

DOI:10.11918/202211002

# 结合注意力机制的轻量化网络高光谱图像分类

张迎豪, 杨芳

(武汉科技大学 信息科学与工程学院 人工智能与信息融合研究院, 武汉 430000)

**摘要:** 高光谱图像分类是遥感领域的核心难题,其挑战主要源于高光谱图像的光谱维度高、空间分辨率低,导致分类任务中难以充分提取高光谱图像的空谱特征。针对现有的卷积神经网络(CNN)高光谱图像分类模型参数量大、计算资源消耗高、分类精度不足的问题,提出了一种基于注意力机制的轻量化网络高光谱图像分类模型(AMLW-CNN)。为了增强网络的特征提取能力,将空谱特征提取模块设计为2个多尺度提取模块,并将空间特征提取模块的各卷积层通过残差结构连接,同时引入注意力机制来强化网络对有效特征的提取。另外,为了减少模型参数量,采用非对称卷积来替代三维卷积核、深度可分离卷积来替换二维卷积核。实验结果表明:AMLW-CNN的分类精度优于对比算法,计算复杂度更低,鲁棒性更强。在Indian Pines、Salinas和Pavia U 3个数据集上的总体分类精度分别达到了98.5%、99.8%、99.9%。

**关键词:** 高光谱图像分类;多尺度提取模块;注意力机制;轻量化网络;非对称卷积;深度可分离卷积

中图分类号: TP391

文献标志码: A

文章编号: 0367-6234(2026)03-0151-13

## Hyperspectral image classification based on lightweight network with attention mechanism

ZHANG Yinghao, YANG Fang

(Institute for Artificial Intelligence and Information Fusion, School of Information Science and Engineering, Wuhan University of Science and Technology, WuHan 430000, China)

**Abstract:** Hyperspectral image (HSI) classification is a challenging task in the field of remote sensing, because the HSI has high spectral dimensionality and low spatial resolution, which makes it difficult to fully extract the spatial-spectral features of hyperspectral images in the classification task. Aiming at solving the problems of the existing convolutional neural network (CNN)-based HSI classification models, such as large parameter size, high computational cost and low classification accuracy, a lightweight network hyperspectral image classification model based on attention mechanism (AMLW-CNN) is proposed in this paper. In order to enhance feature extraction ability of the network, the spatial-spectral feature extraction module is designed based on two multiscale extraction modules. In addition, we use the residual structures to connect the convolutional layers of spatial feature extraction module and incorporate the attention mechanism to enhance the extraction of useful features. Furthermore, to reduce the number of model parameters, an asymmetric convolution and a depthwise separable convolution are introduced to replace the 3D and 2D convolution kernels, respectively. The experimental results show that classification accuracy of AMLW-CNN is better than that of the comparison algorithms, with lower computational complexity and higher robustness. The overall classification accuracies on the datasets of Indian Pines, Salinas and Pavia U has attain 98.5%, 99.8% and 99.9%, respectively.

**Keywords:** hyperspectral image classification; multi-scale extraction module; attention mechanism; lightweight network; asymmetric convolution; depthwise separable convolution

20世纪80年代初,遥感技术在理论和实际应用中取得了显著进展,随着成像光谱理论的提出,光学遥感正式迈入到高光谱遥感阶段<sup>[1]</sup>。高光谱图像包含了丰富的地物空间和光谱维度信息,随着航空航天事业的发展以及高光谱成像技术的完善,高光谱图像分类展现出了巨大的实用价值和发展潜

力,其在灾害监测、军事侦察、刑事调查、国土规划<sup>[2-3]</sup>等领域的应用价值和发展潜力日益凸显,分类任务也成为遥感领域的研究热点。

高光谱图像分类的精度高度依赖于训练样本的数量,但是高光谱图像的空间分辨率比较低,导致训练样本不足,容易出现同谱异物和异物同谱现象。

收稿日期: 2022-11-01; 录用日期: 2023-01-28; 网络首发日期: 2024-06-27

网络首发地址: <https://link.cnki.net/urlid/23.1235.T.20240626.1616.007>

基金项目: 国家自然科学基金(62101392)

作者简介: 张迎豪(1998—),男,硕士研究生;杨芳(1989—),女,副教授,硕士生导师

通信作者: 杨芳, yangfang\_idif@wust.edu.cn

同时,由于高光谱图像的光谱分辨率很高(通常包含上百个波段),在处理过程中容易出现休斯(Hughes)现象<sup>[4]</sup>,导致分类精度下降。现有的高光谱图像分类算法主要分为传统的机器学习算法和深度学习算法两类,前者存在对特征利用不充分的问题,后者则普遍面临参数量过大的挑战。针对这些问题,本文提出基于注意力机制的轻量化网络高光谱图像分类模型(AMLW-CNN),该模型主要由三维卷积神经网络(3D-CNN)、二维卷积神经网络(2D-CNN)和线性分类单元组成。首先通过主成分分析法(principal component analysis,PCA)<sup>[5]</sup>将高光谱图像的频带数进行缩减,把降维后的数据输入 3D-CNN 进行多尺度空谱特征提取;然后通过通道注意力机制和空间注意力机制,根据重要性为输入数据的不同通道和空间区域赋予相应的权重;再将注意力机制模块输出的数据输入结合了深度可分离卷积和残差结构的 2D-CNN,提取更高级特征;最后将数据输入到线性分类单元输出分类结果。

## 1 相关研究

传统的高光谱图像分类方法通过特征选择和特征提取,然后将处理后的数据输入分类模型进行分类,如支持向量机(SVM)、K-近邻算法(KNN)<sup>[6]</sup>、逻辑回归<sup>[7]</sup>等。这类算法依赖人工设提取低级和中级特征,未能充分利用高光谱图像的内部关联,破坏了图像中原有的空谱结构,不能综合提取数据的特征,因此特征提取能力有限、分类效果欠佳;此外,传统的机器学习算法泛化能力较差,往往只对特定数据集有效,而高光谱图像的分类复杂多样,选择合适的特征提取模型有一定的困难。

近年来,深度学习技术快速发展,堆叠式自编码器(SAE)<sup>[8]</sup>、深度信念网络(DBN)<sup>[9]</sup>、递归神经网络(RNN)<sup>[10]</sup>和卷积神经网络(CNN)<sup>[11]</sup>等模型被广泛应用于高光谱图像分类,衍生出胶囊网络(CapNet)<sup>[12]</sup>、生成对抗网络(GAN)<sup>[13]</sup>、图卷积网络(GCN)<sup>[14]</sup>、注意力机制网络<sup>[15]</sup>等方法。Zhu 等<sup>[16]</sup>提出了深度卷积胶囊网络,结合 CNN 和胶囊网络缓解参数量过大导致的过拟合问题;Lin 等<sup>[17]</sup>利用 GAN 扩增数据,减轻空间分辨率低的影响,在样本点有限的场景中提高分类精度;Hong 等<sup>[18]</sup>提出基于 Transformer 的高光谱图像分类算法,提升模型对光谱连续信息的分析能力,减小光谱维度的特征损失;Li 等<sup>[19]</sup>提出双分支双注意力机制网络,通过上下两个分支依次

对应通道注意力机制模块和空间注意力机制模块,增强模型对空谱信息的提取能力。

由于在空间结构信息的特征提取方面 CNN 有显著优势,因此,本文主要分析基于 CNN 的高光谱图像分类算法。根据提取特征类别可以分为 3 类:1D-CNN、2D-CNN 和 3D-CNN。

1D-CNN 通过一维卷积核遍历高光谱图像,提取像素关联的深层光谱特征。Hu<sup>[20]</sup>将 1D-CNN 应用在高光谱图像分类中,该网络包含 5 个带权层,分别是输入层、卷积层、最大池化层、全连接层和输出层,可深度提取光谱特征,但是无法获得图像空间维度信息,对高光谱图像的空间特征利用不充分。

2D-CNN 通过二维卷积核遍历高光谱图像,提取不同波段的空谱信息,在自然图像中应用广泛。但是由于高光谱图像波段比较多,若直接处理需要为每个特征图都配置一个卷积核,导致参数量剧增、计算资源消耗大且易引起过拟合,因此需先用 PCA 对高光谱图像进行降维处理。Xu 等<sup>[21]</sup>提出的 RpNet 引入了多尺度提取模块,提升了空间特征提取能力,进而提高了分类精度。但是 2D-CNN 无法利用光谱特征,对光谱特征的提取不够充分,导致分类效果不佳。高光谱图像同时包含了空间和光谱信息,单独使用 1D-CNN 或 2D-CNN 对特征的提取都不充分。为了解决这个问题,Yang 等<sup>[22]</sup>设计双通道 CNN,分别用 1D-CNN 和 2D-CNN 提取高光谱图像的光谱特征和空间特征并融合,可以实现更加充分的特征提取,达到了较好的分类水平。

3D-CNN 通过三维卷积核在空间和光谱维度移动,同步提取空间和光谱特征。Li 等<sup>[23]</sup>提出了 3D-CNN 的基本架构,不用对原始高光谱图像降维就可以进行特征提取,最后将高级特征输入分类器进行分类,得到了较好的分类结果。He 等<sup>[24]</sup>设计多尺度全卷积 3D-CNN,提取高光谱图像的光谱和空间信息,进一步提升了分类水平。尽管 3D-CNN 对特征的提取更加充分,实现了很好的分类效果,但其参数量过大,计算成本较高。Lee 等<sup>[25]</sup>提出混合 CNN,首先利用 3D-CNN 提取高光谱图像的空谱联合特征,再输入 2D-CNN 提取高级特征,在保证精度的同时,也大幅降低了模型的参数量。

为了更细致地提取高光谱图像深层次的空谱信息,同时降低模型的参数量,本文结合高光谱图像结构特点,基于混合 CNN,提出了融合深度可分离卷积、多尺度提取模块、残差结构的轻量化网络模型;

为了提高模型的分类精度,引入 CBAM(convolutional block attention module)注意力机制模块<sup>[26]</sup>。本文的主要贡献如下:

1)优化混合 CNN 结构。采用残差结构加深网络深度,提高模型的非线性拟合能力,以适配不同类型的数据集。同时采用非对称卷积替代 3D 卷积核、深度可分离卷积替代 2D 卷积核,大幅降低网络参数量,在保证精度的同时减少计算资源消耗。

2)引入注意力机制。在经典的基于 CNN 的高光谱图像分类过程中,特征在前向传播时,各通道和空间区域对下一层特征提取的贡献度相同,未能区分特征重要性,影响分类精度。本文引入注意力机制,把计算资源更有效地分配给权重较大的特征,提

高网络对重要信息的提取能力。

## 2 基于注意力机制的轻量化网络

高光谱图像为包含空间信息和光谱信息的三维立方体数据,模型的输入为三维张量  $X \in \mathbf{R}^{H \times W \times B}$  ( $H$  为图像高度、 $W$  为宽度、 $B$  为波段数,每个像素都包含  $B$  个光谱维度),模型的输出分类结果图为  $Y = \{y_1, y_2, \dots, y_C\} \in \mathbf{R}^{H \times W}$  ( $y$  为地物类别分类结果,  $C$  为地物类别数)。AMLW-CNN 网络架构如图 1 所示,分为 5 个模块:预处理模块、空谱特征提取模块、注意力机制模块、空间特征提取模块、线性分类模块。每个模块依次对高光谱图像进行处理,然后将数据输入下一个模块。

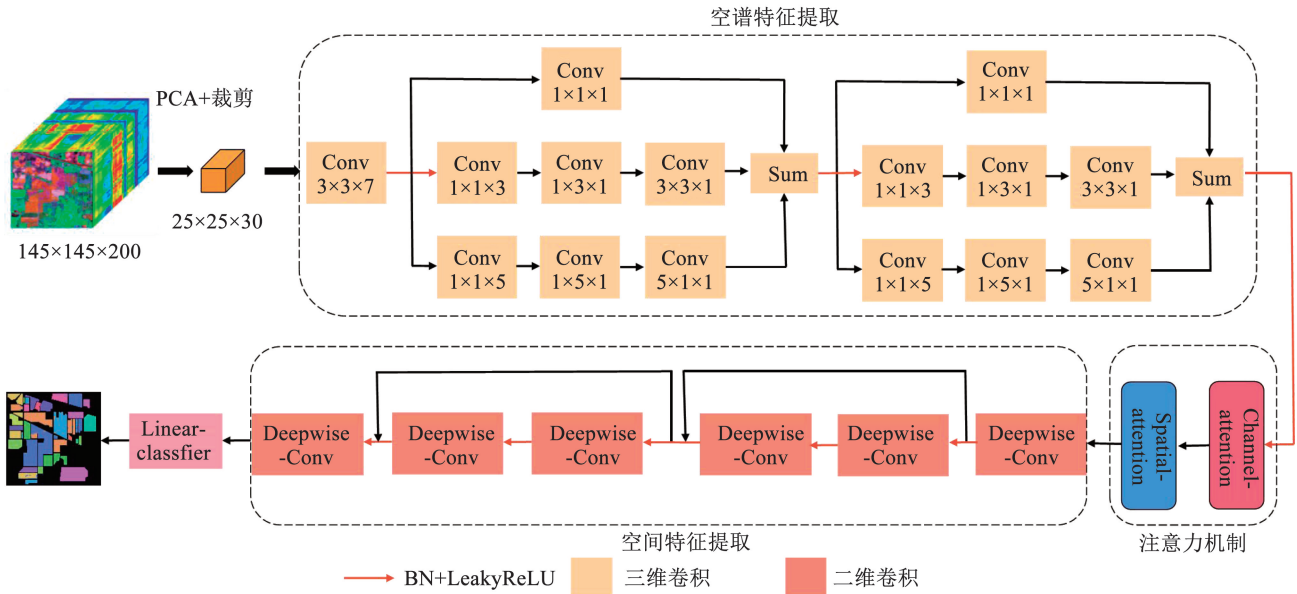


图 1 AMLW-CNN 网络框架图

Fig. 1 AMLW-CNN network framework diagram

### 2.1 预处理模块

高光谱图像的高维特征会导致卷积层参数量过大、计算资源消耗高、训练时间延长。为此,采用 PCA 对图像进行线性降维,在尽可能保留原始信息的前提下压缩光谱维度。PCA 通过将数据投影到信息量最大的线性投影维度上,把高维数据映射到低维空间。如将光谱维度为  $D$  的特征投影到  $B$  维空间 ( $D > B$ ),可以表示为  $X \in \mathbf{R}^{M \times N \times B}$  ( $M$ 、 $N$  为高光谱图像的空间尺寸,  $D$  为原始光谱维度尺寸,  $B$  为投影后的光谱维度尺寸)。具体步骤为:(1)去除高光谱数据的平均值,使数据分布更集中;(2)计算高光谱数据的协方差矩阵;(3)求解协方差矩阵的特征值与特征向量;(4)对特征值按降序排序;(5)保留前  $B$  个最大的特征值对应的特征向量;(6)将高光

谱图映射至所选特征向量构成的低维空间,最终得到降维后的数据。

### 2.2 空谱特征提取模块

考虑到高光谱图像的数据结构特点,该模块通过 3D-CNN 提取降维后数据的空谱联合特征。第一层 3D-CNN 的卷积核尺寸为  $3 \times 3 \times 7$ ,在空间维度用  $3 \times 3$  尺寸提取空间特征,步长设置为 1;在光谱维度用  $1 \times 7$  尺寸提取光谱特征,步长设置为 3。将初步提取的特征输入两个连续的多尺度提取模块 (Inception),进一步强化特征提取。Inception 结构如图 2 所示,包含 3 个并行通道,卷积核尺寸分别为  $1 \times 1 \times 1$ 、 $3 \times 3 \times 3$ 、 $5 \times 5 \times 5$ ,可以同时提取同一像素不同空间范围内的空谱特征。同时,为降低参数量,在保证分类性能的前提下,采用非对称卷积替换

$n \times n \times n$ 卷积核,将其分解为  $1 \times 1 \times n$ 、 $1 \times n \times 1$ 、 $n \times 1 \times 1$  的卷积组合。在减少参数的同时,增加网络深度,提高非线性拟合能力。空谱特征提取完成后,将数据输入注意力机制模块。

另外,在模型训练过程中,每层输入都会受前面所有网络层的影响,网络参数的微小波动会导致输出数据分布不断改变,产生内部协变量偏移(internal covariate shift, ICS)<sup>[27]</sup>。为了消除这种现象,在每层激活函数前加入批量归一化(batch normalization, BN)层,通过固定数据均值和方差实现标准化,减少各层输入数据的分布波动,从而加速训练收敛。同时, BN 层还可以缓解学习率过大导致的梯度爆炸或消失问题,并通过正则化作用降低过拟合风险,无需依赖权重初始化或 Dropout 层。

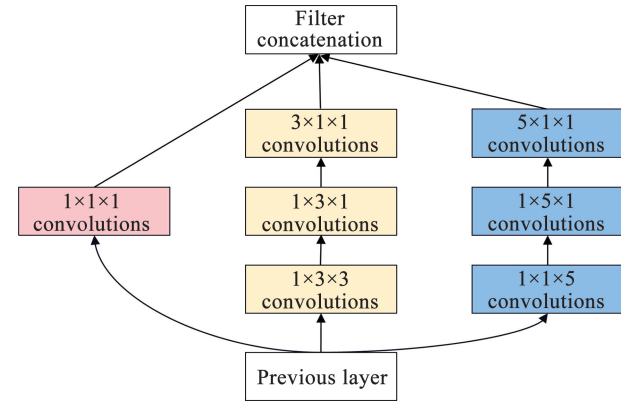


图 2 Inception 结构  
Fig. 2 Inception structure

2.3 注意力机制模块

在 AMLW-CNN 模型中引入 CBAM 模块,将算力集中在对分类结果更有用的特征上,增大特征类间方差,从而提高分类精度。CBAM 由通道注意力机制(CAM)和空间注意力机制(SAM)组成,分别对输入数据的通道以及空间特征分配权重。

1) 通道注意力机制(CAM)

CAM 基于空间维度操作,对输入特征图( $H \times W \times B$ )分别执行全局最大池化和全局平均池化,得到两个  $1 \times 1 \times B$  的特征向量;然后将 2 个特征向量分别输入共享多层感知机(MLP),通过两层卷积提取特征向量后求和,再输入 Sigmoid 激活函数将特征值压缩到  $[0, 1]$  区间,得到各通道的重要性权重;最后将权重与原始输入特征图逐通道相乘,输出通道权重分配后的特征图。其计算过程如式(1)所示:

$$M_c(F) = \sigma(\text{MLP}(\text{Avg Pool}(F)) + \text{MLP}(\text{Max Pool}(F))) = \sigma(W_1(W_0(F_{\text{avg}}^c)) + W_1(W_0(F_{\text{max}}^c))) \quad (1)$$

式中: $F$  为输入特征值; $\sigma$  为 Sigmoid 函数; $F_{\text{avg}}^c$  和  $F_{\text{max}}^c$  分别为经过全局平均池化和全局最大池化后的数据; $W_0$  和  $W_1$  为共享 MLP 层的 2 个卷积层的权重。

2) 空间注意力机制(SAM)

SAM 基于光谱维度操作,对 CAM 输出的特征图分别执行全局最大池化和全局平均池化,得到 2 个  $H \times W \times 1$  的特征图;将 2 个特征图在通道维度拼接为  $H \times W \times 2$  的张量,通过  $7 \times 7$  卷积核(补 0 保证特征图尺寸不变)进行特征提取,经 Sigmoid 激活函数得到空间区域权重;最后将权重与原始输入特征图逐像素相乘,输出空间权重分配后的特征图。其计算过程可以表示为

$$M_s(F) = \sigma(f^{7 \times 7}([\text{Avg Pool}(F); \text{Max Pool}(F)])) = \sigma(f^{7 \times 7}([F_{\text{avg}}^s; F_{\text{max}}^s])) \quad (2)$$

式中: $f^{7 \times 7}$  表示  $7 \times 7$  的卷积运算。

CBAM 通过“通道权重分配→空间权重分配”的两步操作,实现对关键特征的聚焦与无用特征的抑制,随后将加权特征图输入到空间特征提取模块进行进一步处理。通道注意力机制如图 3 所示,空间注意力机制如图 4 所示。

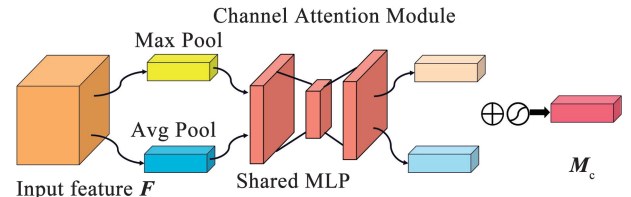


图 3 通道注意力机制模块

Fig. 3 Channel attention mechanism module

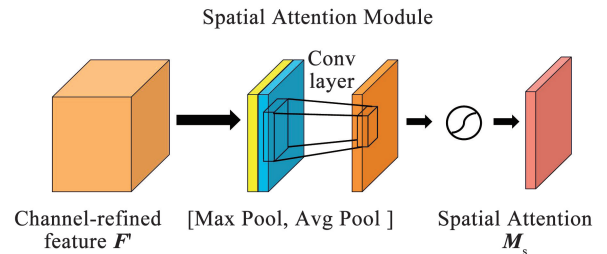


图 4 空间注意力机制模块

Fig. 4 Spatial attention mechanism module

2.4 空间特征提取模块

由于 3D-CNN 庞大的计算量会过度消耗资源,因此空间特征提取模块主要通过 2D-CNN 来提取高光谱图像更深层次的空间特征,设计了 6 层二维卷积层,引入轻量化网络,通过改变计算方式降低网络参数量。

1) 深度可分离卷积优化

将空间特征提取模块中的二维卷积层替换为深度可分离卷积<sup>[28]</sup>,该结构由深度卷积和逐点卷积组成,并在每层后添加 BN 层和 ReLU 激活函数,可以在降低计算量的同时,增强模型的非线性拟合能力。深度可分离卷积和普通卷积的计算方式不同,计算量也有很大差异,两者的计算过程分别如图 5、图 6 所示。

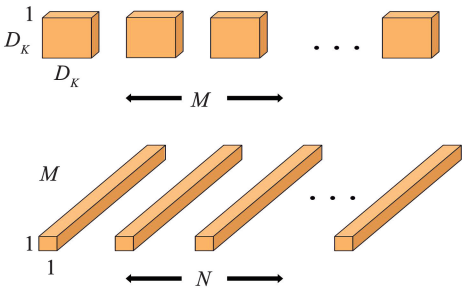


图 5 深度可分离卷积计算过程

Fig. 5 Depth-separable convolution calculation process

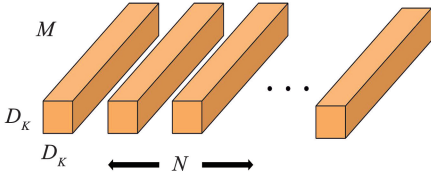


图 6 普通卷积计算过程

Fig. 6 Ordinary convolution computing procedure

以下详细分析两种卷积的计算量差异。一次卷积运算的乘法次数等于卷积核大小,即  $D_k \times D_k \times M$ 。由于有  $N$  个卷积核,并且每个卷积核在空间维度上滑动  $D_f \times D_f$  次,所以普通卷积的乘法次数为  $N \times D_f^2 \times D_k^2 \times M$ 。

深度卷积阶段:卷积核尺寸为  $D_k \times D_k \times 1$  (仅覆盖单个输入通道的空间维度),共配置  $M$  个核以匹配  $M$  个输入通道;这些核在  $M$  个通道上滑动  $D_f \times D_f$  次,单次卷积计算量为  $D_k \times D_k$ ,因此,深度卷积的运算总数为  $M \times D_f^2 \times D_k^2$ 。

逐点卷积阶段:卷积核尺寸为  $1 \times 1 \times M$  (覆盖全部输入通道、空间维度为  $1 \times 1$ ),共配置  $N$  个核以输出  $N$  个通道;同样要滑动  $D_f \times D_f$  次,单次卷积运算为  $1 \times 1 \times M$ ,因此,逐点卷积的总乘法运算量为  $M \times D_f^2 \times N$ 。深度可分离卷积和普通卷积的计算量之比为

$$\frac{D_k \cdot D_k \cdot M \cdot D_f \cdot D_f + M \cdot N \cdot D_f \cdot D_f}{D_k \cdot D_k \cdot M \cdot N \cdot D_f \cdot D_f} = \frac{1}{N} + \frac{1}{D_k^2} \quad (3)$$

在深度卷积与点卷积后分别加入 BN 层与 ReLU 激活函数,既降低计算量,又增强模型的非线性拟合能力。

2) 残差结构设计

随着网络深度增加,梯度在传播过程中可能会出现梯度爆炸或消失现象,训练误差也会增大,导致训练精度下降。为了解决此问题,在增加网络深度时引入残差结构来动态调整模型的非线性拟合能力。通过在浅层网络结构中并联  $y = x$  的恒等映射层,把前层卷积结果与后层卷积结果逐元素相加,在不增加参数量的前提下,减少特征传递过程中的信息损失,进而提升网络的精度,并在网络过深时抑制退化问题。

2.5 线性分类模块

特征提取完成后,将提取到的高光谱图像高级特征输入到由 3 个全连接层组成的线性分类单元,将特征映射到样本标签,完成最终的地物分类。在线性分类单元中采用 Dropout 层(丢弃率为 0.5)来防止过拟合。在整个模型中,为提高模型拟合能力和分类精度,所有激活函数均采用 LeakyReLU 函数,通过引入一个固定负半轴斜率(本文设为 0.01),扩大激活函数的有效输入范围,解决 ReLU 在负输入时的梯度消失问题,使负输入神经元也可以参与参数更新,从而进一步提高模型拟合能力和分类精度。

3 实验结果与分析

为了验证 AMLW-CNN 模型的性能,实验设计如下:在 Indian Pines、Salinas、Pavia U 3 个数据集上进行消融实验,以验证多尺度提取模块和残差结构的有效性;对比 AMLW-CNN 和传统卷积模型的网络参数以及分类精度;再将 AMLW-CNN 和 SVM、KNN、1D-CNN、2D-CNN、3D-CNN、CNN-ASS<sup>[29]</sup> 以及无注意力机制的 LW-CNN 模型的性能进行对比。

3.1 实验数据集

Indian Pines 数据集:由 AVIRIS 机载红外成像光谱仪于 1992 年在美国印第安纳州农业区拍摄,是第一个公开应用于高光谱图像分类的数据集,AVIRIS 成像光谱仪的空间分辨率为  $20 \text{ m} \cdot \text{pixel}^{-1}$ 。原始波段数为 220。剔除吸水波段(104 ~ 108、150 ~ 160、220),保留 200 个波段。图像空间尺寸为  $145 \times 145$  像素,总计 21 025 个像素,其中 10 249 个为已标记样本点,包含 16 类地物,如玉米、大豆、燕麦等

粮食作物。由于样本之间的差别较小,光谱曲线比较接近,而且样本点分布比较混乱,分类任务挑战较大。

**Salinas 数据集:**由 AVIRIS 在美国加利福尼亚州山谷拍摄,空间分辨率为  $3.7 \text{ m} \cdot \text{pixel}^{-1}$ 。原始波段数 224,剔除 108 ~ 112、154 ~ 167、224 等噪声波段,保留 204 个波段。图像空间尺寸为  $512 \times 217$  像素,共包含 111 104 个像素,其中有 54 129 个样本点被标记,包含 16 类地物,如耕地、芹菜等。

**Pavia U 数据集:**由德国的机载成像光谱仪在意大利帕维亚城市区域拍摄。原始波段数为 115,剔除 12 个噪声波段,保留 103 个波段。图像空间尺寸为  $610 \times 340$  像素,共包含 207 400 个像素,其中有 42 776 个像素被标记为样本点,包括 9 类地物,如道路、牧场、树木等。

### 3.2 评价指标

采用全局分类精度(OA)、平均分类精度(AA)与 Kappa 系数评估模型性能。

**Kappa 系数:**基于分类混淆矩阵来衡量分类精度,公式为:

$$P = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,N_c} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,N_c} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N_c,1} & p_{N_c,2} & \cdots & p_{N_c,N_c} \end{bmatrix} \quad (4)$$

$$\text{Kappa} = \frac{N \sum_{m=1}^{N_c} p_{m,m} - \sum_{m=1}^{N_c} p_{m,+} \cdot p_{+,m}}{N^2 - \sum_{m=1}^{N_c} p_{m,+} \cdot p_{+,m}}$$

式中: $N$  为所有地表真实分类中的像元总数; $p_{m,m}$  为混淆矩阵的对角线元素(第  $m$  类正确分类的像元数); $p_{m,+}$  为第  $m$  类被分类的像元总数; $p_{+,m}$  为第  $m$  类真实像元总数。

**全局分类精度(OA):**正确分类的像素数目与被标记的样本点的总数的比值,公式为

$$\text{OA} = \frac{\sum_{m=1}^n w_m}{\sum_{m=1}^n N_m} \quad (5)$$

式中: $n$  表示图像中的类别数目; $N_m$  为第  $m$  类实际像元数目; $w_m$  为正确分类的第  $m$  类像元个数。

**平均分类精度(AA):**所有类别分类精度的平均值,公式为

$$\text{AA} = \frac{1}{n} \sum_{m=1}^n \frac{p_{m,m}}{N_m} \quad (6)$$

### 3.3 实验平台

实验运行环境为 Python 3.6 + PyTorch 框架,使用 GPU 加速模型训练;硬件配置为 Intel® Core™ i9-11900 处理器、NVIDIA GeForce RTX 3090 GPU;使用 Adam 优化器,所有实验都基于此平台完成。

### 3.4 实验参数设置

消融实验中,选取各数据集 10% 的样本点作为训练样本,其余为测试样本;SVM 和 KNN 模型选用 70% 的样本进行训练;1D-CNN 和 2D-CNN 选用 30% 的样本进行训练;3D-CNN、CNN-ASS、LW-CNN 和 AMLW-CNN 选用 10% 的样本训练。

在训练模型阶段,影响训练时间和分类精度的主要因素包括学习率、卷积层层数和卷积核尺寸。学习率决定了目标函数能否收敛到最小值以及收敛速率。合适的学习率可加速训练过程,促进目标函数更快收敛。根据实验结果,将学习率设置为 0.005。在空谱特征提取阶段,单个多尺度提取模块的特征提取能力有限,导致特征提取不充分,但是过多的模块堆叠会增加模型的参数量,带来巨大的算力消耗,同时过深的网络也会导致模型过拟合,造成分类精度下降。因此,本文采用 2 个连续的多尺度卷积模块以有效提取空谱特征。多尺度提取模块的卷积核尺寸设置为  $1 \times 1 \times 1$ 、 $3 \times 3 \times 3$ 、 $5 \times 5 \times 5$  的 3 层独立并行结构,适配不同数据集尺寸。另外,根据实验结果,单个或者不叠加残差块时模型的分类效果不佳,然而叠加过多的残差块会导致模型过拟合,从而使分类精度下降,且过多的残差块会增加模型计算量,带来不必要的算力消耗。因此,空间特征提取模块叠加 2 个残差结构,以减少数据向后传播过程中的损失,提升模型的拟合能力。

### 3.5 消融实验

#### 1) 多尺度提取模块的影响

保持 AMLW-CNN 网络模型其他模块不变,分别对比不加入多尺度提取模块、加入 1 个多尺度提取模块、连续 2 个多尺度提取模块的模型分类结果。其中,不加入多尺度提取模块的模型,将 2 个多尺度提取模块都替换为卷积核尺寸  $3 \times 3 \times 3$  的卷积层;加入 1 个多尺度提取模块的模型,将第 1 个多尺度提取模块替换为卷积核尺寸  $3 \times 3 \times 3$  的卷积层,第 2 个保持不变。实验结果如表 1 所示,3 个数据集上均呈现一致规律:不加入多尺度提取模块的模型分类精度最低,将普通三维卷积层替换为多尺度提取模块后,分类精度有所提升。这说明多尺度提取模

块的特征提取能力较普通卷积层更强,可以更充分利用输入图像的不同局部结构,提取更丰富的特征信息,细化分类结果。

表1 不同数量多尺度提取模块的模型分类精度

Tab.1 Classification accuracy of models with different numbers of multiscale extraction modules

配置模型	Indian Pines 数据集	Pavia U 数据集	Salinas 数据集
无多尺度 提取模块	OA = 0.966	OA = 0.989	OA = 0.993
	AA = 0.958	AA = 0.978	AA = 0.991
	Kappa = 0.961	Kappa = 0.986	Kappa = 0.992
1个多尺度 提取模块	OA = 0.975	OA = 0.994	OA = 0.996
	AA = 0.968	AA = 0.986	AA = 0.995
	Kappa = 0.972	Kappa = 0.992	Kappa = 0.995
2个多尺度 提取模块	OA = 0.985	OA = 0.999	OA = 0.998
	AA = 0.984	AA = 0.997	AA = 0.999
	Kappa = 0.982	Kappa = 0.999	Kappa = 0.998

2) 残差模块的影响

保持网络其他模块不变,对比不加入残差模块、加入1个残差模块、加入2个残差模块的网络分类结果。实验结果如表2所示:在Indian Pines和Pavia U数据集上,无残差模块的模型总体分类精度较低,加入1个残差模块或2个残差模块后,总体分类精度显著提升;在Salinas数据集上,无残差模块的模型分类精度较低,加入1个残差模块后,分类精度有所提升,加入2个残差模块时分类精度和加入1个残差模块的模型保持一致。分析可知,Salinas数据集相较于Indian Pines和Pavia U数据集,光谱分辨率和空间分辨率都较高,因此样本点更多、信息更丰富,

表3 AMLW-CNN模型各模块参数设置

Tab.3 Parameter settings of each module of the AMLW-CNN model

模块类别	卷积类型/步长	卷积核尺寸	输入数据尺寸	参数量
预处理模块	PCA + 裁剪	—	145 × 145 × 200	—
三维卷积层	3D-Conv1/1 × 1 × 1	3 × 3 × 7	4 × 1 × 30 × 25 × 25	1 024
多尺度特征提取模块 × 2	3D-Conv2_1/1 × 1 × 1	1 × 1 × 1	4 × 16 × 24 × 23 × 23	272
	3D-Conv2_2_1/1 × 1 × 1	3 × 1 × 1	4 × 16 × 24 × 23 × 23	784
	3D-Conv2_2_2/1 × 1 × 1	1 × 1 × 3	4 × 16 × 24 × 23 × 23	784
	3D-Conv2_2_3/1 × 1 × 1	1 × 3 × 1	4 × 16 × 24 × 23 × 23	784
	3D-Conv2_3_1/1 × 1 × 1	5 × 1 × 1	4 × 16 × 24 × 23 × 23	1 296
	3D-Conv2_3_2/1 × 1 × 1	1 × 1 × 5	4 × 16 × 24 × 23 × 23	1 296
	3D-Conv2_3_3/1 × 1 × 1	1 × 5 × 1	4 × 16 × 24 × 23 × 23	1 296
通道注意力机制模块	Avg Pool	23 × 23	4 × 384 × 23 × 23	0
	Max Pool	23 × 23	4 × 384 × 23 × 23	0
	2D-Conv/1	1 × 1	4 × 384 × 1 × 1	9 216
	2D-Conv/1	1 × 1	4 × 24 × 1 × 1	9 216

当网络特征提取能力达到一定限度时,新增残差模块无法进一步提升分类性能。综上,残差模块通过残差信号优化,可以显著提高网络拟合能力,提升分类精度,克服有限训练样本下的网络性能不佳问题。

表2 不同数量残差模块的模型分类精度

Tab.2 Classification accuracy of the models with different number of residual modules

配置模型	Indian Pines 数据集	Pavia U 数据集	Salinas 数据集
无残差模块	OA = 0.972	OA = 0.990	OA = 0.996
	AA = 0.952	AA = 0.983	AA = 0.994
	Kappa = 0.968	Kappa = 0.987	Kappa = 0.995
1个残差模块	OA = 0.976	OA = 0.997	OA = 0.998
	AA = 0.972	AA = 0.994	AA = 0.999
	Kappa = 0.972	Kappa = 0.996	Kappa = 0.998
2个残差模块	OA = 0.985	OA = 0.999	OA = 0.998
	AA = 0.984	AA = 0.997	AA = 0.999
	Kappa = 0.982	Kappa = 0.999	Kappa = 0.998

3.6 对比实验分析

1) 模型复杂度对比

输入数据尺寸为145 × 145 × 200,经过预处理模块的降维和裁剪后,变成30 × 25 × 25大小的块;设置训练batch\_size为4,则输入三维卷积层的数据尺寸为4 × 1 × 30 × 25 × 25。依次输入三维卷积层、2个多尺度特征提取模块、通道注意力机制模块、空间注意力机制模块、空间特征提取模块进行特征提取,最后将网络输入分类单元进行分类输出。各个模块的卷积层参数如表3所示。

表 3(续)

模块类别	卷积类型/步长	卷积核尺寸	输入数据尺寸	参数量
空间注意力机制模块	Avg Pool	384 × 1	4 × 384 × 23 × 23	0
	Max Pool	384 × 1	4 × 384 × 23 × 23	0
	2D-Conv/1	7 × 7	4 × 2 × 23 × 23	98
空间特征提取模块	deep_Conv/1	5 × 5	4 × 384 × 23 × 23	9 600
	point_Conv/1	1 × 384	4 × 384 × 19 × 19	24 576
	(deep_Conv2/1	3 × 3	4 × 64 × 19 × 19	576
	point_Conv/1) × 5	1 × 64	4 × 64 × 19 × 19	4 096
分类单元	—	—	4 × 23 104	—
输出结果	—	—	4 × 16	—

表 4 传统卷积模型各模块参数设置

Tab. 4 Parameter setting of each module of the traditional model

模块类别	类型/步长	卷积核尺寸	输入数据尺寸	参数量
预处理模块	PCA + 裁剪	—	145 × 145 × 200	—
三维卷积层	3D-Conv1/1 × 1 × 1	3 × 3 × 7	4 × 1 × 30 × 25 × 25	1 024
多尺度特征提取模块 × 2	3D-Conv2_1/1 × 1 × 1	1 × 1 × 1	4 × 16 × 24 × 23 × 23	272
	3D-Conv2_2/1 × 1 × 1	3 × 3 × 3	4 × 16 × 24 × 23 × 23	6 928
	3D-Conv2_3/1 × 1 × 1	5 × 5 × 5	4 × 16 × 24 × 23 × 23	32 016
通道注意力机制模块	Avg Pool	23 × 23	4 × 384 × 23 × 23	0
	Max Pool	23 × 23	4 × 384 × 23 × 23	0
	2D-Conv/1	1 × 1	4 × 384 × 1 × 1	9 216
	2D-Conv/1	1 × 1	4 × 24 × 1 × 1	9 216
空间注意力机制模块	Avg Pool	384 × 1	4 × 384 × 23 × 23	0
	Max Pool	384 × 1	4 × 384 × 23 × 23	0
	2D-Conv/1	23 × 23	4 × 2 × 23 × 23	98
空间特征提取模块	2D-Conv/1	5 × 5	4 × 384 × 23 × 23	61 464
	(2D-Conv/1) × 5	3 × 3	4 × 64 × 19 × 19	36 928
分类单元	—	—	4 × 23 104	—
输出结果	—	—	4 × 16	—

为了突出 AMLW-CNN 的轻量化特点,将其与没有应用非对称卷积和深度可分离卷积的传统卷积模型的参数量进行对比,传统卷积模型各模块参数如表 4 所示。对比表 3、表 4 的数据可知,2 个模型的三维卷积层和注意力机制模块的卷积层参数量相等,多尺度提取模块和空间特征提取模块参数量存在差异。计算可得,AMLW-CNN 2 个连续多尺度提取模块的参数量合计为 13 024,空间特征提取模块参数量合计为 57 536;传统卷积模型的 2 个多尺度提取模块参数量为 39 216,空间特征提取模块的参数量为 246 104。通过对比可知,AMLW-CNN 的

卷积层参数量比传统卷积模型少了 214 760。参数量的减少不仅可以节约计算资源,还能提高训练效率。

## 2) 分类精度对比

AMLW-CNN 和传统卷积模型分类精确度对比结果如表 5 所示,在 Indian Pines、Pavia U 和 Salinas 数据集上,AMLW-CNN 的分类精度和传统卷积模型相比都有提升。分析可得,相对于传统卷积模型,非对称卷积和深度可分离卷积加深了 AMLW-CNN 的网络深度,提升了模型拟合能力,使其能更好地提取数据深层次的特征。

表 5 传统卷积模型和 AMLW-CNN 模型的分类精度

Tab.5 Classification accuracy of traditional convolutional models and AMLW-CNN model

模型类别	Indian Pines 数据集	Pavia U 数据集	Salinas 数据集
传统卷积模型	OA = 0. 979	OA = 0. 997	OA = 0. 997
	AA = 0. 964	AA = 0. 991	AA = 0. 994
	Kappa = 0. 978	Kappa = 0. 997	Kappa = 0. 997
AMLW-CNN 模型	OA = 0. 985	OA = 0. 999	OA = 0. 998
	AA = 0. 984	AA = 0. 997	AA = 0. 999
	Kappa = 0. 982	Kappa = 0. 999	Kappa = 0. 998

AMLW-CNN 和其他对比算法的分类结果如表 6、表 7、表 8 所示,分类标签和结果如图 7、图 8、图 9 所示。由表可知,由于高光谱图像空间分辨率比较低,训练样本不足,SVM、KNN 等传统的机器学习方法特征提取能力有限,不能有效提取数据的高

级特征,因此分类效果较差,鲁棒性也较差,而且只对特定数据集的效果明显,模型泛化能力较差。1D-CNN 仅提取光谱特征,特征利用不充分,分类精度最低,分类结果图噪声点较多。2D-CNN 只提取高光谱图像的空间特征,且需要对高光谱图像进行降维处理,不能有效提取其光谱特征,分类精度低于 3D-CNN 和 AMLW-CNN。3D-CNN 由于可以充分提取高光谱图像的空谱特征,所以分类结果相较于传统机器学习和 1D-CNN、2D-CNN 表现较好,分类结果图中的噪声点也少了很多,但是其网络结构复杂、参数量巨大,计算成本高。AMLW-CNN 相比 3D-CNN 在 Indian Pines、Pavia U、Salinas 数据集上的总体分类精度分别提高了 4.6%、5.6%、2.5%;与 CNN-ASS 相比,AMLW-CNN 的全局分类精度和平均分类精度更高,分类结果和原始标签更为接近,图像噪声点更少,地物分类更加精确。

表 6 不同方法在 Indian Pines 数据集上的分类精度

Tab.6 Classification accuracy of different methods on the Indian Pines dataset

类别	SVM	KNN	1D-CNN	2D-CNN	3D-CNN	CNN-ASS	LW-CNN	AMLW-CNN
Alfalfa	0	0.67	0	<b>1</b>	0.95	<b>1</b>	<b>1</b>	0.93
Corn-notill	0.69	0.69	0.64	0.951	0.93	0.98	0.98	<b>1</b>
Corn-mintill	0.68	0.67	0.66	0.57	0.92	0.94	0.94	<b>0.98</b>
Corn	0.48	0.57	0.62	0.6	0.91	<b>0.99</b>	0.92	<b>0.99</b>
Grass-pasture	0.92	0.88	0.44	0.74	0.95	0.97	0.99	1
Grass-trees	0.91	0.89	0.87	0.99	0.99	<b>1</b>	0.99	0.99
Grass-pasture-mowed	0	0.88	0	<b>1</b>	0.95	<b>1</b>	0.93	1
Hay-windrowed	0.92	0	0.94	0.7	0.99	<b>1</b>	0.99	1
Oats	0	0.36	0	0.94	0.6	<b>1</b>	0.91	0.95
Soybean-notill	0	0.72	0.64	0.91	0.9	<b>0.97</b>	<b>0.97</b>	<b>0.97</b>
Soybean-mintill	0.79	0.79	0.75	0.93	0.95	0.97	<b>0.98</b>	0.97
Soybean-clean	0.62	0.68	0.76	0.8	0.93	<b>0.99</b>	0.97	0.97
Wheat	0.91	0.92	0.88	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>
Woods	0.93	0.94	0.86	0.84	0.97	0.99	0.98	<b>1</b>
Buildings-Grass-Trees-Drives	0.64	0.51	0.51	0.36	0.87	0.86	0.93	<b>1</b>
Stone-Steel-Towers	0.88	0.98	<b>1</b>	0.96	0.99	0.96	0.98	0.88
OA	0.787	0.786	0.734	0.759	0.939	0.962	0.973	<b>0.985</b>
AA	0.586	0.697	0.598	0.831	0.924	0.981	0.966	<b>0.984</b>
Kappa	0.753	0.755	0.693	0.733	0.931	0.957	0.971	<b>0.982</b>

表 7 不同方法在 Salinas 数据集上的分类精度

Tab.7 Classification accuracy of different methods on the Salinas dataset

类别	SVM	KNN	1D-CNN	2D-CNN	3D-CNN	CNN-ASS	LW-CNN	AMLW-CNN
Broccoli_green_weeds_1	0.99	0.99	0.95	0.98	<b>1</b>	0.98	<b>1</b>	<b>1</b>
Broccoli_green_weeds_2	0.99	0.99	0.97	0.99	<b>1</b>	<b>1</b>	<b>1</b>	0.99
Fallow	0.97	0.99	0.85	0.97	0.99	0.99	<b>1</b>	<b>1</b>
Fallow_rough_plow	0.98	0.99	0.99	0.99	0.99	0.98	0.99	<b>1</b>
Fallow_smooth	0.98	0.99	0.91	0.98	0.99	0.99	0.98	<b>1</b>

表 7(续)

类别	SVM	KNN	1D-CNN	2D-CNN	3D-CNN	CNN-ASS	LW-CNN	AMLW-CNN
Stubble	0.99	0.99	0.99	<b>1</b>	<b>1</b>	0.99	<b>1</b>	<b>1</b>
Celery	0.99	0.99	0.98	0.99	<b>1</b>	0.99	<b>1</b>	<b>1</b>
Grapes_untrained	0.81	0.85	0.79	0.81	0.94	0.94	0.99	<b>1</b>
Soil_vinyard_develop	0.98	0.99	0.98	0.99	0.99	0.99	<b>1</b>	<b>1</b>
Corn_senesced_green_weeds	0.89	0.95	0.87	0.9	0.99	0.97	0.99	<b>1</b>
Lettuce_romaine_4wk	0.92	0.97	0.87	0.91	0.99	0.98	0.99	<b>1</b>
Lettuce_romaine_5wk	0.98	0.99	0.97	0.98	<b>1</b>	0.98	0.99	<b>1</b>
Lettuce_romaine_6wk	0.96	0.98	0.96	0.95	0.99	0.98	<b>1</b>	<b>1</b>
Lettuce_romaine_7wk	0.95	0.95	0.95	0.96	0.99	0.98	0.99	<b>1</b>
Vinyard_untrained	0.59	0.76	0.68	0.6	0.91	0.88	0.99	<b>1</b>
Vinyard_vertical_trellis	<b>0.99</b>	<b>0.99</b>	0.87	<b>0.99</b>	<b>0.99</b>	0.92	<b>0.99</b>	<b>0.99</b>
OA	0.895	0.929	0.878	0.898	0.973	0.952	0.996	<b>0.998</b>
AA	0.935	0.961	0.911	0.937	0.985	0.974	0.997	<b>0.999</b>
Kappa	0.882	0.922	0.864	0.886	0.976	0.957	0.997	<b>0.998</b>

表 8 不同方法在 Pavia U 数据集上的分类精度

Tab. 8 Classification accuracy of different methods on the Pavia U dataset

类别	SVM	KNN	1D-CNN	2D-CNN	3D-CNN	CNN-ASS	LW-CNN	AMLW-CNN
Asphalt	0.88	0.93	0.92	0.9	0.97	0.97	0.99	<b>1</b>
Meadows	0.94	0.94	0.91	0.94	0.95	0.98	0.99	<b>1</b>
Gravel	0.64	0.78	0.47	0.69	0.93	0.94	0.97	<b>1</b>
Trees	0.94	0.93	0.91	0.94	0.96	0.99	0.99	<b>1</b>
Painted metal sheets	0.99	0.99	0.99	0.99	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>
Bare Soil	0.77	0.79	0.65	0.78	0.98	0.96	0.99	<b>1</b>
Bitumen	0.54	0.84	0.83	0.69	0.97	0.96	<b>0.99</b>	<b>0.99</b>
Self-Blocking Bricks	0.84	0.86	0.79	0.85	0.97	0.94	0.98	<b>1</b>
Shadows	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>
OA	0.888	0.909	0.855	0.898	0.943	0.971	0.989	<b>0.999</b>
AA	0.843	0.896	0.831	0.864	0.971	0.973	0.988	<b>0.997</b>
Kappa	0.849	0.878	0.804	0.862	0.926	0.968	0.986	<b>0.999</b>

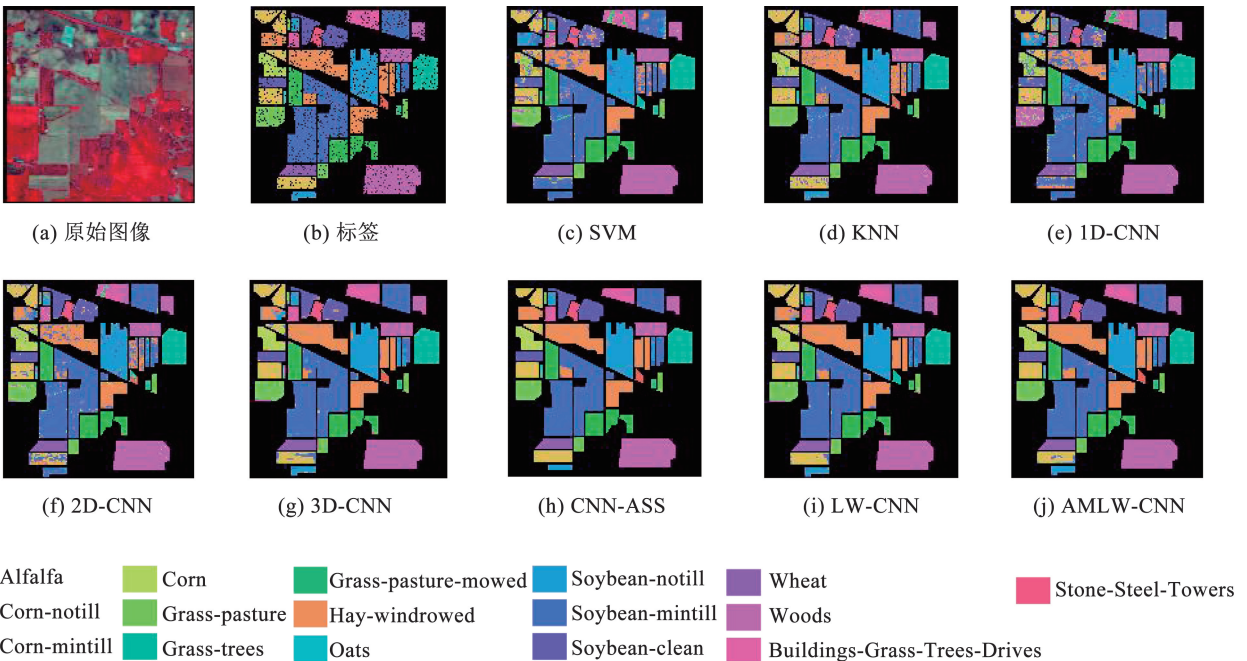


图 7 Indian Pines 分类结果图

Fig. 7 The Indian Pines classification results

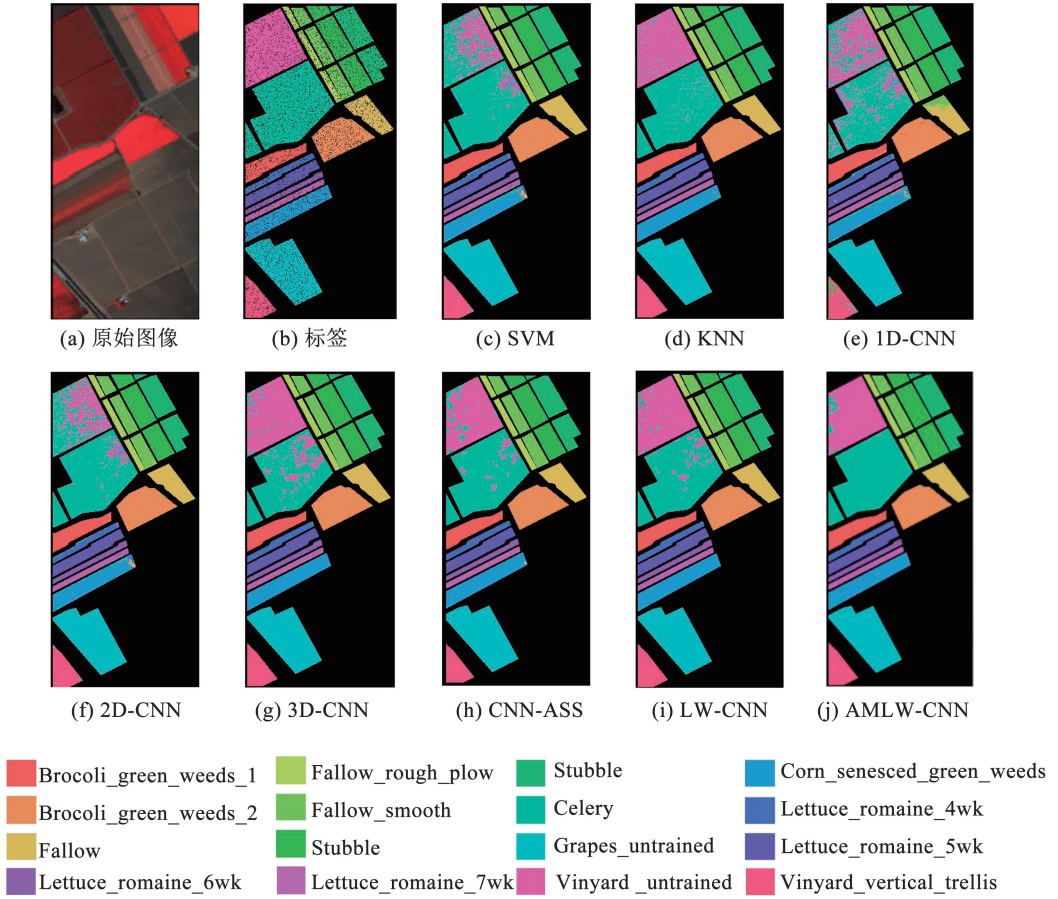


图8 Salinas 分类结果图

Fig. 8 The Salinas classification results

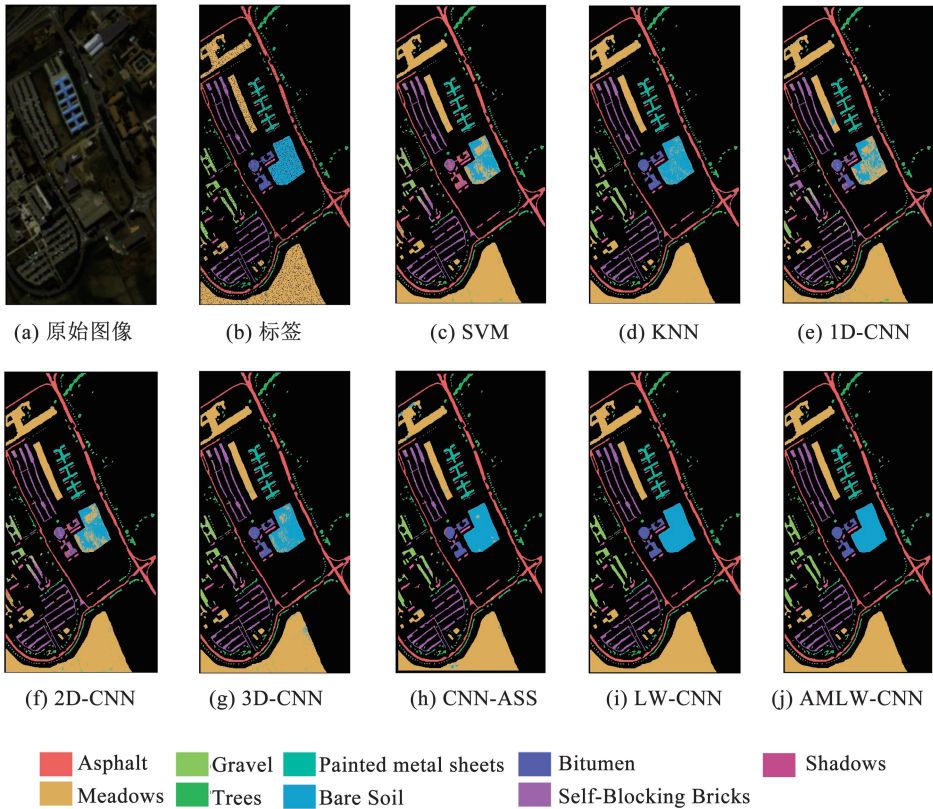


图9 Pavia U 分类结果图

Fig. 9 The Pavia U classification results

分析可得,对于其他方法而言,高光谱图像中部分地物样本点的光谱曲线非常接近,当训练样本不充分、模型对特征的提取能力较弱时就容易造成错分,进而影响分类精度,导致分类结果图中的噪声点显著增多。本文方法能够更充分地结合空间特征来提高分类精度:一方面,高光谱图像中单个样本点的周边样本点更大概率属于同一类别,可通过空间关联性辅助类别判断;另一方面,对于非同类的样本点,本方法可更充分地利用光谱特征差异实现有效区分,最终达成优异的分类效果。

为验证注意力机制的作用,对比了添加注意力机制和没有添加注意力机制模块的算法分类结果,显然,由于本方法在空谱特征提取模块后添加了通道注意力机制(CAM)和空间注意力机制(SAM)模块,分别对输入数据  $x$  的不同通道和不同空间区域,根据其分类结果的重要程度赋予对应权重  $ca(x)$ 、 $sa(x)$ ,具体权重计算如式(7)、式(8)所示:

$$ca(x) = \begin{bmatrix} 0.5152 & 0.5117 & 0.5132 & 0.5130 \\ 0.4771 & 0.4822 & 0.4786 & 0.4819 \\ 0.4833 & 0.4875 & 0.4845 & 0.4830 \\ \vdots & \vdots & \vdots & \vdots \\ 0.5259 & 0.5284 & 0.5265 & 0.5277 \\ 0.4829 & 0.4759 & 0.4813 & 0.4817 \\ -0.5036 & 0.5057 & 0.5063 & 0.5029 \end{bmatrix} \quad (7)$$

$$sa(x) = \begin{bmatrix} 0.4875 & 0.4866 & 0.5044 & \cdots & 0.5142 & 0.5118 & 0.5142 \\ 0.4796 & 0.4820 & 0.5100 & \cdots & 0.5122 & 0.5052 & 0.5063 \\ 0.4836 & 0.4730 & 0.5071 & \cdots & 0.5181 & 0.5095 & 0.5153 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0.4806 & 0.4599 & 0.4870 & \cdots & 0.4892 & 0.4924 & 0.5027 \\ 0.4833 & 0.4602 & 0.4738 & \cdots & 0.4796 & 0.4814 & 0.4888 \\ 0.4640 & 0.4523 & 0.4665 & \cdots & 0.4724 & 0.4817 & 0.4907 \end{bmatrix} \quad (8)$$

由于增大了样本间的类间方差,对空谱特征的提取更加充分,分类精度更高。而且相较传统 3D-CNN 模型,AMLW-CNN 引入了非对称卷积和深度可分离卷积,计算复杂度更低、运行效率更高;通过残差结构加深了网络层的数量,对高光谱图像深层次的特征提取更充分,对不同的光谱数据都有更好的适用性,模型的鲁棒性更强。式(7)中,输入通道注意力机制的数据尺寸为  $4 \times 384 \times 23 \times 23$ ,所以通道注意力机制权重矩阵的数据尺寸为  $384 \times 4$ ,用来对 4 个 batch 的 384 个通道分别赋予不同的权重;式(8)中,输入数据尺寸为  $4 \times 384 \times 23 \times 23$ ,空间注意力机制权重矩阵的尺寸为  $4 \times 23 \times 23$ ,本文选取其中一个 batch 的权重进行展示。

## 4 结 语

针对高光谱图像的自身特性和当前卷积神经网络对空谱特征利用不充分、参数量过大的问题,本文结合 3D-CNN、2D-CNN、多尺度提取模块、注意力机制、深度可分离卷积与残差网络,提出了一种基于注意力机制的多尺度轻量化卷积神经网络模型(AMLW-CNN)。本文的主要贡献在于设计了空谱特征提取能力更优的模型,将参数量庞大的模型进行了参数和结构优化,而且引入了注意力机制,降低了计算复杂度,增强了模型对高光谱图像综合特征的提取能力,且具有良好的鲁棒性。在 3 个高光谱数据集上的实验结果验证了本文所提方法的有效性。在不同的数据集上,本文算法的分类精度都优于对比算法,达到了良好的分类效果。本文方法仍有改进空间,目前在监督学习方向取得了良好的效果,后续研究将向半监督方向拓展,并采用迁移学习、强化学习方法对网络进行更细致的优化。

## 参考文献

- [1] AUDEBERT N, LE SAUX B, LEFÈVRE S. Deep learning for classification of hyperspectral data: A comparative review[J]. IEEE Geoscience and Remote Sensing Magazine, 2019, 7(2): 159. DOI: 10.1109/MGRS.2019.2912563
- [2] 叶珍,白璠,何明一. 高光谱图像空谱特征提取综述[J]. 中国图象图形学报, 2021, 26(8): 1737  
YE Zen, BAI Lin, HE Mingyi. Review of spatial-spectral feature extraction of hyperspectral image[J]. Journal of Image and Graphics, 2021, 26(8): 1737. DOI: 10.11834/jig.210198
- [3] 孔毅,纪定哲,程玉虎,等. 基于光谱注意力图卷积网络的高光谱图像分类[J]. 电子与信息学报, 2022, 44:1  
KONG Yi, JI Dingzhe, CHENG Yuhu, et al. Hyperspectral image classification based on spectral attentional graph convolution network[J]. Journal of Electronics and Information Science, 2022, 44: 1. DOI: 10.11999/JEIT220204
- [4] 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. DOI: 10.1109/TGRS.2004.831865
- [5] LI W, PRASAD S, FOWLER J E, et al. Locality-preserving dimensionality reduction and classification for hyperspectral image analysis[J]. IEEE Transactions on Geoscience and Remote Sensing, 2011, 50(4): 1185. DOI: 10.1109/TGRS.2011.2165957
- [6] BLANZIERI E, MELGANI F. Nearest neighbor classification of remote sensing images with the maximal margin principle[J]. IEEE Transactions on Geoscience and Remote Sensing, 2008, 46(6): 1804. DOI: 10.1109/TGRS.2008.916090
- [7] LI J, BIUCAS-DIAS J M, PLAZA A. Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields[J]. IEEE Transactions on Geoscience and Remote Sensing, 2011, 50(3): 809. DOI: 10.1109/TGRS.2011.2162649

- [8] SHI C, PUN C M. Multiscale super pixel-based hyperspectral image classification using recurrent neural networks with stacked autoencoders[J]. IEEE Transactions on Multimedia, 2019, 22(2): 487. DOI: 10.1109/TMM.2019.2928491
- [9] FANG Z, ROY K, MARES J, et al. Deep learning-based axial capacity prediction for cold formed steel channel sections using Deep Belief Network [J]. Structures, 2021; 2792. DOI: 10.1016/j.istruc.2021.05.096
- [10] GUO Y, HAN S, CAO H, et al. Guided filter based deep recurrent neural networks for hyperspectral image classification[J]. Procedia Computer Science, 2018, 129: 219. DOI:10.1016/j.procs.2018.03.048
- [11] ZHANG M, LI W, DU Q. Diverse region-based CNN for hyperspectral image classification[J]. IEEE Transactions on Image Processing, 2018, 27(6): 2623. DOI: 10.1109/TIP.2018.2809606
- [12] PAOLETTI M E, MORENO-ÁLVAREZ S, HAUT J M. Multiple attention-guided capsule networks for hyperspectral image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60:1. DOI: 10.1109/TGRS.2021.3135506
- [13] ZHONG Z, LI J, CLAUSI D A, et al. Generative adversarial networks and conditional random fields for hyperspectral image classification[J]. IEEE Transactions on Cybernetics, 2019, 50(7): 3318. DOI: 10.1109/TCYB.2019.2915094
- [14] HONG D, GAO L, YAO J, et al. Graph convolutional networks for hyperspectral image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(7): 5966. DOI: 10.1109/TGRS.2020.3015157
- [15] ZHU M, JIAO L, LIU F, et al. Residual spectral-spatial attention network for hyperspectral image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(1): 449. DOI: 10.1109/TGRS.2020.2994057
- [16] ZHU K, CHEN Y, GHAMISI P, et al. Deep convolutional capsule network for hyperspectral image spectral and spectral-spatial classification[J]. Remote Sensing, 2019, 11(3): 223. DOI: 10.3390/rs11030223
- [17] ZHU L, CHEN Y, GHAMISI P, et al. Generative adversarial networks for hyperspectral image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(9): 5046. DOI: 10.1109/TGRS.2018.2805286
- [18] HONG D, HAN Z, YAO J, et al. SpectralFormer: Rethinking hyperspectral image classification with transformers [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1. DOI: 10.1109/TGRS.2021.3130716
- [19] LI R, ZHENG S, DUAN C, et al. Classification of hyperspectral image based on double-branch dual-attention mechanism network [J]. Remote Sensing, 2020, 12(3): 582. DOI: 10.3390/rs12030582
- [20] HU W, HUANG Y, WEI L, et al. Deep convolutional neural networks for hyperspectral image classification [J]. Journal of Sensors, 2015; 258619. DOI: 10.1155/2015/258619
- [21] XU L, ZHANG H, ZHAO M, et al. Integrating spectral and spatial features for hyperspectral image classification using low-rank representation [C]//2017 IEEE International Conference on Industrial Technology (ICIT). Malmö: IEEE, 2017; 1024. DOI: 10.1109/icit.2017.7915502
- [22] YANG J, ZHAO Y, CHAN J C W, et al. Hyperspectral image classification using two-channel deep convolutional neural network [C]//2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Beijing, China, 2016; 5079. DOI: 10.1109/IGARSS.2016.7730324
- [23] LI Y, ZHANG H, SHEN Q. Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network [J]. Remote Sensing, 2017, 9(1):67. DOI: 10.3390/rs9010067
- [24] HE M, LI B, CHEN H. Multi-scale 3D deep convolutional neural network for hyperspectral image classification [C]//2017 IEEE International Conference on Image Processing (ICIP). Beijing, China; IEEE, 2017; 3904. DOI: 10.1109/ICIP.2017.8297014
- [25] LEE H, KWON H. Contextual deep CNN based hyperspectral classification [C]//2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Fort Worth; IEEE, 2017; 3322. DOI: 10.1109/IGARSS.2016.7729859
- [26] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]//2018 European Conference on Computer Vision (ECCV). Cham; Springer, 2018; 3. DOI: 10.1007/978-3-030-01234-2\_1
- [27] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [C]//2015 International Conference on Machine Learning (ICML). New York: JMLR Workshop and Conference Proceedings, 2015; 448. DOI: 10.48550/arXiv.1502.03167
- [28] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications [EB/OL]. (2017-04-14) [2025-11-10]. <https://arxiv.org/abs/1704.04861>
- [29] ZHANG Y, YANG K, YUAN L. A hyperspectral image classification method with CNN based on attention enhanced spectral and spatial features [J]. Journal of Physics: Conference Series, 2021, 2006(1): 012033. DOI: 10.1088/1742-6596/2006/1/012033

(编辑 丁晓清)