

基于小波卷积网络的高光谱图像分类^①

巩传江¹, 臧德厚³, 郭金¹, 孙媛媛^{1,2}, 宋廷强^{1,2}

¹(青岛科技大学 信息科学技术学院, 青岛 266061)

²(青岛科技大学 大数据学院, 青岛 266061)

³(颐中烟草(集团)有限公司, 青岛 266021)

通信作者: 宋廷强, E-mail: 116638307@qq.com



摘要: 高光谱图像波段多、波段之间关联性强,但其空间纹理和几何信息的表达较弱,传统分类模型存在空间光谱特征提取不充分、计算量大的问题,分类性能有待提高.针对此问题,提出一种基于小波变换的多尺度多分辨率注意力特征融合卷积网络(wavelet transform convolutional attention network, WTCAN),采用小波变换思想对光谱波段进行4次分解,通过层次性提取光谱特征可减少计算量.该网络设计了空间信息提取模块,同时引入金字塔注意力机制,通过设计逆向跳跃连接网络结构利用多尺度获取空间位置特征,增强空间纹理表达能力,可以有效改进传统2D-CNN特征提取尺度单一、忽略空间纹理细节等缺陷.本文对所提出的WTCAN模型分别在不同空间分辨率高光谱数据集Indian Pines (IP)、WHU_Hi_HanChuan (HanChuan)、WHU_Hi_HongHu (HongHu)进行实验,通过对比SVM、2D-CNN、DBMA、DBDA、HybridSN模型效果,WTCAN模型取得较好的分类效果,3个数据集的分类总体精度分别达到了98.41%、99.64%、99.67%,可为高光谱图像的分类研究提供参考依据.

关键词: 高光谱图像分类; 特征提取; 小波变换; 二维卷积神经网络; 注意力机制

引用格式: 巩传江,臧德厚,郭金,孙媛媛,宋廷强.基于小波卷积网络的高光谱图像分类.计算机系统应用,2023,32(7):23-34. <http://www.c-s-a.org.cn/1003-3254/9186.html>

Hyperspectral Image Classification Based on Wavelet Convolution Network

GONG Chuan-Jiang¹, ZANG De-Hou³, GUO Jin¹, SUN Yuan-Yuan^{1,2}, SONG Ting-Qiang^{1,2}

¹(School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China)

²(School of Big Data, Qingdao University of Science and Technology, Qingdao 266061, China)

³(Yizhong Tobacco (Group) Co. Ltd., Qingdao 266021, China)

Abstract: Hyperspectral images have multiple bands and a strong correlation between bands, but their spatial texture and geometric information are poorly expressed. The traditional classification model has insufficient extraction of spatial spectral features and large calculation, and its classification performance needs to be improved. To solve this problem, a multi-scale and multi-resolution attention feature fusion convolution network (WTCAN) based on the wavelet transform is proposed. The concept of wavelet transform is applied to decompose the spectral band four times, and the hierarchical extraction of spectral features can reduce the calculation amount. The network has designed the spatial information extraction module and introduced the pyramid attention mechanism. By designing the reverse jump connection network structure, it uses multiple scales to obtain the spatial position features and enhances the expression ability of spatial texture, which can effectively improve the defects of traditional 2D-CNN feature extraction, such as single scale and the ignoring of spatial texture details. The proposed WTCAN model is experimented on the hyperspectral datasets with different spatial resolutions—Indian Pines (IP), WHU_Hi_HanChuan (HanChuan), and WHU_Hi_HongHu (HongHu)

① 基金项目: 山东省自然科学基金青年项目 (ZR2021QC120); 山东省科技研发项目 (2019GGX101047)

收稿时间: 2023-01-05; 修改时间: 2023-02-09; 采用时间: 2023-03-03; csa 在线出版时间: 2023-05-24

CNKI 网络首发时间: 2023-05-25

respectively. By comparing the effects of SVM, 2D-CNN, DBMA, DBDA, and HybridSN models, the WTCAN model achieves excellent classification results. The overall classification precision of the three datasets reaches 98.41%, 99.64%, and 99.67% respectively, which can provide a valuable reference for the research on the classification of hyperspectral images.

Key words: hyperspectral imagery (HSI) classification; feature extraction; wavelet transform; two-dimensional convolutional neural network (2D-CNN); attention mechanism

高光谱图像 (hyperspectral imagery, HSI) 的光谱分辨率可达纳米级, 分辨率越高, 地物信息越详细, 越有利于地物区分. 与多光谱图像相比, 高光谱图像在光谱维度上包含更多的波段数量, 富含更多地物特质, 但也存在大量噪声以及数据冗余的问题. 高光谱图像不仅光谱数量多, 空间信息也十分丰富, 因此, 对于地物分类而言, 通过提取高光谱图像的空间特征和光谱特征, 可有效实现地面物体的识别及区分. 目前, 高光谱图像在城市遥感^[1]、目标检测^[2]、海洋应用^[3]以及变化检测^[4]等领域都有很好的应用, 高光谱图像分类已成为一个研究热点.

高光谱图像分类为高光谱应用研究奠定了基础, 如何对高光谱图像中的每个像素进行快速准确的判别是主要研究的问题. 不同地物具有不同的光谱曲线, 最初期一些基于机器学习的利用光谱特征和数据统计特征传统的方法被应用于分类任务中, 包括最近改进的 K 近邻算法 (K nearest neighbors, KNN)^[5]、支持向量机 (support vector machines, SVM)^[6]、稀疏表示分类 (sparse representation, SR)^[7]、随机森林 (random forest, RF)^[8]等. 但是上述方法都存在一个共同的问题, 即利用光谱信息对像素进行分类, 没有考虑光谱空间相关性. 这样出现的一个问题就是属于某一类像素的光谱信息很可能与其他类别像素的光谱信息混合在一起, 另外一个问题就是在较小的样本也很难学习出一个高精度的分类器, 也就是著名修斯现象^[9], 仅浅层特征提取会导致分类精度差强人意. 基于深度学习的方法克服了提取浅层特征的局限性, 在提取深层和非线性特征方面具有优势, 因此被逐渐应用在高光谱图像分类任务中, 如深度信念网络 (deep belief network, DBN)^[10]和栈式自编码器 (stacked autoencoder, SAE)^[11]以及卷积神经网络 (convolutional neural network, CNN). DBN 和 SAE 这两种方法的局限性在于网络模型的输入要求是一维数据, 使得原本二维空间向量被迫变换成一维向量, 从而造成空间信息的丢失. 卷积神经网络广泛应用于图

像处理和模型识别等方面, CNN 和前两者相比, 能够在不破坏空间结构的前提下提取空间特征, 因此很多 CNN 的方法更广泛引入到高光谱图像分类. 郭国栋^[12]针对小样本高光谱数据样本少的问题采用像素对匹配的方式增加样本, 最后采用一维卷积神经网络 (one-dimensional convolutional neural network, 1D-CNN) 对高光谱图像进行分类, 但它的局限性在于 1D-CNN 只能提取光谱维度上的特征, 忽略了空间信息. Xu 等人^[13]提出使用 RpNet (random patches network) 模型提取高光谱图像特征, 首先使用一些降维算法对图像进行预处理, 然后将图像中的随机块直接作为卷积核, 无需训练, 使其拥有了多尺度、耗时少优点. 比较常用的降维算法为主成分分析 (principal component analysis, PCA), 但是 PCA 降维的方法是通过线性变换来提取很多相似样本的相似信息, 而高光谱数据本质上是非线性的, 所以 PCA 会导致丢失很多有用信息. 本文选用因子分析 (factor analysis, FA) 的降维方法, 原因在于它能够反映出不同相关光谱和重叠光谱的差异性, 有助于模型更好地对相似样本进行分类, 且它是非线性变换. 针对同一类别像素的光谱特征方差较大时分类效果非常不理想这一问题, Li 等人^[14]提出一种基于自动聚类的双通道 CNN, 首先在光谱维通过自动聚类的方法降低类间方差, 然后用 CNN 提取光谱维信息. 实验结果证明该方法针对光谱信息类内方差较大的物质有很好的区分效果. 传统单一的 2D-CNN 更多地关注了图像空间信息, 但在一定程度上造成光谱特定信息丢失. 三维卷积神经网络 (three-dimensional convolutional neural network, 3D-CNN)^[15,16]可以分别从图像空间和光谱维度上提取到特征, 相比单一 1D-CNN 和 2D-CNN 模型而言它可以明显提高分类精度. Zhong 等人^[17]基于 3D-CNN 提出光谱-空间残差网络 (SSRN), 促进了梯度的反向传播, 同时提取到了更深层次的光谱特征, 缓解了深度学习模型精度下降的问题. 混合像元这一问题一直存在于高光谱图像分类任务, Gao 等人^[18]对

3D-CNN 进行改进并提出了一种基于谱特征增强的“三明治”卷积模型 (sandwich CNN based on spectral feature enhancement), 可以提取更加抽象的谱特征以抑制混合像元的干扰, 特别是在面对低分辨率遥感图像时效果提升更加明显. 针对单一网络模型对小样本高光谱数据特征提取有限、模型计算量大这一问题, Roy 等人^[19]提出了 2D-CNN 和 3D-CNN 的混合卷积神经网络 (HybridSN), 3D-CNN 有助于获取高光谱图像空间光谱联合特征, 2D-CNN 进一步学习更加抽象的空间特征表达, 降低了模型的复杂度, 在小样本数据上具有比单一 3D-CNN 更好的分类性能.

近年来, 注意力机制^[20,21]被引入到卷积神经网络中, 使得模型能够有选择地识别目标特征, 忽略一些对分类无用的信息. Li 等人^[22]提出一种双支双注意力机制网络 (double-branch dual-attention mechanism network, DBDA), 双支分别提取光谱和空间特征来减少之间的干扰, 并在各自分支添加注意力机制确保更加有效利用光谱空间特征, 最后通过特征融合完成分类. Wu 等人^[23]进一步提出了双支空谱注意力机制网络, 设计了带有注意力模块的双支并行 3D-CNN 模型, 一支为专注于提取光谱特征并自适应于学习较强的光谱通道, 另一支为专注于探索空间特征并自适应学习较强的空间结构, 最后将双支 3D-CNN 提取到的特征进行融合, 通过分类器完成图像分类. Sun 等人^[24]提出了一种光谱空间注意力网络 (SSAN), 引入注意力模块抑制干扰像素的影响, 提取出对分类有较大贡献的光谱空间特征. 尽管上述深度学习方法取得了不错的分类性能, 但是基于 3D-CNN 的模型参数需要大量训练成本, 计算成本也大, 应用不够广泛, 另一方面多数模型卷积核比较单一, Qing 等人^[25]提出了多尺度残差注意力机制网络结构, 该神经网络嵌入了 Wang 等人^[26]提出的轻量级改进的通道注意机制, 即高效通道注意力网络 (ECA-NET), 使用局部跨通道交互而不降维, 有效地提取了图像的空间特征和光谱特征, 减少了训练样本信息的冗余. 利用不同尺度和大小的卷积核和新的残差结构, 多次提取图像的空间谱特征, 在小样本数据集上表现出不错的泛化能力, 但特征融合能力仍然可以进一步提高, 在小样本及样本数量较多的复杂区域分类精度仍有提升空间.

针对上述问题, 本文提出了一种用于高光谱图像分类任务的基于小波变换的多尺度多分辨率注意力卷

积模型 (WTCAN). 为了减少原始数据的冗余度、训练时间, 使用因子分析 (FA) 的方法对光谱波段进行降维. 为了降低提取重要光谱特征网络的计算量和复杂度, 采取 Harr 小波变换层次变换思想获取有效波段信息. 为了充分提取空间纹理结构和细节信息, 本文设计了逆向跳跃网络结构的空信息提取模块, 利用设计的网络结构学习高光谱图像的空间位置特征, 创建得到空间光谱向量, 引入金字塔注意力机制, 采用多尺度分组卷积来提取光谱-空间特征, 通过多尺度卷积既能提取全局信息, 又能关注到局部信息, 可以增强高光谱图像特征的表达性, 提高网络性能.

1 背景知识

1.1 多分辨率小波变换

小波卷积神经网络是小波变换与卷积神经网络的结合. 为了能够给图像去噪, 充分获取高光谱图像的纹理和细节信息, 采用多分辨率 Harr 小波分解对光谱信息提取. 二维图像 Harr 变换通过低通和高通滤波器从水平和竖直两个方向上进行分解, 图像的一次 Harr 分解过程如图 1 所示: 光谱降维之后的图像实质上是一个矩阵, 首先沿着矩阵的每一行做一维 Harr 变换, 其次沿着矩阵每一列做一维 Harr 变换, 得到 4 个子图, 分别表示数字图像的近似信息、水平细节信息、垂直细节信息、对角细节信息, 重复上述步骤完成层次划分. 本文采用 4 次 Harr 小波层次变换提取光谱特征.



图 1 小波分解

二维 Harr 小波利用以下 4 个核 $f_{L,L}$, $f_{L,H}$, $f_{H,L}$, $f_{H,H}$ 进行小波变换.

$$f_{L,L} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, f_{L,H} = \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix},$$

$$f_{H,L} = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}, f_{H,H} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

假设在 Harr 变换之前高光谱图像块 x 大小为 $S \times S$, (i, j) 表示光谱位置, Harr 变换公式可以记作:

$$Harr(i, j) = x(2i-1, 2j-1) + x(2i-1, 2j) + x(2i, 2j-1) + x(2i, 2j) \quad (1)$$

1.2 EPSABlock

Zhang 等人^[27]引入基于 SENet 模块^[28]的金字塔注意力模块,自适应对通道注意力权重进行调整.为了从不同尺度上建立更长距离的光谱通道之间关系,丰富多尺度特征表达,在卷积层之后添加金字塔注意力模块.

EPSABlock 如图 2 所示,输入特征 input 依次经过 1×1 卷积、PSA 模块、 1×1 卷积,并将最后结果矩阵与原输入特征 input 矩阵相加得到最终特征图.

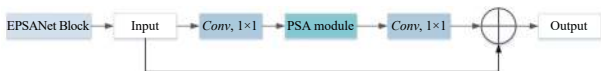


图 2 EPSABlock 模块

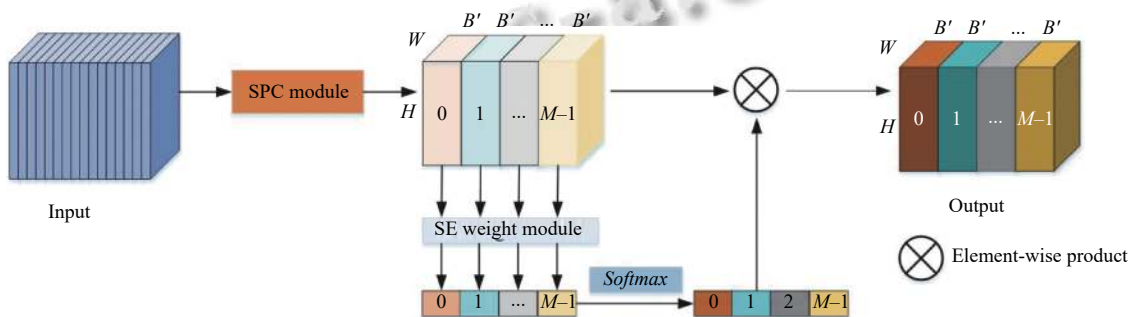


图 3 PSA 模块

将 M 个特征图沿着波谱通道进行拼接,输出一个信息更为丰富的空谱特征向量.

PSA 最重要组成部分即为 SPC 模块,如图 4 所示.对于特征图 x 按照波谱数量将其切割成 M 个部分.如果 x 的波谱数量为 B ,每一部分的波谱数量 $C = B/M$,每一部分特征图满足 $x_i \in R^{h,w,C}$, $i = 0, 1, 2, \dots, M-1$,每一

金字塔注意力机制模块 (pyramid split attention, PSA) 模块如图 3 所示, input 经过 SPC (split and concat) 模块后得到多尺度融合特征图, SPC 过程会在下一段落详细阐述.之后经过 SE 模块通过波谱通道注意力权重计算得到一个输出.

$$Z_i = SE(F_i), i = 0, 1, 2, \dots, M-1 \quad (2)$$

其中, F_i 即表示多尺度特征融合图, M 表示特征图数量.进一步使用 Softmax 进行权重重标定.

$$Softmax(Z_i) = \frac{\exp(Z_i)}{\sum_{i=0}^{M-1} \exp(Z_i)} \quad (3)$$

部分使用多尺度分组卷积提取不同空间信息,其中,分组卷积可以降低网络参数量卷积计算输出.

$$F_i = Conv(K_i, G_i)(x_i), i = 0, 1, \dots, M-1 \quad (4)$$

其中, K_i 代表卷积核大小, G_i 代表 group 大小, $G_i = 2^{\frac{K_i-1}{2}}$, 最终结果特征图按照维度拼接,即:

$$F = concat(F_0, F_1, \dots, F_{M-1}) \quad (5)$$

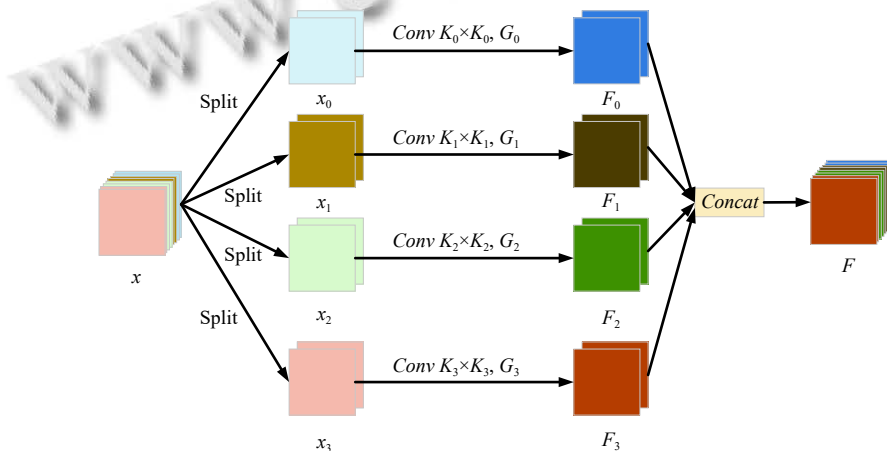


图 4 SPC 模块

2 本文方法

2.1 整体架构

WTCAN 模型整体架构设计如图 5 所示, 该模型基于小波卷积神经网络对高光谱图像提取空谱特征. 假设高光谱数据的输入维度为 $M \times N \times R$, M 、 N 、 R 分别表示高光谱数据的空间大小和波段数量. 数据预处理过程使用因子分析的降维方法, 降维后数据维度变为 $M \times N \times B$, B 表示降维后的波段数量. 图 5 网络模型以 Indian Pines 数据为例, 选择大小为 $S \times S \times B$ 的 HSI 图像块作为网络模型输入, 对图像块作 4 次小波变换处理, 通过光谱波段叠加串联, 可以得到 4 个不同空间尺度大小高光谱数据块, 光谱降维之后维度为 3, 然后将其输入到空间信息提取模块, 空间信息提取模块中

得到的特征图在通道上进行级联, 逆向跳跃连接地网络结构也达到了空间信息融合的效果, 同时在卷积层中添加金字塔注意力模块从多尺度提取纹理和空间特征, 通过金字塔注意力模块的 SPC 模块从多尺度集成输入特征图信息, 能够有效地从每一个通道特征图提取不同尺度空间信息, 更精确地合并相邻尺度的上下文特征, 然后利用通道注意力权重构建跨维度交互, 最后通过 *Softmax* 重新校准相应通道的注意力权重, 进而获取到空间光谱融合向量, 最终经过全局平均池化, 将其输入到全连接层, 实现像素分类. 卷积核大小设置为 3×3 , padding 设置为 1, 池化选用最大池化, 同时使用 dropout 层避免过拟合, 最后通过 *Softmax* 得到最终分类结果.

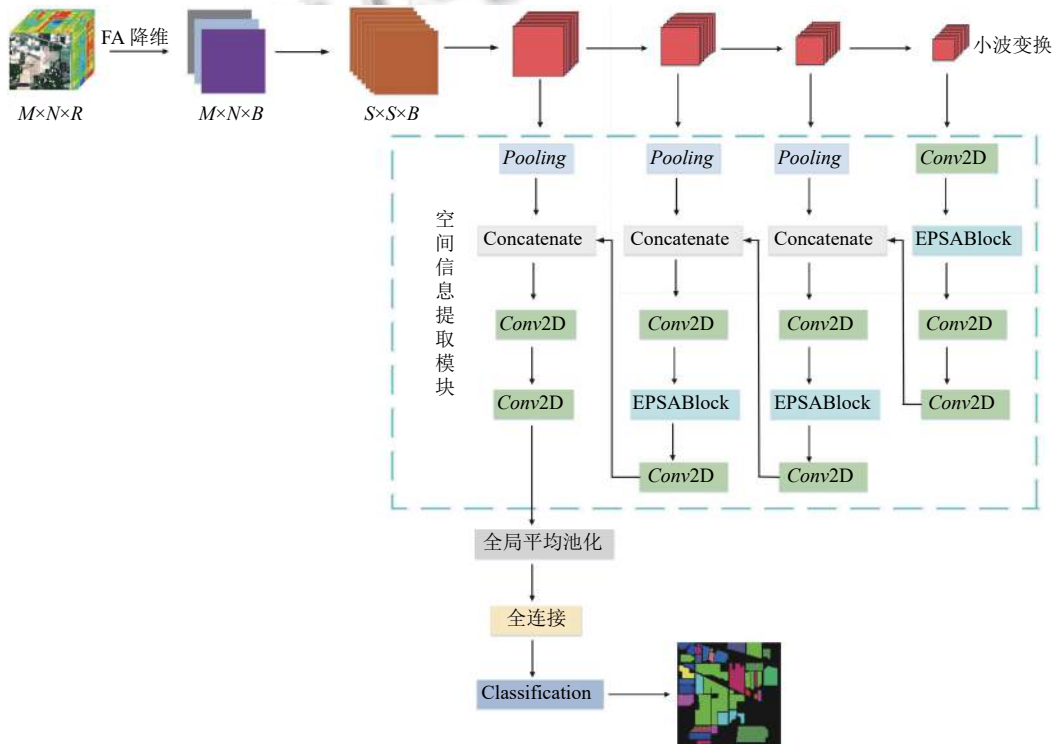


图 5 WTCAN 网络结构图

2.2 空间信息提取模块

本文设计的空间信息提取模块结构如图 6 所示, 假设输入特征 $x_0 \in R^{h_0, w_0, c}$, 其中, $h_0 = w_0$, $c = 3$. h , w 表示空间的高度和宽度, c 代表降维后波段数量. 小波变换 x_i 的输出大小为:

$$x_i = \begin{cases} x_i \in R^{h_i, w_i, 4c}, & i = 0 \\ x_i \in R^{\frac{h_{i-1}}{2}, \frac{w_{i-1}}{2}, 4c}, & i = 1, 2, 3, 4 \end{cases} \quad (6)$$

经过小波变换之后得到 x_1 、 x_2 、 x_3 、 x_4 这 4 个分辨率不同的特征图, 对 x_1 、 x_2 、 x_3 作最大池化处理, 既能保留更多图像纹理信息, 也能减少网络参数量, 再各自通过卷积提取空间信息, 进而捕捉空谱信息, 本文使用 3×3 卷积核提取多分辨率图特征, 其中, x_1 、 x_2 、 x_3 每个通道分别使用 64、128、256 个滤波器. 本文设计了逆向跳跃连接的网络结构, 既能实现特征复用, 达

到特征融合的效果,也一定程度促进了网络的反向传播,避免梯度消失.对应的输出表示为:

$$y_j = \begin{cases} Conv(x_j), & j = 4 \\ Conv(pooling(x_j + y_{j+1})), & j = 1, 2, 3 \end{cases} \quad (7)$$

其中, *Conv* 代表卷积操作、归一化、激活操作,归一化可以加速网络的收敛, *pooling* 表示池化操作,使用 ReLU 激活函数增加网络的非线性,然后逐步向前跳跃连接达成通道级联再次通过卷积提取特征,获取最终空谱融合特征,如图 6 所示 y_1 为最后的输出.

3 实验

3.1 数据集

本文采用 Indian Pines (IP), 武汉大学公开数据集 WHU-Hi-HanChuan 以及 WHU-Hi-HongHu 进行实验^[29,30]. IP 数据集地址为 <https://www.ehu.es/ccwintco/>

index.php/Hyperspectral_Remote_Sensing_Scenes, 武汉大学公开数据集地址为 http://rsidea.whu.edu.cn/resource_WHUHi_sharing.htm.

3 个数据集地物假彩色图以及地物真实图如图 7-图 9 所示.

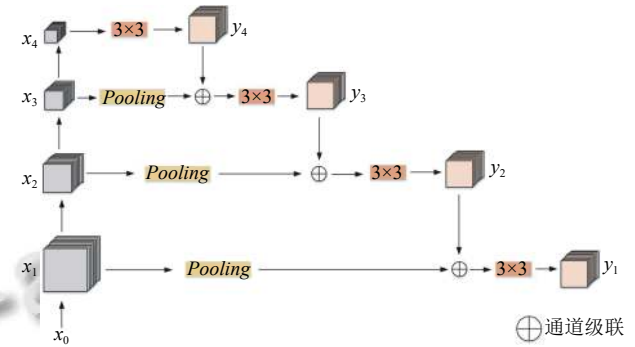


图 6 空间信息提取模块

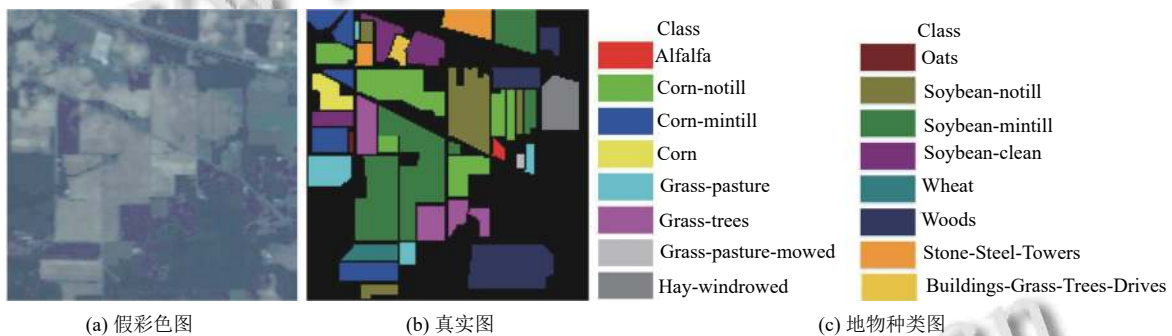


图 7 IP 数据集

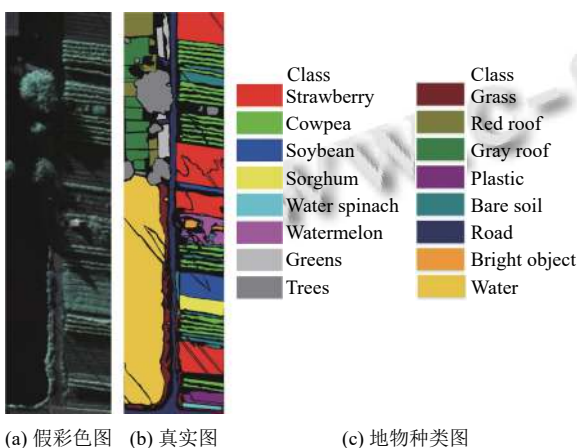


图 8 WHU-Hi-HanChuan 数据集

Indian Pines 数据集是由 AVIRIS 传感器在印第安纳州北部采集的,波长范围 400-600 nm,该数据集主要包含一些农作物、植被、道路和一些建筑,Indian

Pines 地表真实图以及所有种类如图 7 所示.

WHU-Hi-HanChuan 数据集于 2016 年 6 月一天下午在湖北汉川市拍摄采集,研究区域是一个城乡接合部. WHU-Hi-HanChuan 数据集地物真实值以及种类如图 8 所示.

WHU-Hi-HongHu 数据集是在 2017 年 11 月采集,拍摄期间天气多云,实验区是一个复杂的农业场景,作物种类繁多,该地区种植的是同一农作物类型的不同品种,具体种类如图 9 所示.

各个数据集其他详细说明如表 1 内容所示,从表 1 中可以看出 IP 空间分辨率相比 WHU-Hi-HanChuan 以及 WHU-Hi-HongHu 数据集空间分辨率低,WHU-Hi-HongHu 空间分辨最高,达到了厘米级.

3.2 实验设置及评价指标

实验环境系统为 Linux 操作系统,服务器计算版

本为 CUDA 10.0, 处理器配置为 CPU 为 Intel Xeon Silver 4114, Memory 为 256 GB, GPU 显卡配置为 NVIDIA GeForce TITAN Xp. 对比实验方法有 SVM, 残差 2D-CNN, 3D-CNN 选用 DBMA, DBDA, 混合卷积选用 HybridSN 模型, 本模型选用 SGD 优化器, 同时设置动量大小为 0.9, 用来优化目标函数, 学习率设置

为 0.001, 设置批次大小为 30, 迭代次数为 150. IP 数据集的空间输入大小为 64×64, WHU-Hi-HanChuan 以及 WHU-Hi-HongHu 数据的空间输入大小为 24×24. 实验的 Loss 曲线图与 Acc 曲线图如图 10 和图 11 所示, 横坐标代表迭代次数, 纵坐标分别表示 Loss、Acc, 用以证明实验结果的收敛性且没有冗余.

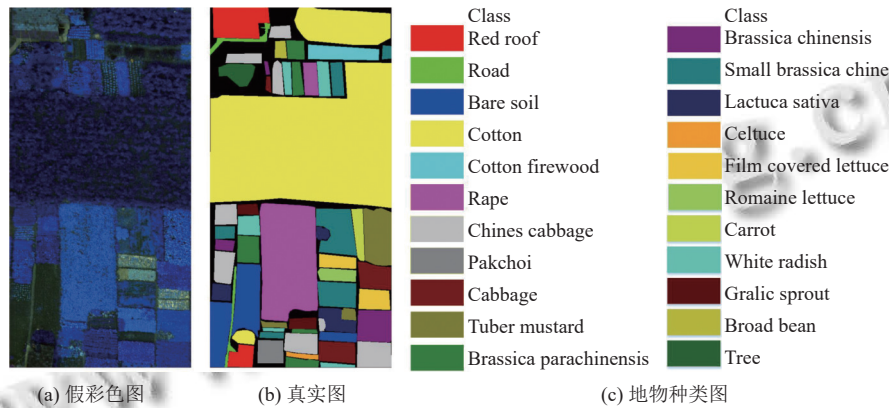


图 9 WHU-Hi-HongHu 数据集

表 1 3 个数据集介绍说明

数据名称	空间大小	空间分辨率 (m)	波段数量	种类数量
IP	145×145	3.7	224	16
WHU-Hi-HanChuan	1217×303	0.109	274	22
WHU-Hi-HongHu	940×475	0.043	270	16

$$OA = \frac{\sum_{i=1}^k n_{ii}}{\sum_{i=1}^k K_i} \quad (8)$$

其中, k 表示图像中地物种类数量, K_i 表示第 i 类像素的数目, n_{ii} 表示第 i 类地物被准确分类的像素数目. 从式 (8) 表达我们可以看出 OA 能够估量所有地物的分类效果, 但无法细致到每一类地物分类情况, OA 这个指标受样本是否均匀的影响.

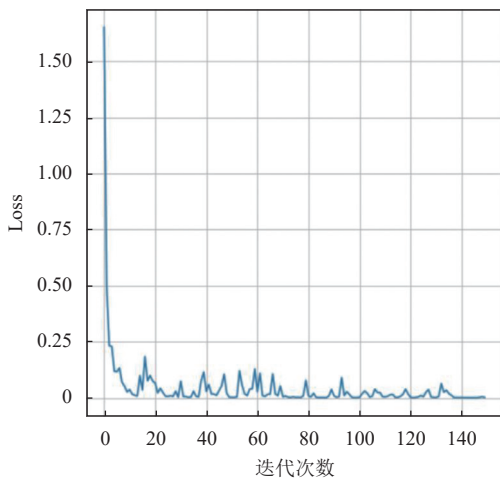


图 10 Loss 曲线图

高光谱图像分类结果的评价指标主要包括: 总体精度 (overall average, OA)、平均精度 (average accuracy, AA) 以及卡帕系数 ($Kappa$).

OA 表示所有测试样本中分类准确的样本数量与所有测试样本的比值, 表达公式如下.

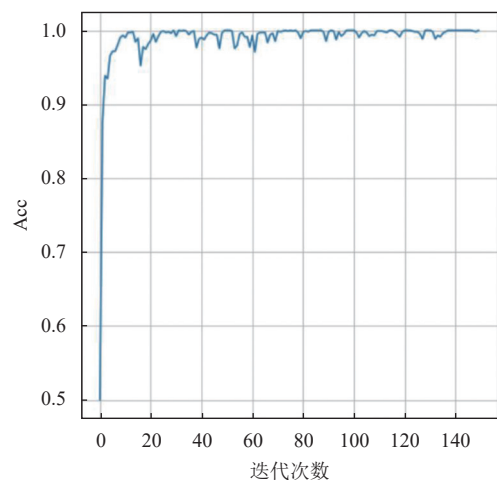


图 11 Acc 曲线图

AA 表达每类分类精度的平均值, 公式如式 (9) 所示. 从式 (9) 可以看出, AA 的值受每一类分类精度的影响, 相比 OA 更加客观.

$$AA = \frac{1}{k} \sum_{i=1}^k \frac{n_{ii}}{K_i} \quad (9)$$

$Kappa$ 系数可以描述真实值与预测值的差异情况, 它与混淆矩阵密切相关, 具体如式 (10) 所示.

$$N = \begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1m} \\ n_{21} & n_{22} & \cdots & n_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ n_{m1} & n_{m2} & \cdots & n_{mm} \end{bmatrix}$$

$$Kappa = \frac{K \sum_{i=1}^k n_{ii} - \sum_{i=1}^k \left(\sum_{j=1}^k n_{ji} \sum_{i=1}^k n_{ii} \right)}{K^2 - \sum_{i=1}^k \left(\sum_{j=1}^k n_{ji} \sum_{i=1}^k n_{ii} \right)} \quad (10)$$

其中, N 表示一个混淆矩阵, n_{ij} 表示第 i 类地物被分到第 j 类的数量, 也就是说当 i 与 j 数值相等时, 表示该类测试样本正确分类的数量, 两者不相等时为错误分类的数量. K 为所有样本数量, 主对角线上的数字越大, 分类效果越理想. 相比 OA 、 AA 、 $Kappa$ 利用了混淆矩阵, 反映出的分类情况更全面.

3 个指标值越高, 分类效果越好. Indian Pines 数据集样本数量较少, 空间分辨率低, 且标记样本也最少, WHU-Hi-HanChuan 以及 WHU-Hi-HongHu 这两个数据集空间分辨率高, 标记样本多. 各自取其 10% 的样本作为训练集, 剩余 90% 作为测试样本. 训练测试集采用随机划分策略, 重复进行 5 次实验取平均值. 表 2—表 4 统计了不同算法在 3 个数据集上的实验结果, 图 12—图 14 列出了不同算法的分类结果图.

表 2—表 4 已经详细统计了每一个数据集每一类的评价指标值, 图 12—图 14 详细展示了不同算法对 3 个数据集的分类结果图. 表 2—表 4 结果对比可以看出 WTCAN 模型相比于其他 5 种方法在 OA 、 AA 、 $Kappa$ 这 3 个指标上表现最佳.

3.3 实验结果分析

(1) IP 数据集实验结果分析. 共计 16 类地物, IP 数据集的不同类具有很高的相似性, 比如 corn-notill、corn-mintill 以及 corn, grass-pasture、grass-trees 以及 grass-pature-mowed, soybean-notill、soybean-mintill 以及 soybean-clean, 它们是同一品种的不同类别, 同时有

限的标记样本也给分类造成了很大的困难. 使用 10% 的样本训练, 本文模型 OA 、 AA 、 $Kappa$ 指标分别达到了 98.41%、98.52%、0.9840, 相比 SVM、2D-CNN、DBMA、DBDA、HybridSN 这 5 种方法 OA 分别提升了 18.47%、18.48%、3.3%、0.54%、2.02%. 从表 2 可以看出, 同一品种不同类别的 corn, grass 以及 soy 分类准确率均达到了 95% 以上. 从图 12 可以看出, SVM、2D-CNN、DBMA 方法的分类结果图中均存在地物 corn-notill 错分成 soybean-clean、地物 soybean-notill 错分成 corn-notill, 地物 corn-mintill 错分 soybean-mintill 的情况, 此外结果图中均出现噪声, 出现噪声的原因是对原始冗余数据没有充分处理, WTCAN 采用因子分析的降维方法起到了不错效果. 本文提出的 WTCAN 模型的分类结果图中在以上提到的这些地物中没有出现错分情况, 结果图也很平滑. 相对来说 DBDA 和 HybridSN 分类结果图也比较令人满意, 但错分地物情况仍然比本文模型多, 如 corn-notill 与 soybean-mintill. 本文模型 IP 多数地物分类上表现良好, 但 buildings-grass-trees-drives 这类地物在所有地物中分类精度最低, 一方面是因为这类样本数量较少、标记样本少, 这导致模型在此类训练不足, 另一方面 buildings-grass-trees-drives 分辨率低, 特征自身有限, 模型的空间信息提取模块池化部分也会丢失一部分信息, 可能导致最终提取到的空谱特征不足, 分类效果不够理想.

(2) WHU-Hi-HanChuan 数据实验分析. 空间分辨率比较高, 波段数量最多, 共有 16 类地物. 其中 red roof 与 gray roof 较为相似, 其余地物之间差异较为明显. 从表 3 的实验数据来看, WTCAN 模型可以将这两类 roof 进行区分, 且能够取得不错的分类效果. 总体 OA 为 99.67%, 相比 SVM、2D-CNN、DBMA、DBDA、HybridSN 方法, OA 各自提升了 16.06%、14.36%、4.56%、3%、7.81%. AA 与 $Kappa$ 也分别达到了 99.05%、0.9939. 从表 3 中也可清晰发现所有地物的分类精度超过 90%, 其中有 12 类地物分类精度达到 99% 以上. 与 IP 数据集分类结果图类似, SVM、2D-CNN、DBMA 方法的分类结果图中出现了噪声, 但噪声明显减少很多, 高空间分辨率图像的信息更有利于地物的精细分类. DBDA、HybridSN 模型方法在略复杂地物区域分类效果不够理想, 容易错分地物.

表 2 不同算法使用 IP 数据集实验结果

算法	准确率 (%)																		
	Alfalfa	Corn-notill	Corn-mintill	Grass-pasture	Grass-trees	Grass-pasture-mowed	Hay-windrowed	Oats	Soybean-notill	Soybean-mintill	Soybean-n-clean	Wheat	Stone-Steel-Towers	Buildings-Grass-Trees-Drives	OA (%)	AA (%)	Kappa		
SVM	60.83	73.30	76.64	62.01	91.67	88.35	83.42	89.89	73.59	75.50	76.22	71.50	91.30	91.44	76.25	97.24	79.94	79.93	0.7698
2DCNN	33.33	67.48	72.16	79.01	97.44	97.25	55.56	88.19	75	77.85	76.79	57.91	88.29	89.34	89.96	88.77	79.93	77.77	0.7624
DBMA	87.80	100	91.00	94.17	96.89	98.12	92.30	100	92.30	94.57	90.84	97.09	100	97.96	94.09	96.05	95.11	95.20	0.9441
DBDA	100	99.20	99.10	100	98.48	99.15	100	100	100	96.38	99.43	98.53	100	99.41	97.12	94.80	97.87	97.85	0.9772
HybridSN	93.94	97.30	96.54	96.74	96.81	95.36	100	92.87	52.94	95.38	96.85	91.99	100	99.91	99.06	86.02	96.39	94.67	0.9588
WTCAN	98.56	99.37	95.29	99.55	99.11	99.78	100	99.54	100	97.18	97.39	97.15	99.96	100	99.57	89.31	98.41	98.52	0.98

表 3 不同算法使用汉川数据集实验结果比较

算法	准确率 (%)																		
	Strawberry	Cowpea	Soybean	Sorghum	Water spinach	Watermelon	Greens	Trees	Grass	Red roof	Gray roof	Plastic	Bare soil	Road	Bright object	Water	OA (%)	AA (%)	Kappa
SVM	87.46	66	73.33	91.30	29.92	29.84	62.84	75.68	62.24	95.78	84.51	54.38	55.81	82	94.11	99.21	83.61	71.52	0.8078
2DCNN	87.41	86.46	87.81	89.57	86.44	83.85	65.27	82.65	84.78	89.96	84.65	82.70	87.27	86.62	75.51	89.17	85.31	81.26	0.8452
DBMA	87.80	100	91	94.17	96.89	98.12	92.31	100	92.31	94.57	90.84	97.09	100	97.96	94.10	96.05	95.10	95.20	0.9440
DBDA	96.64	95.68	96.44	99.19	96	95.07	95.58	98.82	91.73	99.76	98.37	97.71	91.61	98.98	92.37	99.85	96.67	96.49	0.9628
HybridSN	86.44	95.24	98.01	96.70	86.79	89.22	87.82	87.87	90.29	92.31	93.30	92.08	75.15	88.30	97.62	97.17	91.86	90.90	0.9041
WTCAN	99.80	99.74	99.07	99.71	99.54	97.78	99.44	99.43	98.11	94.79	99.57	99.46	98.21	99.80	99.50	99.99	99.64	99.72	0.9934

表 4 不同算法使用洪湖数据集实验结果比较

算法	准确率 (%)																								
	Red roof	Road	Bare soil	Cotton	Cotton firewood	Rape	Chines cabbage	Pakchoi	Cabbage	Tuber mustard	Brassica parachinensis	Small brassica chinensis	Lactuca sativa	Celluce lettuce	Film covered lettuce	Romaine lettuce	Carrot	White radish	Gralic sprout	Broad bean	Tree	OA (%)	AA (%)	Kappa	
SVM	95.97	71.95	79.37	93.58	59.50	86.95	69.09	32.26	95.71	69.99	58.92	56.81	61.24	82.51	84.37	83.05	78.91	54.31	76.44	71.97	40.28	57.39	83.86	70.93	0.7993
2DCNN	97.66	94.76	91.35	98.89	80.03	95.94	85.11	45	98.57	88.52	85.45	80.95	84.06	93.31	92.25	97.29	78.31	82.25	92.99	93.34	61.55	82.32	93.40	86.36	0.9165
DBMA	98.53	87.52	97.07	99.43	94.28	98.88	95.18	91.58	99.23	97.11	94.13	94.21	95.26	99.07	90.28	97.87	92.43	94.45	97.47	96.39	90.75	93.80	97.76	95.22	0.9716
DBDA	99.02	93.72	95.99	99.65	94.81	99.07	95.11	95.99	99.61	97.19	97.57	96.92	96.29	99.01	95.37	99.47	95.13	97.26	96.27	91.11	95.79	96.23	98.21	96.67	0.9774
HybridSN	98.81	93.70	93.30	99.21	89.74	97.84	90.30	86.59	97.15	94.79	93.41	89.44	89.22	93.11	87.54	97.34	95.05	93.57	91.17	92.65	91.10	98.20	96.12	93.33	0.9509
WTCAN	99.67	99.52	99.77	99.95	99.63	99.86	99.07	99.75	99.93	99.71	99.47	99.38	99.42	99.43	96.32	99.66	99.81	98.25	99.26	99.30	98.83	99.78	99.67	99.73	0.9932

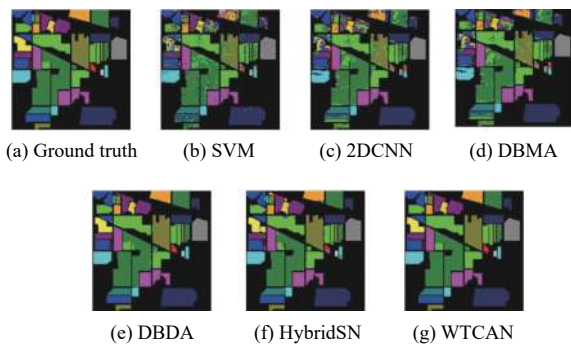


图 12 不同算法对 IP 数据分类结果

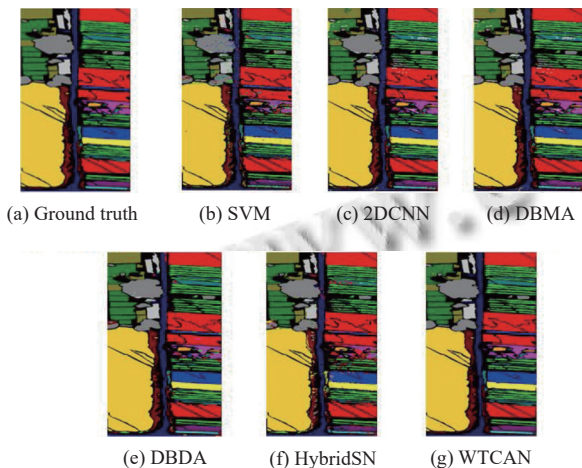


图 13 不同算法对 WHU-Hi-HanChuan 数据分类结果

从图 13 可以看出, watermelon 这个地物区域较为复杂, 特别是一些小区域, SVM、2D-CNN、DBMA、DBDA、HybridSN 在此区域中均有把 watermelon 错分成 cowpea 的情况, WTCAN 模型分类结果图中可以看到不管区域大小都能将 watermelon 精准分类, 没有错分, 说明模型空间多尺度特征提取到了利于地物识别和检测的关键信息. WTCAN 模型表现出了良好的分类性能, 在地物边界区域也能比较准确将其识别, 足以表现出模型对地物空谱特征提取充分.

WHU-Hi-HongHu 数据实验分析. 空间分辨率达到了 0.043 m, 波段数量也较多. 地物种类数量最多, 共有 22 类地物, 18 种农作物, 相似类型较多, 主要有 cotton 与 cotton firewood, brassica chinesis 与 small brassica chinensis, carrot 与 white radish, 地物区域相对比较复杂. 从表 4 可以看出, WTCAN 模型的 OA 相比 SVM、2D-CNN、DBMA、DBDA、HybridSN 分别提升了 15.86%, 4.32%, 1.96%, 1.51%, 3.6%, AA 和 $Kappa$ 也达到了 99.73%、0.9973. 所有地物种类准确率达到 95%, 19 类地物分类精度达到了 99% 以上. 对相似地物也较

为准确进行了分类. 从图 14 结果图可以看出, SVM、2D-CNN 分类结果仍然出现一些噪声, DBMA、DBDA、Hybrid SN 算法总体来说对不复杂区域分类较为准确, 但是在一些紧邻的小区域的地物分类错分情况比较多, 比如 DBDA 模型把 small brassica chinensis 错分成 brassica chinensis, rape 错分成 bare soil, HybridSN 模型把 rape 错分成 small brassica chinensis, small brassica chinensis 错分成 cotton, WTCAN 模型总体对这些紧邻小区域地物、地物之间边界以及相似地物的分类能够达到令人满意结果.

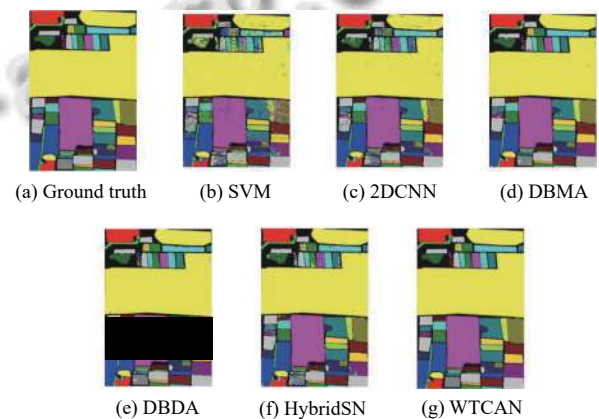


图 14 不同算法对 WHU-Hi-HongHu 数据分类结果

综上所述, WTCAN 在低空间分辨率和高光谱数据上实验取得了较理想的分类结果. WTCAN 模型在 WHU-Hi-HanChuan 数据、WHU-Hi-HongHu 数据的分类效果优于 Indian Pines 数据. SVM 分类器破坏了原有光谱空间结构, 传统 2D-CNN 对光谱信息提取严重缺乏, DBMA、DBDA 的 3D-CNN 模型训练参数量大, HybridSN 模型泛化能力较差, 相比 SVM、2D-CNN、DBMA、DBDA 以及 HybridSN, WTCAN 模型优点是能够利用小波变换层次性有效获取到光谱信息, 相比 3D-CNN 能显著减少计算量, 空间信息提取模块从不同尺度提取空间特征、融合特征信息, 使得最终结果图像的分类边界和边缘更平滑, 复杂地域的地物分类识别也更准确, 同时空间信息提取模块中也引入了金字塔注意力机制, 将注意力集中在更具辨别性的信息上, 提供更详细、更准确的分类图, 但是单独考虑光谱、空间信息也有一些缺点, 小波变换的方式虽然降低了计算量, 但是一定程度上也会忽略部分光谱信息, 空间信息提取模块不同分辨率特征图采用最大池化缩略也会丢失一小部分空间纹理构造信息, 但总体来说 WTCAN 模型能够更快更有效的提取光谱空间特征,

进而提高分类精度。

3.4 消融实验

为了捕捉到更多有效光谱空间位置信息,减少细节信息丢失,在空间信息提取模块中引入金字塔注意力机制,特别对于高光谱图像一些复杂细小区域或者边界,利用不同尺度充分获取当前的局部像素所包含空间纹理信息。为了验证空间信息提取模块添加金字

塔注意力机制的有效性,本文在3个数据集上进行了消融实验。从表5可以看出,添加金字塔注意力模块的评价指标要比不添加的指标要高,IP数据OA、AA、Kappa分别提升0.95%、1.31%、0.0153, HanChuan数据OA、AA、Kappa各自提升1.02%、1.13%、0.0113, HongHu数据OA、AA、Kappa对应提升0.08%、0.38%、0.0002,证明了添加注意力模块的有效性。

表5 消融实验结果对比

数据集名称	评价指标 (不添加注意力模块)			评价指标 (添加注意力模块)		
	OA (%)	AA (%)	Kappa	OA (%)	AA (%)	Kappa
IP	97.46	97.21	0.9687	98.41	98.52	0.9840
WHU-Hi-HanChuan	98.62	98.59	0.9821	99.64	99.72	0.9934
WHU-Hi-HongHu	99.64	99.35	0.9932	99.72	99.73	0.9934

4 结束语

高光谱图像中存在光谱高维度、光谱数量多、信息冗余等特征,造成空间和纹理特征提取困难这一问题,本文提出了一种基于小波变换的2D-CNN的方法,采用多层小波分解方法对高光谱图像光谱特征提取,有效解决了信息冗余问题,相比3D-CNN也能够大幅减少参数量,本文设计了空间信息提取模块充分获取空间纹理信息,引入金字塔注意力机制从不同尺度提取高光谱图像空间光谱特征,最终实现地物分类。在Indian Pines、WHU_Hi_HanChuan、WHU_Hi_HongHu这3个数据集上进行实验,结果表明高光谱图像光谱降成三维之后在低空间分辨率和高空间分辨率数据上对各个目标仍能有效分类,分类精度能有所提升。但是当该算法面对更多光谱波段数量的数据时,使用波段数量过少可能不仅没有解决光谱信息冗余问题,或许还会导致丢失不少重要光谱信息,进而出现分类准确率低的问题。后续可以继续重点研究面向大数据量、光谱波段数量更多的空谱特征提取方法,如何高效利用高光谱图像全局信息,改善网络结构,加速模型收敛速度,实现快速实时分类是我们要继续探究的问题。

参考文献

- Weber C, Aguejdad R, Briottet X, *et al.* Hyperspectral imagery for environmental urban planning. Proceedings of the 2018 IEEE International Geoscience and Remote Sensing. Valencia: IEEE, 2018. 1628–1631.
- Yan LB, Zhao M, Wang XH, *et al.* Object detection in hyperspectral images. IEEE Signal Processing Letters, 2021, 28: 508–512. [doi: 10.1109/LSP.2021.3059204]
- Ibrahim A, Franz B, Ahmad Z, *et al.* Atmospheric correction for hyperspectral ocean color retrieval with application to the hyperspectral imager for the coastal ocean (HICO). Remote Sensing of Environment, 2018, 204: 60–75. [doi: 10.1016/j.rse.2017.10.041]
- Zhou J, Kwan C, Ayhan B, *et al.* A novel cluster kernel RX algorithm for anomaly and change detection using hyperspectral images. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(11): 6497–6504. [doi: 10.1109/TGRS.2016.2585495]
- Zhang YS, Ma Y, Dai XB, *et al.* Locality-constrained sparse representation for hyperspectral image classification. Information Sciences, 2021, 546: 858–870. [doi: 10.1016/j.ins.2020.09.009]
- Okwuashi O, Ndehedehe CE. Deep support vector machine for hyperspectral image classification. Pattern Recognition, 2020, 103: 107298. [doi: 10.1016/j.patcog.2020.107298]
- Peng JT, Jiang X, Chen N, *et al.* Local adaptive joint sparse representation for hyperspectral image classification. Neurocomputing, 2019, 334: 239–248. [doi: 10.1016/j.neucom.2019.01.034]
- Zhang YQ, Cao G, Li XS, *et al.* Active semi-supervised random forest for hyperspectral image classification. Remote Sensing, 2019, 11(24): 2974. [doi: 10.3390/rs11242974]
- Kuo BC, Ho HH, Li CH, *et al.* A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7(1): 317–326. [doi: 10.1109/JSTARS.2013.2262926]
- Zhong P, Gong ZQ, Li ST, *et al.* Learning to diversify deep belief networks for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing, 2017,

- 55(6): 3516–3530. [doi: [10.1109/TGRS.2017.2675902](https://doi.org/10.1109/TGRS.2017.2675902)]
- 11 Ma XR, Wang HY, Geng J. Spectral-spatial classification of hyperspectral image based on deep auto-encoder. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, 9(9): 4073–4085. [doi: [10.1109/JSTARS.2016.2517204](https://doi.org/10.1109/JSTARS.2016.2517204)]
- 12 郭国栋. 基于深度学习的高光谱图像分类和异常检测研究 [硕士学位论文]. 北京: 北京化工大学, 2018.
- 13 Xu YH, Du B, Zhang F, *et al.* Hyperspectral image classification via a random patches network. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 142: 344–357. [doi: [10.1016/j.isprsjprs.2018.05.014](https://doi.org/10.1016/j.isprsjprs.2018.05.014)]
- 14 Li Y, Xu QZ, Li W, *et al.* Automatic clustering-based two-branch CNN for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(9): 7803–7816. [doi: [10.1109/TGRS.2020.3038425](https://doi.org/10.1109/TGRS.2020.3038425)]
- 15 郑宗生, 刘海霞, 王振华, 等. 改进 3D-CNN 的高光谱图像地物分类方法. *自然资源遥感*. 2022: 1–7. <http://kns.cnki.net/kcms/detail/10.1759.P.20220729.1043.012.html>.
- 16 郑姗姗, 刘文, 单锐, 等. 一种改进多尺度三维残差网络的高光谱图像分类方法. *计算机工程*, 2020, 46(12): 215–221. [doi: [10.19678/j.issn.1000-3428.0056791](https://doi.org/10.19678/j.issn.1000-3428.0056791)]
- 17 Zhong ZL, Li J, Luo ZM, *et al.* Spectral-spatial residual network for hyperspectral image classification: A 3-D deep learning framework. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(2): 847–858. [doi: [10.1109/TGRS.2017.2755542](https://doi.org/10.1109/TGRS.2017.2755542)]
- 18 Gao HM, Chen ZH, Li CM. Sandwich convolutional neural network for hyperspectral image classification using spectral feature enhancement. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 3006–3015. [doi: [10.1109/JSTARS.2021.3062872](https://doi.org/10.1109/JSTARS.2021.3062872)]
- 19 Roy SK, Krishna G, Dubey SR, *et al.* HybridSN: Exploring 3-D–2-D CNN feature hierarchy for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 2020, 17(2): 277–281. [doi: [10.1109/LGRS.2019.2918719](https://doi.org/10.1109/LGRS.2019.2918719)]
- 20 徐沁, 梁玉莲, 王冬越, 等. 基于 SE-Res2Net 与多尺度空谱融合注意力机制的高光谱图像分类. *计算机辅助设计与图形学学报*, 2021, 33(11): 1726–1734.
- 21 王欣, 樊彦国. 基于改进 DenseNet 和空谱注意力机制的高光谱图像分类. *激光与光电子学进展*, 2022, 59(2): 0210014.
- 22 Li R, Zheng SY, Duan CX, *et al.* Classification of hyperspectral image based on double-branch dual-attention mechanism network. *Remote Sensing*, 2020, 12(3): 582. [doi: [10.3390/rs12030582](https://doi.org/10.3390/rs12030582)]
- 23 Wu HJ, Li D, Wang YJ, *et al.* Hyperspectral image classification based on two-branch spectral-spatial-feature attention network. *Remote Sensing*, 2021, 13(21): 4262. [doi: [10.3390/rs13214262](https://doi.org/10.3390/rs13214262)]
- 24 Sun H, Zheng XT, Lu XQ, *et al.* Spectral-spatial attention network for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(5): 3232–3245. [doi: [10.1109/TGRS.2019.2951160](https://doi.org/10.1109/TGRS.2019.2951160)]
- 25 Qing Y, Liu WY. Hyperspectral image classification based on multi-scale residual network with attention mechanism. *Remote Sensing*, 2021, 13(3): 335. [doi: [10.3390/rs13030335](https://doi.org/10.3390/rs13030335)]
- 26 Wang QL, Wu BG, Zhu PF, *et al.* ECA-Net: Efficient channel attention for deep convolutional neural networks. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 11531–11539.
- 27 Zhang H, Zu KK, Lu J, *et al.* EPSANet: An efficient pyramid squeeze attention block on convolutional neural network. *Proceedings of the 16th Asian Conference on Computer Vision on Computer Vision*. Macao: Springer, 2023. 541–557.
- 28 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 7132–7141.
- 29 Zhong YF, Hu X, Luo C, *et al.* WHU-Hi: UAV-borne hyperspectral with high spatial resolution (H^2) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with CRF. *Remote Sensing of Environment*, 2020, 250: 112012. [doi: [10.1016/j.rse.2020.112012](https://doi.org/10.1016/j.rse.2020.112012)]
- 30 Zhong YF, Wang XY, Xu Y, *et al.* Mini-UAV-borne hyperspectral remote sensing: From observation and processing to applications. *IEEE Geoscience and Remote Sensing Magazine*, 2018, 6(4): 46–62. [doi: [10.1109/MGRS.2018.2867592](https://doi.org/10.1109/MGRS.2018.2867592)]

(校对责编: 牛欣悦)