核, 减少了模型计算量. 同时, LiteFCTMN 模型还使用通道分割操作减少卷积层特征通道数, 进一步降低模型复杂度. 具体来说, 在 Indian Pines 数据集中, LiteFCTMN 模型较 LiteDenseNet 模型在参数量方面压缩了 231.60 倍, 计算量减少 89.60%. 与 LiteDepthwiseNet 模型相比, 所提模型参数量压缩 14.05 倍, 计算量减少 63.53%. 其余两组数据集上也表现出相似的实验结果. 综上所述, 所提 LiteFCTMN 模型能够兼具分类性能和模型复杂度.

4.5 不同张量分解方法比较

最后, 对比 CP 分解、tucker 分解、TT 分解和 FCTN 分解在本文模型中的效果. 为公平比较, 实验仅将 FCTNMConv3D3 和 FCTNMConv3D4 单元替换为对应的

表 5 不同模型在三个数据集上的复杂度

模型	Indian Pines		Kennedy Space Cente		Houston	
	Params/k	FLOPs/M	Params/k	FLOPs/M	Params/k	FLOPs/M
3DCNN	1 803.80	273.20	1 802.90	273.20	1 803.50	273.20
SSRN	364.17	95.58	327.23	83.91	278.13	68.35
DBDA	382.35	108.17	338.21	94.99	280.08	77.41
LiteDenseNet	852.30	171.81	748.44	150.56	610.32	122.22
LiteDepthwiseNet	51.71	49.00	46.35	42.97	39.56	34.93
LiteFCTMN	3.68	17.87	3.41	15.67	3.47	12.74

基于 CP 分解、tucker 分解和 TT 分解的卷积单元,其余设置均保持不变.对比单元的秩根据 LiteFCTMN 模型的复杂度确定,即 CP 分解的秩设置为 6,构建 CP-6 模型. Tucker 分解的秩取 3,构建 Tucker-3 模型. TT 分解的秩取 4,构建 TT-4 模型.

表 6 为不同张量分解模型的性能、参数量以及计算量.从表中可以看出,在模型复杂度相近情况下,所提模型具有最高 OA 值,基于 CP 分解的模型次之,基于 tucker 分解的模型分类性能最低.具体来说,在 Indian Pines 数据集中, LiteFCTMN 模型较 CP-6 模型性能

提升 0.28%;与基于 tucker 分解的模型相比, LiteFCTMN 模型 OA 值比 Tucker-3 模型提高 2.78%;与基于 TT 分解的模型相比,所提模型获得 0.69% 性能增益.其余两组数据集上也表现出相似的实验结果.

上述结果是由于 FCTNMConv3D 单元的全连接结构能够充分表征任意两个因子张量之间的相关性,同时具有转置不变性,使得模型能够获得更好的准确性.而其余分解形式只连接部分因子张量,导致模型分类性能较低.总体而言,本文所提 LiteFCTMN 模型能够在低复杂度的情况下取得出色的高光谱图像分类效果.

表 6 不同张量分解模型分类性能与复杂度比较

模型	Indian Pines			Kennedy Space Center			Houston		
	OA/%	Params/k	FLOPs/M	OA/%	Params/k	FLOPs/M	OA/%	Params/k	FLOPs/M
CP-6	96.28	3.48	18.30	95.84	3.22	16.05	84.45	3.66	13.04
Tucker-3	93.73	3.03	18.09	94.33	2.81	15.86	83.71	3.30	12.88
TT-4	95.82	4.52	20.36	95.39	4.15	17.85	83.90	4.65	14.50
LiteFCTMN	96.51	3.68	17.87	96.92	3.41	15.67	86.49	3.47	12.74

5 结论

本文基于 FCTN 分解提出了轻量级全连接张量映射网络 LiteFCTMN 模型用于高光谱图像分类研究,该模型具有出色的分类性能和极低的复杂度.通过设计两个低复杂度的三维张量映射卷积单元(FCTNMConv3D4 和 FCTNMConv3D3),并构建 RDT 模块, LiteFCTMN 模型在多个广泛使用的高光谱图像数据集上均能够以较低的模型参数量和计算量获得出色的分类性能.

参考文献

- [1] RAYA-SERENO M D, ALONSO-AYUSO M, PANCORBO J L, et al. Residual effect and N fertilizer rate detection by high-resolution VNIR-SWIR hyperspectral imagery and solar-induced chlorophyll fluorescence in wheat[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-17.
- [2] 赵旭东,陶然,李伟,等.高光谱图像分数域信息提取理

论与方法进展[J].电子学报,2022,50(12):2874-2883.

ZHAO X D, TAO R, LI W, et al. Recent developments in fractional information extraction theory and methods of hyperspectral image[J]. Acta Electronica Sinica, 2022, 50(12): 2874-2883. (in Chinese)

- [3] HÄNSCH R, HELLWICH O. Fusion of multispectral LiDAR, hyperspectral, and RGB data for urban land cover classification[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 18(2): 366-370.
- [4] 方健,杨劲翔,肖亮.基于多层小波深度聚合网络的高光谱图像超分辨率方法[J].电子学报,2024,52(1):201-216. FANG J, YANG J X, XIAO L. MW-DAN: Multilevel wavelet-deep aggregation network for hyperspectral image super-resolution[J]. Acta Electronica Sinica, 2024, 52(1): 201-216. (in Chinese)
- [5] KHAN M H, SALEEM Z, AHMAD M, et al. Hyperspectral imaging-based unsupervised adulterated red chili content transformation for classification: Identification of red

- chili adulterants[J]. *Neural Computing and Applications*, 2021, 33(21): 14507-14521.
- [6] LI R, DUAN C. LiteDenseNet: A lightweight network for hyperspectral image classification[EB/OL]. (2020-04-17) [2024-05-08]. <https://arxiv.org/abs/2004.08112>.
- [7] 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). Piscataway: IEEE, 2017: 2261-2269.
- [8] ZHANG X Y, ZHOU X Y, LIN M X, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 6848-6856.
- [9] CUI B L, DONG X M, ZHAN Q Q, et al. LiteDepthwiseNet: A lightweight network for hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-15.
- [10] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd International Conference on Machine Learning. Cambridge: JMLR, 2015: 448-456
- [11] CHENG Y, WANG D, ZHOU P, et al. Model compression and acceleration for deep neural networks: The principles, progress, and challenges[J]. *IEEE Signal Processing Magazine*, 2018, 35(1): 126-136.
- [12] DENG L, LI G Q, HAN S, et al. Model compression and hardware acceleration for neural networks: A comprehensive survey[J]. *Proceedings of the IEEE*, 2020, 108(4): 485-532.
- [13] HU W S, LI H C, DENG Y J, et al. Lightweight tensor attention-driven ConvLSTM neural network for hyperspectral image classification[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2021, 15(3): 734-745.
- [14] OSELEDETS I V. Tensor-train decomposition[J]. *SIAM Journal on Scientific Computing*, 2011, 33(5): 2295-2317.
- [15] MA T Y, LI H C, WANG R, et al. Lightweight tensorized neural networks for hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-16.
- [16] ZHAO Q, ZHOU G, XIE S, et al. Tensor ring decomposition[EB/OL]. (2016-06-17)[2024-05-08]. <https://arxiv.org/pdf/1606.05535>.
- [17] ZHENG Y B, HUANG T Z, ZHAO X L, et al. Fully-connected tensor network decomposition and its application to higher-order tensor completion[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, 35(12): 11071-11078.
- [18] LI H C, LIN Z X, MA T Y, et al. Hybrid fully connected tensorized compression network for hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 1-16.
- [19] LEBEDEV V, GANIN Y, RAKHUBA M, et al. Speeding-up convolutional neural networks using fine-tuned CP-decomposition[EB/OL]. (2015-04-24) [2024-05-08]. <https://arxiv.org/pdf/1412.6553>.
- [20] 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). Piscataway: IEEE, 2016: 770-778.
- [21] Su J H, Li J L, Bhattacharjee B, et al. Tensorial neural networks: Generalization of neural networks and application to model compression[EB/OL]. (2018-12-08) [2024-05-08]. <https://arxiv.org/pdf/1805.10352>.
- [22] Grupo de Inteligencia Computacional de la Universidad del PAÍS Vasco. Hyperspectral remote sensing scenes [EB/OL]. (2021-07-12) [2024-05-08]. https://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes.
- [23] XU X D, LI W, RAN Q, et al. Multisource remote sensing data classification based on convolutional neural network[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(2): 937-949.
- [24] CHANG C C, LIN C J. LIBSVM: A library for support vector machines[J]. *ACM Transactions on Intelligent Systems and Technology*, 2011, 2(3): 1-27.
- [25] PAOLETTI M E, HAUT J M, PLAZA J, et al. Deep learning classifiers for hyperspectral imaging: A review [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2019, 158: 279-317.
- [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] 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): 582-606.

作者简介



林知心 女, 1998年5月出生于福建省福州市。毕业于西南交通大学信息科学与技术学院。主要研究方向为张量分解、神经网络压缩和高光谱图像处理等。

E-mail: zhixin_lin@163.com



郑玉棒 男, 1993年11月出生于安徽省宿州市。现为西南交通大学信息科学与技术学院讲师、硕士生导师。主要研究方向包括高维信号与图像处理、张量建模与智能计算、机器学习及其数学理论等。

E-mail: zhengyubang@163.com



马天宇 男, 1996年7月出生于河北省邯郸市。现为西南交通大学信息科学与技术学院在读博士生。主要研究方向为张量分解、神经网络压缩和高光谱图像处理等。

E-mail: mty9678@my.swjtu.edu.cn



王蕊 女, 1978年3月出生于甘肃省平凉市。现为西南交通大学信息科学与技术学院副教授、硕士生导师。主要研究方向为高光谱图像解混、遥感图像处理。

E-mail: wangrui@swjtu.edu.cn



李恒超 男, 1978年11月出生于山东省临沂市。现为西南交通大学信息科学与技术学院教授、博士生导师。主要研究方向为遥感图像处理与解译、计算智能与模式识别。中国电子学会会员编号: E190092013M。

E-mail: heli@swjtu.edu.cn