



## 基于混合卷积与三重注意力的高光谱图像分类网络

王瑞婷, 王海燕, 陈晓, 耿信哲, 雷涛

引用本文:

王瑞婷,王海燕,陈晓,耿信哲,雷涛. 基于混合卷积与三重注意力的高光谱图像分类网络[J]. 智能系统学报, 2023, 18(2): 260–269.

WANG Ruiting, WANG Haiyan, CHEN Xiao, GENG Xinzhe, LEI Tao. Hyperspectral image classification based on hybrid convolutional neural network with triplet attention[J]. *CAAI Transactions on Intelligent Systems*, 2023, 18(2): 260–269.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202204002>

## 您可能感兴趣的其他文章

### 结合全局注意力机制的实时语义分割网络

Global attention mechanism with real-time semantic segmentation network

智能系统学报. 2023, 18(2): 282–292 <https://dx.doi.org/10.11992/tis.202208027>

### 面向表情识别的重影非对称残差注意力网络模型

A ghost asymmetric residual attention network model for facial expression recognition

智能系统学报. 2023, 18(2): 333–340 <https://dx.doi.org/10.11992/tis.202201003>

### 基于双向LSTM卷积网络与注意力机制的自动睡眠分期模型

Automatic sleep staging model based on the bi-directional LSTM convolutional network and attention mechanism

智能系统学报. 2022, 17(3): 523–530 <https://dx.doi.org/10.11992/tis.202103013>

### 改进MobileNet的图像分类方法研究

Research on the improved image classification method of MobileNet

智能系统学报. 2021, 16(1): 11–20 <https://dx.doi.org/10.11992/tis.202012034>

### 4D卷积神经网络的自闭症功能磁共振图像分类

Classification of the functional magnetic resonance image of autism based on 4D convolutional neural network

智能系统学报. 2021, 16(6): 1021–1029 <https://dx.doi.org/10.11992/tis.202009022>

### 层次化双注意力神经网络模型的情感分析研究

Hierarchical double-attention neural networks for sentiment classification

智能系统学报. 2020, 15(3): 460–467 <https://dx.doi.org/10.11992/tis.201812017>

DOI: 10.11992/tis.202204002

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20221009.1214.010.html>

# 基于混合卷积与三重注意力的高光谱图像分类网络

王瑞婷, 王海燕, 陈晓, 耿信哲, 雷涛

(陕西科技大学 电子信息与人工智能学院, 陕西 西安 710021)

**摘 要:** 针对高光谱图像光谱维度高、现有网络无法提供深度级的多层次特征, 从而影响分类精度和速度的问题。首先采用核主成分分析对高光谱图像进行降维, 使降维后的数据具有最佳区分度, 提出了一种基于混合卷积与三重注意力的卷积神经网络 (hybrid convolutional neural network with triplet attention, HCTA-Net) 模型, 该模型设计了一种基于三维、二维和一维卷积的混合卷积神经网络, 通过不同维度卷积神经网络的融合, 提取高光谱图像精细的光谱-空间联合特征。在二维卷积中加入深度可分离卷积, 减少了模型参数, 同时引入三重注意力机制, 使用三支结构实现跨维度信息交互, 抑制无用的特征信息。在 Indian Pines、Salinas 和 Pavia University 数据集上的实验结果表明, 本文提出的模型优于其他对比方法, 总体分类精度分别达到了 99.16%、99.87% 和 99.76%。

**关键词:** 遥感; 高光谱图像分类; 深度学习; 特征提取; 降维; 深度可分离卷积; 注意力机制; 混合卷积神经网络  
**中图分类号:** TP751   **文献标志码:** A   **文章编号:** 1673-4785(2023)02-0260-10

中文引用格式: 王瑞婷, 王海燕, 陈晓, 等. 基于混合卷积与三重注意力的高光谱图像分类网络 [J]. 智能系统学报, 2023, 18(2): 260-269.

英文引用格式: WANG Ruiting, WANG Haiyan, CHEN Xiao, et al. Hyperspectral image classification based on hybrid convolutional neural network with triplet attention[J]. CAAI transactions on intelligent systems, 2023, 18(2): 260-269.

## Hyperspectral image classification based on hybrid convolutional neural network with triplet attention

WANG Ruiting, WANG Haiyan, CHEN Xiao, GENG Xinzhe, LEI Tao

(School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, China)

**Abstract:** To solve the problems of a high spectral dimension of hyperspectral images and the failure of the existing network to provide multilevel features at the depth level, which affects the classification accuracy and speed, the kernel principal component analysis is used to reduce the dimensionality of hyperspectral images to have the best data differentiation after dimensionality reduction, and a hybrid convolutional neural network with triplet attention (HCTA-Net) model is proposed to design a hybrid model based on 3D, 2D, and 1D CNN to extract the fine spectral-spatial joint features through the fusion of different dimension convolutions. The model also adds depthwise separable convolution into the 2D-CNN to reduce the model parameters and simultaneously introduces a triplet attention mechanism, which uses a three-branch structure to achieve cross-dimensional information interaction to inhibit useless feature information. Experimental results on the Indian Pines, Salinas, and Pavia University datasets show that the proposed model is superior to other comparison methods, and the overall classification accuracy reaches 99.16%, 99.87, and 99.76%, respectively.

**Keywords:** remote sensing; hyperspectral image classification; deep learning; feature extraction; dimension reduction; depth-separable convolution; attention mechanism; hybrid convolutional neural network

高光谱图像 (hyperspectral images, HSI) 是由

一维光谱特征和二维空间信息组成的含有上百通道且波段连续的三维图像, 被广泛应用于国防安全、海洋勘探<sup>[1]</sup>及地物分类<sup>[2]</sup>等领域。但由于 HSI 波段多、空间分辨率低和 Hughes 现象<sup>[3]</sup>等现

收稿日期: 2022-04-02. 网络出版日期: 2022-10-10.

基金项目: 国家自然科学基金重点项目 (62031021).

通信作者: 王海燕. E-mail: [hywang@sust.edu.cn](mailto:hywang@sust.edu.cn).

实问题,使 HSI 分类问题仍面临很大的挑战,因此研究一个高效提升 HSI 分类性能的模型具有重要的应用价值。

HSI 分类需要同时考虑光谱和空间特征信息。传统的 HSI 分类方法主要侧重于研究连续光谱,利用可区分的光谱特征对数据进行分类,典型的分类器包括 K 近邻 (K-nearest neighbor algorithm, KNN)<sup>[4]</sup>、支持向量机 (support vector machine, SVM)<sup>[5]</sup>、字典学习<sup>[6]</sup>等,其中 KNN 分类器需要大量的训练样本, SVM 虽能缓解 Hughes 现象,但分类精度不佳,因此对于高维小样本的 HSI 数据集,这些分类方法效果往往并不理想。研究表明,空间信息和光谱特征的互补决策融合,有助于提高分类性能<sup>[7]</sup>,于是 Fauvel 等<sup>[8]</sup>提出集成决策树、Zhong 等<sup>[9]</sup>提出深度信念网络的分类方法用于决策融合,但由于特征表达能力有限,仍不能完全表示光谱-空间信息的丰富度,难以有效提高分类水平。

近年来,深度学习凭借其优异的特征提取能力,被广泛用于 HSI 分类领域<sup>[10]</sup>。Hu 等<sup>[11]</sup>首次将 1D-CNN 应用于 HSI 分类,分类结果略优于传统的分类方法,但会造成空间特征丢失的问题。Xu 等<sup>[12]</sup>提出一种双支 CNN 模型,将并行的 2D-CNN 与 1D-CNN 用于提取光谱-空间特征,通过简单的融合,提高了分类性能,但未考虑光谱-空间特征间的关联性。针对这一问题,Zhang 等<sup>[13]</sup>提出一种 3D 密集连接网络 SSDANet,直接对三维数据进行光谱-空间特征提取,充分发挥了 HSI“图谱合一”的特点,但随着模型深度的加深,细节特征将逐渐丢失,还带来了模型复杂、计算量大等问题。为此,Guo 等<sup>[14]</sup>提出一种深度协同网络 CACNN,将 3D-CNN 与 2D-CNN 用一种深度多层特征融合策略相结合用于 HSI 分类,得到较为满意的结果,但对光谱-空间相关特征仍不够关注。

此外,在提取特征进行分类时,研究者受人类视觉机制的启发,提出用注意力机制来抑制冗余信息,聚焦特定特征并分配相应权重,从而提高模型的分类性能。Hu 等<sup>[15]</sup>提出了压缩激活网络 (squeeze and excitation network, SENet),是第一个通过对通道特征进行加权建模的注意力机制。在此基础上,为了将空间与通道两个维度的特征结合起来重新进行校准,Woo 等<sup>[16]</sup>提出了卷积块注意力模块 (convolutional block attention module, CBAM),对输入的特征进行约束增强处理。另外,基于注意力机制的 CNN 模型在 HSI 分类领域

逐渐取得了较为显著的成果。如 Fang 等<sup>[17]</sup>提出了单注意力网络 MSDN-SA,提高了光谱波段间的相关性。为进一步加强空间信息特征,Roy 等<sup>[18]</sup>在  $A^2S^2K$ -ResNet 网络中引入基于注意力的自适应光谱-空间核模块,实现了光谱-空间信息的联合挖掘,增强了网络结构的特征表达能力。Ahmad 等<sup>[19]</sup>将三支注意力融合块引入 AfNet 结构,自适应地强调重要特征,显著抑制了冗余和无效信息。

针对现有高光谱分类模型较复杂、未充分考虑光谱-空间联合特征,无法提供深度级的多层次特征等问题,本文提出一种基于混合卷积与三重注意力网络模型 (hybrid convolution neural network with triplet attention, HCTA-Net),主要贡献包括以下 4 个方面:

1)设计了一个由 3D-CNN、改进的 2D-CNN 和 1D-CNN 组成的混合卷积神经网络结构,充分利用光谱-空间域与深空间域互补特征,提升整个网络分类性能。

2)在 2D-CNN 模块中加入深度可分离卷积 (depthwise separable convolution, DSC)<sup>[20]</sup>,降低了卷积的参数数量和计算量。

3)引入三支注意力融合块 Triplet Attention<sup>[21]</sup>,将注意力聚焦重要区域,以捕获更重要的光谱-空间特征作为后续网络的输入。

4)利用 3 个评价指标对 Indian Pines、Salinas 和 Pavia University 3 个公开的 HSI 数据集的分类结果进行综合评估,实验结果表明本文方法均优于比较方法。

## 1 HCTA-Net 模型设计

本文提出的基于混合卷积模块与三重注意力机制的高光谱图像分类方法 HCTA-Net,整体流程如图 1 所示,主要由三部分组成:数据降维、分类识别网络和结果预测。首先,采用核主成分分析方法 (kernel principal component analysis, KPCA)<sup>[22]</sup>将维度为  $C \times H \times W$  ( $C$  为波段数,  $W$ 、 $H$  分别为单个波段的高和宽)的原始 HSI 数据降维到  $B$  个波段,有效地消除数据冗余并保留主要空谱信息,减少后续图像处理的时间和空间复杂度。然后选取大小为  $K \times K \times B$  的三维图像块输入到分类识别网络中,其中  $K$  为窗口大小。其次,分类识别网络包括依次连接的 3D-CNN、Triplet Attention、改进后的 2D-CNN 和 1D-CNN 模块构成的混合卷积网络,提取更强的频谱空间信息和重要的通道特征用于分类识别,最后使用 softmax 分类器得到最终的预测分类结果。

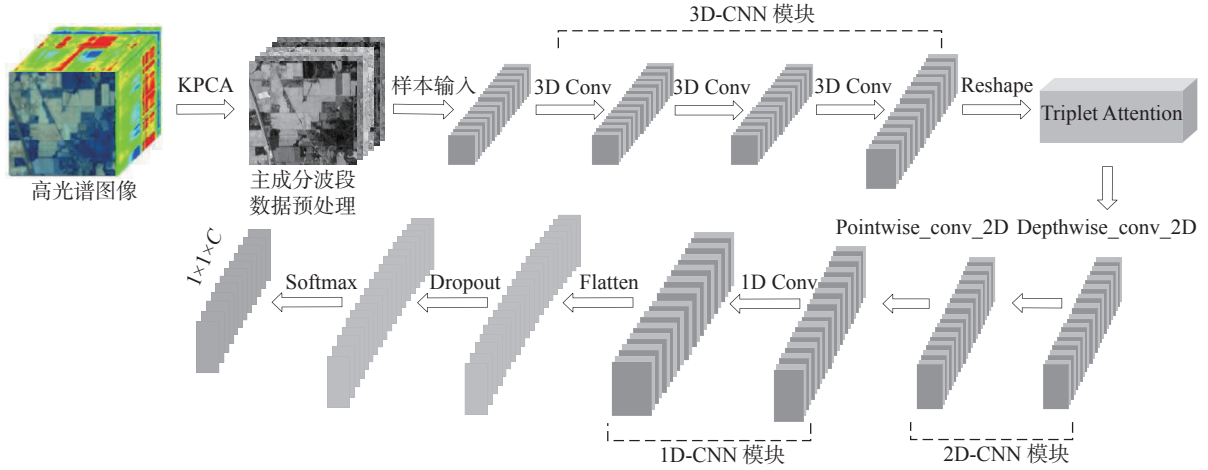


图 1 HCTA-Net 模型结构

Fig. 1 HCTA-Net structure diagram

## 1.1 混合卷积神经网络模块

### 1.1.1 3D-CNN 模块

研究表明, 3D-CNN 可同时对 HSI 数据进行光谱特征和空间特征的联合提取, 但随着网络深度的增加, 会带来细节特征逐渐丢失, 计算量大等问题。因此本文采用三层的 3D-CNN 对 HSI 数据进行特征信息提取, 图 2 给出了三维卷积的示意图。

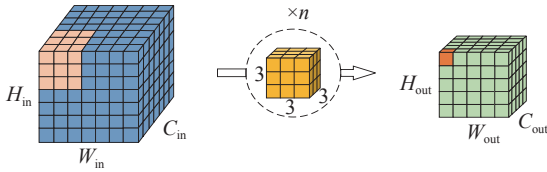


图 2 三维卷积示意

Fig. 2 3D convolution diagram

为进一步理解 3D-CNN 特征提取原理, 将位置为  $(x, y, z)$  的神经元的激活值  $v_{i,j}^{x,y,z}$  公式表示为

$$v_{i,j}^{x,y,z} = F \left( b_{i,j} + \sum_{\tau} \sum_{\lambda=0}^{T_i-1} \sum_{\rho=0}^{P_i-1} \sum_{r=0}^{R_i-1} \omega_{i,j,\tau}^{\lambda,\rho,r} v_{i-1,\tau}^{x+\lambda,y+\rho,z+r} \right) \quad (1)$$

式中: 激活函数为  $F(\cdot)$ , 表示当前第  $i$  层与第  $i-1$  层的索引; 卷积核的宽、高与光谱维度分别为  $T_i$ 、 $P_i$  和  $R_i$ ; 位置  $(\lambda, \rho, r)$  连接到第  $\tau$  个特征值的权重用  $\omega_{i,j,\tau}^{\lambda,\rho,r}$  表示; 第  $i-1$  层中第  $\tau$  个特征图位置为  $(x+\lambda, y+\rho, z+r)$  的值可表示为  $v_{i-1,\tau}^{x+\lambda,y+\rho,z+r}$ ;  $b_{i,j}$  表示偏置。

### 1.1.2 2D-CNN 模块

2D-CNN 模块沿着空间维进行操作并加强 HSI 的局部空间特征。因此, 通过上述 3D-CNN 模块获取到联合的光谱-空间特征后, 为进一步增强 HSI 的局部空间特征信息, 本文在 3D-CNN 基础上经过一层 2D-CNN 模块, 图 3 给出了二维卷积的示意图。

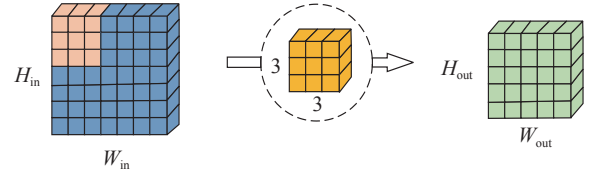


图 3 二维卷积示意

Fig. 3 2D convolution diagram

在第  $i$  层第  $j$  个特征图上, 空间位置为  $(x, y)$  的神经元的值  $v_{i,j}^{x,y}$ , 公式表示为

$$v_{i,j}^{x,y} = F \left( b_{i,j} + \sum_{\tau} \sum_{\lambda=0}^{T_i-1} \sum_{\rho=0}^{P_i-1} \omega_{i,j,\tau}^{\lambda,\rho} v_{i-1,\tau}^{x+\lambda,y+\rho} \right) \quad (2)$$

式中: 激活函数为  $F(\cdot)$ ;  $\tau$  表示当前第  $i$  层与第  $i-1$  层的索引;  $T_i$ 、 $P_i$  分别为卷积核的宽和高; 位置  $(\lambda, \rho)$  连接到第  $\tau$  个特征值的权重用  $\omega_{i,j,\tau}^{\lambda,\rho}$  表示; 第  $i-1$  层中第  $\tau$  个特征图位置为  $(x+\lambda, y+\rho)$  的值可表示为  $v_{i-1,\tau}^{x+\lambda,y+\rho}$ ;  $b_{i,j}$  为第  $i$  层第  $j$  个特征图的偏置参数。

### 1.1.3 1D-CNN 模块

为进一步增强光谱特征, 本文在 3D-CNN 与 2D-CNN 模块基础上, 再经过一层 1D-CNN 模块, 通过不同维度 CNN 的融合, 充分提取了 HSI 精细的光谱-空间联合特征, 图 4 给出了一维卷积示意图。

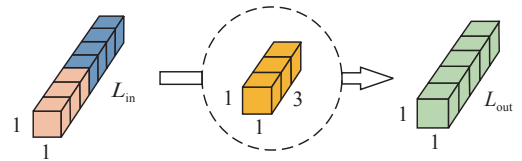


图 4 一维卷积示意

Fig. 4 1D convolution diagram

### 1.1.4 改进的 2D-CNN 模块

为了能够既充分发挥 CNN 网络的巨大优势, 同时减少可学习参数, 提高训练效率, 本文在 2D-



CNN 模块中加入深度可分离卷积,以确保模型在近似没有损失的情况下,捕获更加丰富的空间特征信息来区分不同波段的空间信息。

深度可分离卷积是传统二维卷积的一种变换形式,如图5所示,它对不同的通道采用不同的卷积核进行卷积,将传统的卷积分解为深度卷积(depthwise convolution)与逐点卷积(pointwise convolution)两部分。

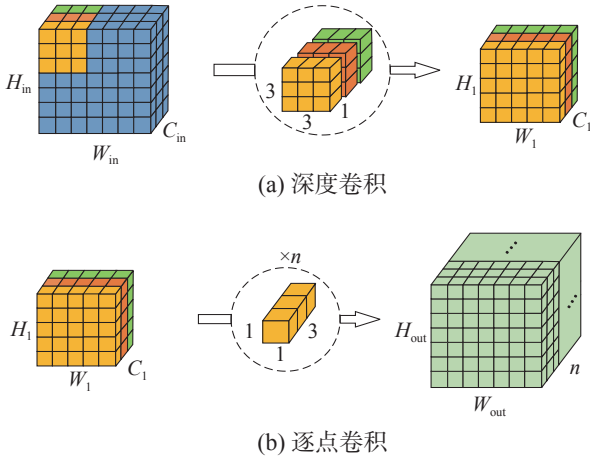


图5 深度可分离卷积

Fig. 5 Depthwise separable convolution

假设输入特征图的尺寸为 $S \times S \times M$ ,卷积核的尺寸为 $L \times L \times M$ ,其数量为 $N$ 。假设对应特征图空间位置中的每一个点都会进行一次卷积操作,

则对 $N$ 个传统的卷积,其总计算量为

$$S \times S \times L \times L \times M \times N \quad (3)$$

对深度可分离卷积而言,其计算总计算量为

$$S \times S \times L \times L \times M \times M \times N \times S \times S \quad (4)$$

那么相对于传统卷积,深度可分离卷积的计算量与传统卷积的比值 $P$ 为

$$P = \frac{1}{N} + \frac{1}{L^2} \quad (5)$$

由此可得,加入深度可分离卷积不但能够实现空间与通道相关性计算的充分解耦,同时网络模型计算效率和内存开销远低于传统卷积。

## 1.2 Triplet Attention 模块

人类视觉中的注意力会选择性地集中在部分给定的信息而忽略其余信息,这一机制有助于在保留信息上下文的同时细化感知的信息。受此启发,一些研究方法提出将注意力机制有效整合到CNN结构中,来提高视觉分类任务的性能。本文基于HSI分类问题的特点,引入轻量化的Triplet Attention注意力机制,利用三支结构捕获跨维交互来计算注意权重,通过旋转操作和残差变换建立维度间的依存关系,如图6所示,顶部分支负责计算跨通道维度 $C$ 和空间维度 $W$ 的注意力权重,中间的分支负责计算通道维度 $C$ 和空间维度 $H$ 的注意力权重,最后一个分支用于捕获 $H$ 和 $W$ 的空间维度依赖关系,最后通过平均法对权重进行聚合。

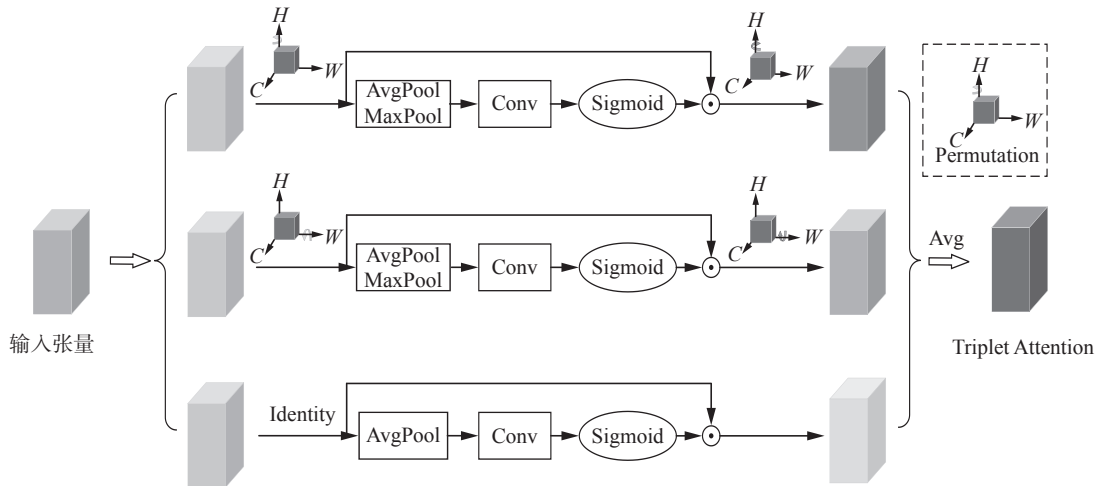


图6 Triplet Attention 原理

Fig. 6 Triplet Attention schematic

图7给出了Triplet Attention的模型结构,当给定输入为 $C \times H \times W$ 时,输入3个分支具体实现如下:

1)第1分支为通道维度 $C$ 和空间维度 $H$ 交互捕获分支,输入特征图先沿 $H$ 轴逆时针旋转 $90^\circ$ ,变成形状为 $W \times H \times C$ 的特征图,进而在 $W$ 维度上进行Z-Pool,Z-Pool是对输入进行平均池化和最

大池化,能够保留特征的丰富表示,同时缩小其深度,得到形状为 $2 \times H \times C$ 的特征,然后通过 $7 \times 7$ 的卷积层和BN层,经过Sigmoid激活层生成相应的注意权重。最后经过残差变换,沿 $H$ 轴顺时针旋转 $90^\circ$ ,将注意力权重乘回原始特征图,得到 $C \times H \times W$ 维度特征图。

2)第2分支为通道 $C$ 和空间 $W$ 维度交互捕

获分支,输入特征图先沿  $W$  轴逆时针旋转  $90^\circ$ ,得到维度为  $H \times C \times W$  的特征图,进而在  $H$  维度上进行 Z-Pool,变成形状为  $2 \times C \times W$  的特征图。最后再通过  $7 \times 7$  卷积和 BN 层,经过激活函数,再利用残差变换和旋转操作将注意力权重乘回原始特征图,变为  $C \times H \times W$  维度特征图。

3) 第 3 分支为空间注意力计算分支,输入特征图经过 Channel Pool,变成形状为  $2 \times H \times W$  的特征图,进而经过  $7 \times 7$  卷积层、BN 层,经由激活函数生成空间注意力权重。最后利用残差变换和旋转操作将注意力权重乘回原始特征图,变为  $C \times H \times W$  维度特征图。

4) 最后对 3 个分支输出的  $C \times H \times W$  维度特征进行相加取平均。

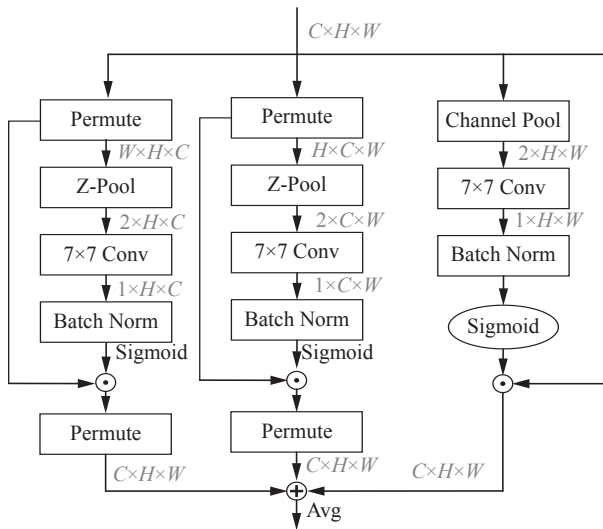


图 7 Triplet Attention 结构

Fig. 7 Triplet Attention architecture diagram

## 2 实验数据获取与预处理

### 2.1 实验数据集

为了将 HCTA-Net 模型与目前的主流方法进行对比,本文选用 Indian Pines、Salinas 和 Pavia University 3 个公开的数据集进行实验,数据集参数信息如表 1 所示。

表 1 HSI 数据集参数  
Table 1 HSI dataset parameter

参数	Indian Pines	Salinas	Pavia University
采集地点	印第安纳州	加州	意大利北部
采集设备	AVIRIS	AVIRIS	ROSIS
谱范围/ $\mu\text{m}$	0.4~2.5	0.4~2.5	0.43~0.86
大小/像素	145×145	512×217	610×340
空间分辨率/m	20	3.7	1.3

续表 1

参数	Indian Pines	Salinas	Pavia University
波段数	224	224	115
去噪波段数	200	204	103
样本量	10249	54129	42776
类别量	16	16	9

图 8~10 分别给出了 Indian Pines、Salinas 和 Pavia University 数据集相应的伪彩色图与真实参考图。

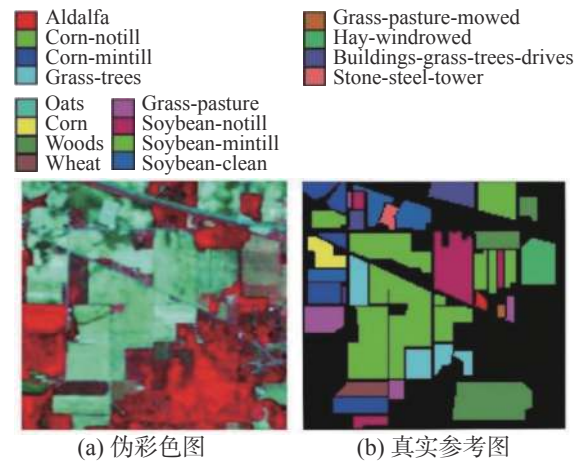


图 8 Indian Pines 数据集

Fig. 8 Indian Pines dataset

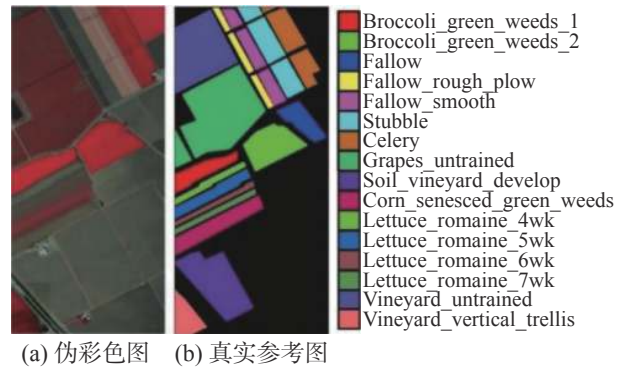


图 9 Salinas 数据集

Fig. 9 Salinas dataset

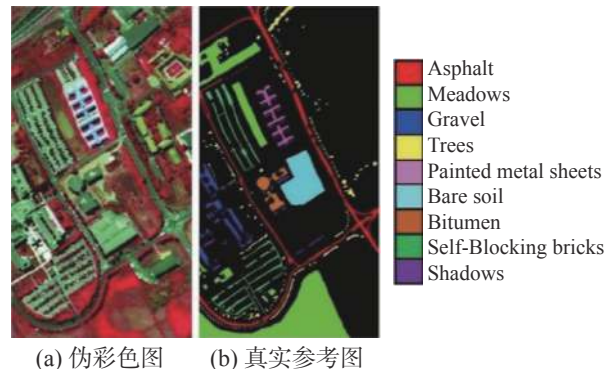


图 10 Pavia University 数据集

Fig. 10 Pavia University dataset

## 2.2 数据降维

HSI 具有高维度和高特征冗余的特点,利用 KPCA 对 HSI 进行降维,通过引用核函数,增强对 HSI 数据的非线性信号处理能力,从而使用更少的特征数据,最大限度地表征原有信息。通过求解 HSI 数据各主成分的贡献率,本文采用 Indian Pines 数据集前 30 个主成分,贡献率为 99.25%,采用 Salinas 数据集前 15 个主成分,贡献率为 99.9%,采用 Pavia University 数据集前 15 个主成分,贡献率为 99.9%。

## 2.3 实验参数设置

本文实验环境为 Intel(R) Core(TM) i7-9700K CPU @3.60 GHz,内存为 32.0 GB, GPU NVIDIA GeForce RTX 3070,软件环境采用开源的 TensorFlow 2.0 框架来搭建网络模型并进行实验。对于 Indian Pines 数据集,从每个类别中随机选取 10%、10% 和 80% 的样本组成训练集、验证集与测试集,对于 Salinas 和 Pavia University 数据集,从每个类别中随机选取 5%、5% 和 90% 的样本组成训练集、验证集与测试集。HCTA-Net 网络模型的输入窗口大小为  $25 \times 25$ ,训练迭代次数为 50 次,学习率初始化设置为 0.001,之后采用 Adam 优化器对网络进行学习训练。

在 3D-CNN 模块中,3 个 3D 卷积核大小分别为 (3,3,7)、(3,3,5) 和 (3,3,3),相对应的滤波器个数为 8、16 和 32,在改进 2D-CNN 模块中,2D 可分离卷积核大小为 (3,3),对应的滤波器数量为 64,另外,在 1D-CNN 模块中,采用 128 个卷积核大小为 3 的滤波器。实验利用 ReLu 和 Dropout 的组

合策略来缓解训练过程中可能存在的过度拟合的问题,从而有效提高了训练模型的泛化性能。最后,本文采用总体精度 (OA)、平均精度 (AA)、Kappa 系数 ( $K$ ) 3 个评价指标,来评价不同模型之间的性能差异,其中,OA 为分类正确数与总像元数之间的比值;AA 为总体精度在各个类别中的均值;Kappa 系数是用来判断一致性的程度,即分类预测结果与真实参考图之间的一致性。

## 3 5 种分类算法性能对比分析

为验证本文提出的 HCTA-Net 模型的正确性和有效性,将其与 SVM<sup>[23]</sup>、2D-CNN<sup>[24]</sup>、3D-CNN<sup>[25]</sup>、HybridSN<sup>[26]</sup>、DTAResNet<sup>[27]</sup>、CNN-ASS<sup>[28]</sup> 模型进行实验对比分析,为保证实验结果的公平性,实验均在相同环境中进行。不同方法在 Indian Pines、Salinas 和 Pavia University 数据集上的分类结果如图 11~13 所示,各类别分类精度如表 2~4 所示。以 Indian Pines 为例,由图 11 可以看出,仅使用 SVM 方法的结果图错分现象明显,3D-CNN 方法通过 3D 卷积与残差网络思想优化了这一现象,HybridSN 通过 3D 与 2D 卷积融合明显提高了分类准确度,DTAResNet 与 CNN-ASS 通过特征提取和引入注意力模块进一步地降低了错分程度。而本文提出的 HCTA-Net 方法分类效果最接近真实情况,另外,由表 2 可知,除 Oats 与 Soybean-clean 外,其他 14 类都达到了 98% 以上的分类精度,证明 HCTA-Net 模型能够聚焦融合更为丰富的光谱-空间特征信息,以达到最佳的分类效果。

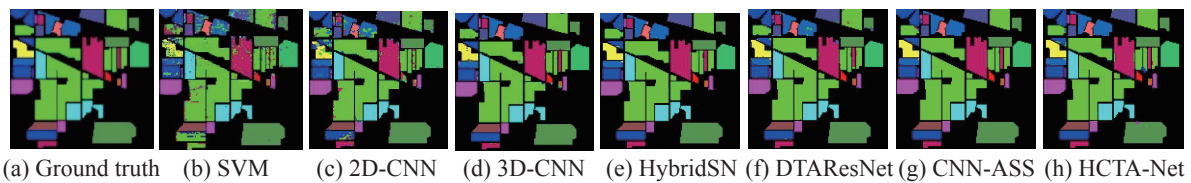


图 11 不同方法下 Indian Pines 分类结果

Fig. 11 Classification results of Indian Pines dataset with different methods

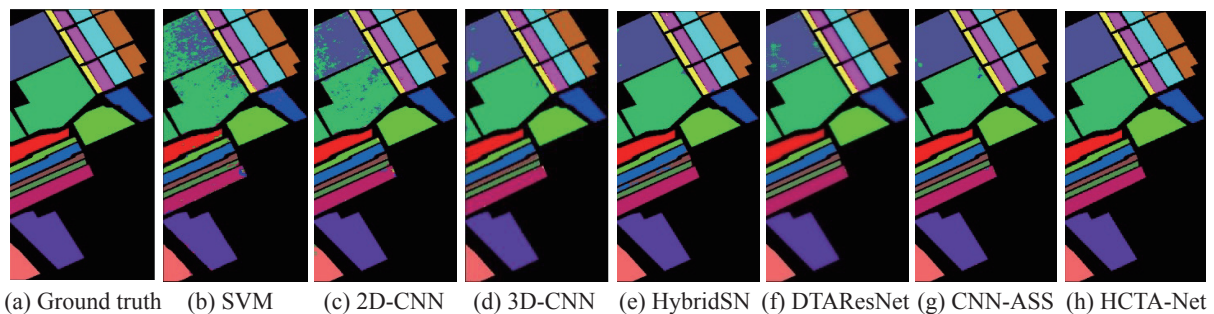


图 12 不同方法下 Salinas 数据集分类结果

Fig. 12 Classification results of Salinas dataset with different methods



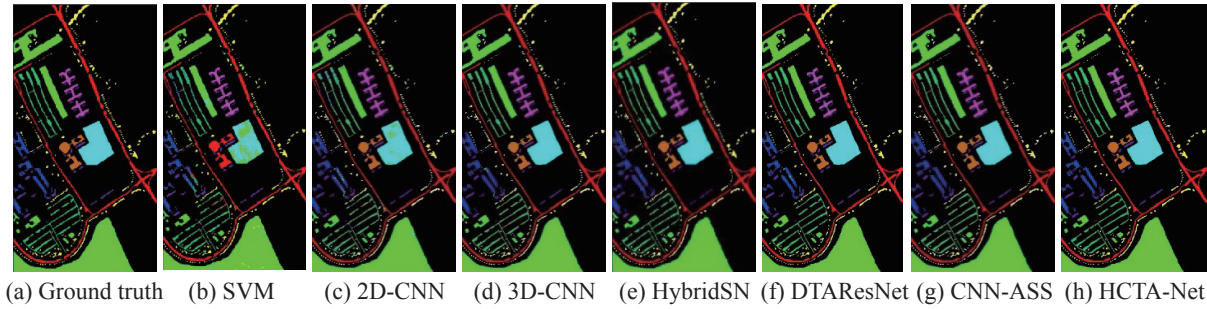


图 13 不同方法下 Pavia University 数据集分类结果图

Fig. 13 Classification results of Pavia University dataset with different methods

表 2 不同方法下 Indian Pines 数据集各类别分类结果

Table 2 Classification results of each category of Indian Pines dataset with different methods %

类别	SVM	2D-CNN	3D-CNN	HybridSN	DTAResNet	CNN-ASS	HCTA-Net
Alfalfa	59.52	100	100	100	100	100	100
Corn-notill	68.35	83.99	90.48	95.87	97.77	98.24	98.70
Corn-mintill	51.00	84.83	96.77	95.28	97.33	99.80	98.80
Corn	44.86	90.48	99.32	99.49	100	95.95	98.81
Grass-pasture	88.74	94.56	97.76	99.75	97.89	96.60	99.70
Grass-trees	93.46	97.56	98.67	99.83	99.10	96.90	99.03
Grass-pasture-mowed	84.62	100	100	100	37.04	100	100
Hay-windrowed	97.68	94.61	99.23	99.24	99.64	100	100
Oats	16.67	66.67	100	5.26	31.58	85.71	76.47
Soybean-notill	62.40	90.18	95.99	99.87	95.46	99.47	99.70
Soybean-mintill	84.71	88.58	99.69	99.01	97.54	99.52	99.13
Soybean-clean	47.19	78.29	99.38	98.11	98.31	98.02	97.58
Wheat	93.51	93.26	100	96.55	100	96.85k	100
Woods	96.66	98.40	98.05	99.71	99.87	98.44	100
Buildings-grass-trees-drives	56.03	89.80	100	98.13	99.55	99.57	100
Stone-steel-towers	90.48	94.67	98.31	94.52	90.63	96.43	98.39
OA	76.29	89.87	97.28	98.36	97.55	98.66	99.16
AA	70.99	78.19	85.85	91.96	90.11	96.64	98.27
K	72.62	88.42	96.89	98.13	97.21	98.44	99.05

表 3 不同方法下 Salinas 数据集各类别分类结果

Table 3 Classification results of each category of Salinas dataset with different methods %

类别	SVM	2D-CNN	3D-CNN	HybridSN	DTAResNet	CNN-ASS	HCTA-Net
Brocoli_green_weeds_1	97.64	100	100	100	100	100	99.67
Brocoli_green_weeds_2	99.92	99.83	99.91	100	100	99.86	99.58
Fallow	96.54	98.58	100	100	99.39	99.63	100
Fallow_rough_plow	98.94	97.91	99.92	98.54	97.89	99.03	99.21
Fallow_smooth	98.00	97.93	99.83	96.95	99.62	99.33	100
Stubble	99.76	100	100	100	100	100	100
Celery	99.21	99.85	100	100	100	100	100
Grapes_untrained	91.44	89.64	99.41	97.91	96.18	98.89	99.90
Soil_vinyard_develop	99.30	99.63	99.98	100	100	100	100
Corn_senesced_green_weeds	93.39	97.69	99.83	100	99.59	99.77	99.90



续表 3

类别	SVM	2D-CNN	3D-CNN	HybridSN	DTAResNet	CNN-ASS	HCTA-Net
Lettuce_roumaine_4wk	94.29	98.81	96.81	99.81	99.28	100	100
Lettuce_roumaine_5wk	99.73	99.24	99.66	99.25	99.54	99.84	100
Lettuce_roumaine_6wk	98.16	99.42	100	98.23	100	99.64	99.76
Lettuce_roumaine_7wk	92.23	96.41	99.89	100	100	99.32	99.90
Vinyard_untrained	52.35	88.86	95.33	99.93	99.95	99.20	99.79
Vinyard_vertical_trellis	98.95	98.88	100	100	100	100	99.81
OA	90.58	95.78	99.12	99.29	99.01	99.52	99.87
AA	94.37	97.64	99.41	99.53	99.41	99.64	99.84
K	89.48	95.30	99.02	99.21	98.90	99.47	99.86

表 4 不同方法下 Pavia University 数据集各类别分类结果

Table 4 Classification results of each category of Pavia University dataset with different methods

%

类别	SVM	2D-CNN	3D-CNN	HybridSN	DTAResNet	CNN-ASS	HCTA-Net
Asphalt	94.98	96.32	99.97	99.25	99.29	98.53	99.49
Meadows	97.44	96.73	99.45	99.74	99.66	99.82	99.98
Gravel	73.08	84.42	99.75	99.29	92.55	98.47	98.82
Trees	84.44	97.21	99.85	98.01	99.38	99.13	99.63
Painted metal sheets	99.69	100	100	100	100	99.92	99.91
Bare soil	70.24	96.12	99.17	99.71	99.48	99.98	100
Bitumen	1.42	96.29	99.92	98.44	97.84	99.25	100
Self-Blocking bricks	88.14	87.13	90.76	97.36	91.58	98.35	99.29
Shadows	100	97.75	98.84	99.50	99.87	97.84	100
OA	88.08	95.30	98.73	99.27	98.46	99.34	99.76
AA	78.83	93.33	98.63	98.38	97.74	98.29	99.68
K	83.88	93.75	98.32	99.04	97.96	99.13	99.69

表 2~4 给出了不同方法分类结果的评价指标值, 由表可知, HCTA-Net 方法的 OA、AA、Kappa 系数均达到了最优。在 Indian Pines 数据集上, OA 分别为 99.16%, 与 SVM、2D-CNN、3D-CNN、HybridSN、DTAResNet、CNN-ASS 方法相比, 分别提高了 22.87%、9.29%、1.88%、0.8%、1.81% 与 0.5%, 与其他方法中性能最好的 CNN-ASS 方法相比, AA 高出了 1.63%, Kappa 系数高出了 0.61%。在 Salinas 数据集上, OA 达到了 99.87%, 比 SVM、2D-CNN、3D-CNN、HybridSN、DTAResNet、CNN-ASS 方法分别提高了 9.29%、4.09%、0.75%、0.58%、0.86%、0.35%, 与分类效果较为理想的 CNN-ASS 方法相比, AA 提高了 0.2%, Kappa 系数提高

了 0.39%, 这是因为 Salinas 数据集本身样本分布较为均衡, 光谱与空间维度信息有较大差异, 所以有利于分类。在 Pavia University 数据集上, OA 达到了 99.76%, 比 SVM、2D-CNN、3D-CNN、HybridSN、DTAResNet、CNN-ASS 方法分别提高了 11.68%、4.46%、1.03%、0.49%、1.3%、0.42%, 与分类效果较为理想的 CNN-ASS 方法相比, AA 提高了 1.39%, Kappa 系数提高了 0.59%, 此外, 由于本文仅使用了三层 3D、一层 2D 和一层 1D 卷积层, 并引入深度可分离卷积, 使模型更为轻量化, 训练时间也得以进一步减少, 实验结果如表 5 所示, 由表可得, 本文提出的 HCTA-Net 方法网络参数最少, 所用训练时间最短, 为 42.5 s。

表 5 不同分类方法网络参数与训练时间

Table 5 Network parameters and training time of different classification methods

参数	3D-CNN	HybridSN	DTAResNet	CNN-ASS	HCTA-Net
网络参数	199 153	4 844 793	21 975 725	4 787 507	573 989
训练时间/s	124.2	76.1	441.4	63.7	42.5

在实验中,一般通过训练和验证损失率来分析判断网络模型的稳定性与拟合度。图 14 给出了 HCTA-Net 方法应用在 Indian Pines 数据集上的训练和验证损失率的变化曲线。由图可以看出,训练和验证的损失率在训练过程中均基本处于不断下降的状态,在迭代至 20 轮左右达到稳定状态,训练和验证损失基本接近 0,说明 HCTA-Net 模型通过训练学习到了数据集的有效特征,而且模型相对稳定,收敛速度快,具有良好的学习能力和一定的泛化能力,最终总体分类精度可达到 99.16%,意味着可以较为准确地地区分出 HSI 数据集中绝大部分的地物类别。

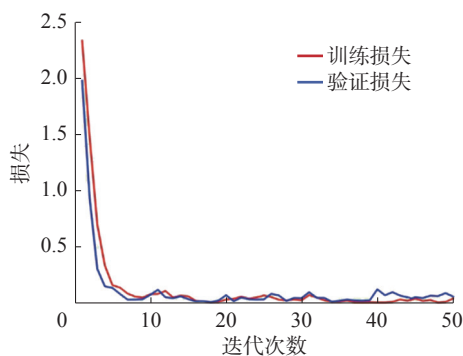


图 14 Indian Pines 数据集训练和验证损失变化曲线

Fig. 14 Training and validation loss rate of Indian Pines dataset

## 4 结束语

本文针对 HSI 数据维度高、特征信息表征不到位及参数过多等问题,提出一种基于混合卷积与三重注意力的卷积神经网络 (HCTA-Net)。该网络模型主要设计了一种由 3D 卷积层、2D 卷积层和 1D 卷积层构成的混合卷积神经网络模块,有效融合不同维度的卷积层,充分利用光谱-空间域与深空间域互补特性,综合提取不同层次更加丰富的光谱-空间特征。为进一步捕获跨空间维度和通道维度间的交互信息,加入了三重注意力模块,同时通过引入深度可分离卷积,有效减少了网络参数,缩短了训练时间。该方法在 Indian Pines、Salinas 和 Pavia University 数据集上 OA 分别达到了 99.16%、99.87% 和 99.76%,证明了本文方法可以有效提高 HSI 的分类性能。虽然本文模型具有较好的分类效果,但也存在一些不足,在后续的研究中,将进一步改进网络结构,使用更轻量级的模型达到更好的分类精度。

## 参考文献:

[1] 王锦锦, 李真, 朱玉玲. 高光谱影像在海洋环境监测中

的应用[J]. *卫星应用*, 2019(8): 36–40.

- [2] 高暖. 高光谱图像地物及目标识别[D]. 西安: 西安电子科技大学, 2019: 1–5.  
GAO Nuan. Terrain and target detection in hyperspectral remote sensing imagery[D]. Xi'an: Xidian University, 2019: 1–5.
- [3] HUGHES G. On the mean accuracy of statistical pattern recognizers[J]. *IEEE transactions on information theory*, 1968, 14(1): 55–63.
- [4] 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): 1804–1811.
- [5] TARABALKA Y, FAUVEL M, CHANUSSOT J, et al. SVM- and MRF-based method for accurate classification of hyperspectral images[J]. *IEEE geoscience and remote sensing letters*, 2010, 7(4): 736–740.
- [6] 陈善学, 何宇峰. 基于空谱字典的加权联合稀疏表示高光谱图像分类[J]. *光学学报*, 2023, 43(1): 0110002–1–11.  
CHEN Shanxue, HE Yufeng. Weighted joint sparse representation hyperspectral image classification based on spatial-spectral dictionary[J]. *Acta optica sinica*, 2023, 43(1): 0110002–1–11.
- [7] 熊余, 单德明, 姚玉, 等. 多特征融合下的高光谱图像混合卷积分类[J]. *红外技术*, 2022, 44(1): 9–20.  
XIONG Yu, SHAN Deming, YAO Yu, et al. Hyperspectral image hybrid convolution classification under multi-feature fusion[J]. *Infrared technology*, 2022, 44(1): 9–20.
- [8] FAUVEL M, TARABALKA Y, BENEDIKTSSON J A, et al. Advances in spectral-spatial classification of hyperspectral images[J]. *Proceedings of the IEEE*, 2013, 101(3): 652–675.
- [9] ZHONG Ping, GONG Zhiqiang. A hybrid DBN and CRF model for spectral-spatial classification of hyperspectral images[J]. *Statistics, optimization & information computing*, 2017, 5(2): 75–98.
- [10] LIU Jinxiang, ZHANG Kefei, WU Suqin, et al. An investigation of a multidimensional CNN combined with an attention mechanism model to resolve small-sample problems in hyperspectral image classification[J]. *Remote sensing*, 2022, 14(3): 785.
- [11] HU Wei, HUANG Yangyu, WEI Li, et al. Deep convolutional neural networks for hyperspectral image classification[J]. *Journal of sensors*, 2015, 2015: 258619.
- [12] XU Xiaodong, LI Wei, RAN Qiong, 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.
- [13] ZHANG Xin, WANG Yongcheng, ZHANG Ning, et al. SSDANet: spectral-spatial three-dimensional convolu-

- tional neural network for hyperspectral image classification[J]. IEEE access, 2020, 8: 127167–127180.
- [14] GUO Hao, LIU Jianjun, YANG Jinlong, et al. Deep collaborative attention network for hyperspectral image classification by combining 2-D CNN and 3-D CNN[J]. IEEE journal of selected topics in applied earth observations and remote sensing, 2020, 13: 4789–4802.
- [15] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132–7141.
- [16] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[M]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3–19.
- [17] FANG Bei, LI Ying, ZHANG Haokui, et al. Hyperspectral images classification based on dense convolutional networks with spectral-wise attention mechanism[J]. Remote sensing, 2019, 11(2): 159.
- [18] ROY S K, MANNA S, SONG Tiecheng, et al. Attention-based adaptive spectral-spatial kernel ResNet for hyperspectral image classification[J]. IEEE transactions on geoscience and remote sensing, 2021, 59(9): 7831–7843.
- [19] AHMAD M, KHAN A M, MAZZARA M, et al. Attention mechanism meets with hybrid dense network for hyperspectral image classification[EB/OL].(2022–01–04) [2022–12–12].<https://arxiv.org/abs/2201.01001>.
- [20] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 1800–1807.
- [21] CUI Ying, YU Zikun, HAN Jiacheng, et al. Dual-triple attention network for hyperspectral image classification using limited training samples[J]. IEEE geoscience and remote sensing letters, 2022, 19: 1–5.
- [22] 李昌元, 刘国栋, 谭博. 基于 PCA 和 KPCA 的高光谱遥感数据降维对比研究 [J]. 地理空间信息, 2022, 20(7): 89–93, 103.  
LI Changyuan, LIU Guodong, TAN Bo. Comparative study on hyperspectral remote sensing data dimensionality reduction based on PCA and KPCA[J]. Geospatial information, 2022, 20(7): 89–93, 103.
- [23] 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.
- [24] CHEN Yushi, JIANG Hanlu, LI Chunyang, 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.
- [25] ZHONG Zilong, LI J, LUO Zhiming, 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.
- [26] ROY S K, KRISHNA G, DUBEY S R, et al. HybridSN: exploring 3-D–2-D CNN feature hierarchy for hyperspectral image classification[J]. IEEE geoscience and remote sensing letters, 2020, 17(2): 277–281.
- [27] 王爱丽, 刘美红, 薛冬, 等. 结合动态卷积和三重注意力机制的高光谱图像分类 [J]. 激光与光电子学进展, 2022, 59(10): 341–351.  
WANG Aili, LIU Meihong, XUE Dong, et al. Hyperspectral image classification combined dynamic convolution with triplet attention mechanism[J]. Laser & optoelectronics progress, 2022, 59(10): 341–351.
- [28] ZHANG Yangming, YANG Kun, YUAN Lei. 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.

#### 作者简介:



王瑞婷, 硕士研究生, 主要研究方向为复杂环境智能信息感知。



王海燕, 教授, 博士生导师, 主要研究方向为智能感知技术。负责国家自然科学基金、国家科技重大专项、国家 863 计划等项目 20 余项, 发表学术论文 80 余篇。



陈晓, 讲师, 主要研究方向为多目标跟踪理论和 AI 技术。