

# 基于全局注意力信息交互的高光谱图像分类<sup>①</sup>

王雷全, 周家梁, 林 瑶

(中国石油大学(华东)青岛软件学院、计算机科学与技术学院, 青岛 266580)  
通信作者: 王雷全, E-mail: [richiewlq@gmail.com](mailto:richiewlq@gmail.com)



**摘 要:** 近年来, 研究者们发现基于双分支结构的高光谱图像分类方法可以更有效地提取图像的光谱特征和空间特征用于分类. 但在双分支结构中, 各分支只侧重于细化、提取光谱特征或空间特征, 忽略了对光谱-空间跨维特征交互的研究, 且两分支各自提取的部分交互不明显, 因此影响了分类的性能. 针对这一问题, 本文提出了一种基于全局注意力信息交互的高光谱图像分类方法. 首先采用密集连接网络分两个分支分别细化图像的光谱特征和空间特征, 然后结合全局注意力机制 (GAM) 得到通道全局注意力特征和空间全局注意力特征, 最后通过一个信息交互的模块实现光谱和空间信息的交互, 更充分地利用光谱和空间信息实现分类. 本文提出的方法分别在 Pavia University (PU) 和 Salinas Valley (SV) 两个数据集上进行了实验, 相较于其他的 4 种方法, 本文提出的方法在分类性能上取得了明显的提升.

**关键词:** 图像处理; 高光谱图像分类; 信息交互; 注意力机制; 跨维信息

引用格式: 王雷全, 周家梁, 林瑶. 基于全局注意力信息交互的高光谱图像分类. 计算机系统应用, 2023, 32(5): 28-35. <http://www.c-s-a.org.cn/1003-3254/9120.html>

## Hyperspectral Image Classification Based on Global Attention Information Interaction

WANG Lei-Quan, ZHOU Jia-Liang, LIN Yao

(Qingdao Institute of Software & College of Computer Science and Technology, China University of Petroleum, Qingdao 266580, China)

**Abstract:** In recent years, researchers have found that the hyperspectral image classification method based on dual branch structure can more effectively extract the spectral and spatial features of the image for classification. However, in the dual branch structure, each branch only focuses on refining and extracting spectral or spatial features, with the study on cross-dimensional spectral-spatial feature interaction ignored, and the partial interaction extracted by the two branches respectively is not obvious, which affects the performance of classification. To solve this problem, this study proposes a hyperspectral image classification method based on global attention information interaction. First, the dense connection network is used to divide the image into two branches to refine the spectral and spatial features, respectively, and then the channel global attention features and spatial global attention features are obtained by combining the global attention mechanism (GAM). Finally, an information interaction module is used to realize the interaction of spectral and spatial information, which makes full use of spectral and spatial information to achieve classification. The method proposed in this study has been tested on Pavia University (PU) and Salinas Valley (SV) datasets, respectively. Compared with that of the other four methods, the classification performance of the method proposed in this study is significantly improved.

**Key words:** image processing; hyperspectral image classification; information interaction; attention mechanism; cross-dimensional information

① 基金项目: 国家自然科学基金 (62071491); 中央高校基本科研业务费专项 (19CX05003A-11)

收稿时间: 2022-11-09; 修改时间: 2022-12-23; 采用时间: 2023-01-06; csa 在线出版时间: 2023-03-24

CNKI 网络首发时间: 2023-03-27

## 1 引言

高光谱图像 (hyperspectral image, HSI) 是一种高维图像, 它比普通图像具有更丰富的光谱信息, 可反映地物的空间信息和光谱信息, 有庞大的数据量, 目前已广泛应用于军事、农业、地质、海洋、城市规划等各个领域, 有着广阔的前景。为了充分分析原始高光谱图像, 高光谱图像分类起着至关重要的作用, 是走向许多遥感应用的先决条件。

早期的高光谱图像分类中, 学者们将 SVM<sup>[1]</sup>, 随机森林<sup>[2]</sup>, KNN<sup>[3]</sup> 等传统分类算法应用到高光谱图像分类上, 但由于 HSI 的高维特性, 传统的分类方法无法充分挖掘其深层特征, 导致分类效果较差。Chen 等<sup>[4]</sup> 首次将深度学习的概念引入到 HSI 分类中, 通过提取空间-光谱联合特征得到了更优分类结果, 由此揭示了深度学习在 HSI 分类中的巨大潜力。ResNet<sup>[5]</sup> 和 DenseNet<sup>[6]</sup> 的提出, 有效解决了深度神经网络的退化问题和梯度消失问题, 因此, 研究者们开始基于这两种神经网络研究 HSI 的分类方法。Wang 等<sup>[7]</sup> 结合 DenseNet 的思想, 提出一种端到端的快速密集光谱空间卷积神经网络 (FDSSC)。Zhong 等<sup>[8]</sup> 结合 ResNet 提出一种光谱空间残差网络 (SSRN), 其中的残差块可以高效学习和判别丰富的空间特征和光谱特征。

虽然上述模型较传统方法而言具有更优的分类性能, 但它们忽视了卷积核提取的不同特征的重要性对分类结果的影响, 而随着 SENet<sup>[9]</sup>、CBAM<sup>[10]</sup>、BAM<sup>[11]</sup>、TAM<sup>[12]</sup> 等各种注意力机制被提出和应用, 并在图像描述<sup>[13]</sup>、目标检测<sup>[14]</sup>、图像分类<sup>[15]</sup> 等不同的任务中取得了优异的效果, 基于双分支结构结合不同的注意力机制的 HSI 分类方法成为研究热点。Ma 等<sup>[16]</sup> 提出了一种基于卷积块注意力 (CBAM) 机制的双分支多注意神经网络 (DBMA), Li 等<sup>[17]</sup> 提出了一种端到端的双分支双注意力机制网络 (DBDA), 王欣等<sup>[18]</sup> 提出了一种基于改进 DenseNet 和空谱注意力机制的高光谱图像分类, 杨国亮等<sup>[19]</sup> 提出了一种基于快速双分支密集连接网络和双注意力机制的高光谱图像分类。

上述基于双分支结构的方法通过在不同分支中对光谱和空间领域分别应用不同的注意力机制, 有效选择出与任务关联性较高的光谱特征和空间特征, 最终取得了令人满意的分类效果。然而, 这些方法中的光谱分支和空间分支通常是两个单独的支路, 两者之间没有实质性的联系, 最终也只是简单地将提取的光谱特

征和空间特征进行相加, 并不能使光谱和空间信息的互补性得到很好的利用。本文提出了一种基于全局注意力信息交互的高光谱图像分类方法, 该方法首先利用 DenseNet 结合 3D-CNN 有效地提取图像的光谱特征和空间特征, 然后引入了全局注意力机制 (GAM)<sup>[20]</sup> 来重点关注特征图中全局维的交互信息, 最终通过一种显式的信息交互方法来得到光谱特征和空间特征的混合信息并利用该信息特征实现分类。

## 2 网络框架

本文方法概况如下: 首先通过一个特征提取网络分别对图像的光谱领域和空间领域的特征进行细化, 之后结合 GAM 得到各自领域的全局注意力特征, 最后利用一个显式的信息交互模块进行光谱信息和空间信息的交互并通过一个全连接层得到最终分类的分类结果。

### 2.1 总体架构

本文方法的总体架构如图 1 所示, 主要包括特征提取模块、通道全局注意力和空间全局注意力模块、信息交互模块 3 个主要部分。该方法首先分为光谱分支和空间分支两个分支, 基于 DenseNet 框架结合卷积核大小不一致的 3D-CNN 分别对高光谱图像的光谱特征和空间特征进行细化, 然后在两个分支中引入全局注意力 GAM 对提取的特征进行权重分配, 再将两个分支中得到的注意力特征输入到一个显式的光谱-空间信息交互模块中, 最终通过显式交互后得到高光谱图像的分类结果。

### 2.2 光谱特征和空间特征的提取

在本方法的特征提取部分, 我们采用了 3D-DenseNet, 即 DenseNet 结构结合 3D 卷积作为特征提取网络。相较于探索单一特征信息的 2D 卷积, 3D 卷积可以从具有高维特性的高光谱图像中同时学习空间信息和光谱信息, 挖掘出更丰富的特征知识。而 DenseNet 采用的密集连接机制, 互相连接所有的层, 每个层将前面所有层的输出在 channel 上进行 concat 操作后作为当前层的输入, 进而实现特征重用, 不仅缓解了梯度消失的现象, 也使得它在参数和计算量更少的情况下能实现比 ResNet 更优的性能。因此, 在特征提取部分, 3D-DenseNet 会随着参数数量不断增加而持续提高精度, 防止网络性能下降和过拟合现象发生。

在特征提取模块, 我们首先使用一个光谱分支和

一个空间分支分别对输入的高光谱图像的光谱特征和空间特征进行提取, 得到其对应的光谱特征图 and 空间特征图. 光谱分支和空间分支的具体操作如下: 输入的

初始高光谱图像块大小分配为  $p \times p \times B$ , 其中  $p \times p$  为输入高光谱图像的空间块大小, 本方法中  $p$  的取值为 11,  $B$  为高光谱图像的通道数.

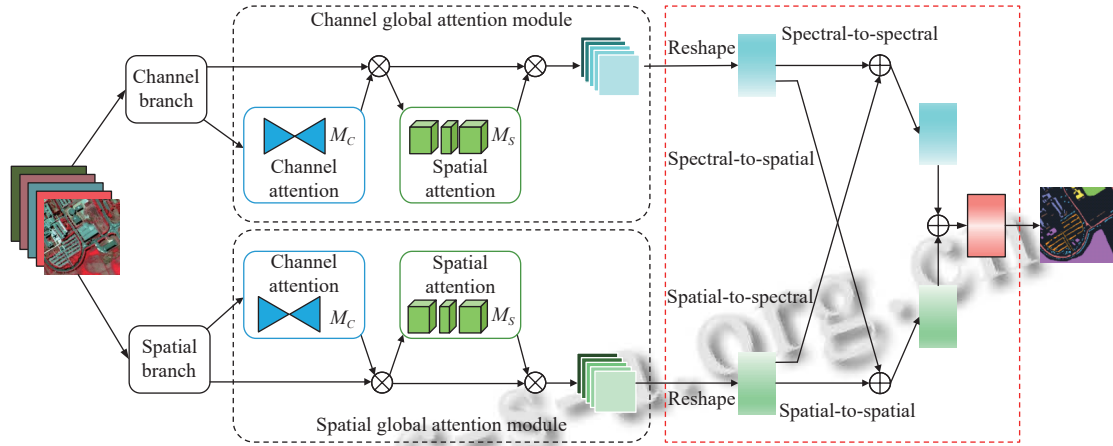


图1 基于全局注意力信息交互的高光谱图像分类结构图

如图 2 所示, 通道分支中, 首先, 使用由 24 个核大小为  $1 \times 1 \times 7$  的 3D 卷积核组成的 3D 卷积块, 以 (1, 1, 2) 的步长向下采样, 从而减少波段数量并得到光谱特征映射 ( $p \times p \times c, 24$ ), 其中  $c = (B - p + 1) / 2$ . 之后, 将得到的光谱特征输入到通道 Dense 块中. 通道 Dense 块由 3 个 BN 和 3D 卷积块结合的模块组成, 每个 3D 卷积块由 12 个核大小为  $1 \times 1 \times 7$  的 3D 卷积核组成, 其采样步长为 (1, 1, 1), 且为了保持特征图输入输出大小的一致性, 我们将其填充属性设置为 (0, 0, 3). 光谱 Dense 块使光谱特征的传递得到了加强, 并且在一定程度上减少了参数量. 经过 Dense 块操作后, 光谱特征的通道数增加到 60, 得到一个 ( $p \times p \times c, 60$ ) 的特征映射. 因此, 通道分支输出特征映射公式为:

$$f_3 = F_3([f_0, f_1, f_2]) \quad (1)$$

其中,  $[f_0, f_1, f_2]$  表示通道维度中来自第 0 层至第 2 层的特征映射.  $F_3$  是一个包含 3 层 Dense 操作的模块, 模块中包含卷积、激活和批量处理归一化操作. 最终, 经过一个 ( $1 \times 1 \times c, 60$ ) 的 3D 卷积块, 得到的光谱特征为  $\Phi$ , 光谱分支的实现细节如表 1 所示.

如图 3 所示, 在空间分支中, 输入的高光谱数据为 ( $p \times p \times B$ ). 首先, 使用一个由 24 个卷积核大小为  $1 \times 1 \times B$  的 3D 卷积核组成的 3D 卷积块将高光谱图像光谱波段压缩到一维, 这部分着重考虑空间信息的提取, 减少光谱通道信息的干扰作用. 然后, 得到空间特征映射 ( $p \times p \times 1, 24$ ). 随后, 将得到的空间特征输入到空间 Dense 块中. 空间 Dense 块同样是由 3 个 3D 卷积块和

BN 块结合的模块组成, 但在空间 Dense 块中, 每个 3D-CNN 的卷积核大小为  $3 \times 3 \times 1$ , 数量为 12 个, 向下采样步长为 (1, 1, 1). 与上述通道分支中 Dense 模块功能基本相同, 都是为了对输入的高光谱图像特征进行特征细化和提取. 空间 Dense 块加强了空间特征的传递, 更有效地利用了空间特征, 最终得到一个 ( $p \times p \times 1, 60$ ) 的空间特征  $\Phi$ . 空间分支的实现细节如表 2 所示.

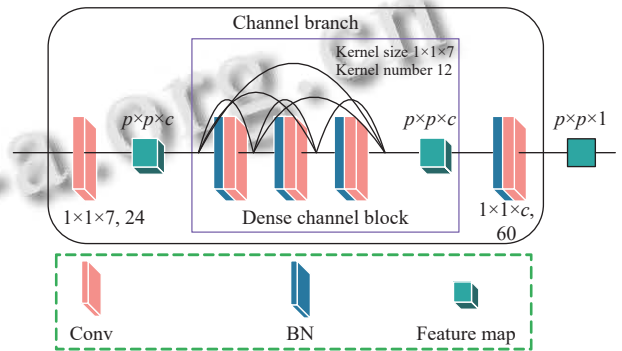


图2 通道分支结构图

### 2.3 光谱全局注意力和空间全局注意力

经过 3D-DenseNet 细化后得到的光谱特征图和空间特征图中, 依然存在另一个维度的特征信息, 因此, 考虑到跨维度的信息相关性, 本方法引入了 GAM 来提取空间特征图和光谱特征图中的全局维度交互特征. 将 3D-DenseNet 细化后得到的光谱特征和空间特征输入到 GAM 注意力模块中, 我们分别得到光谱全局注意力特征和空间全局注意力特征.

表1 通道分支实现细节

Layer name	Kernel size	Output size
Input	—	$(p \times p \times B)$
Conv	$(1 \times 1 \times 7)$	$(p \times p \times c, 24)$
BN-Conv	$(1 \times 1 \times 7)$	$(p \times p \times c, 12)$
Concatenate	—	$(p \times p \times c, 36)$
BN-Conv	$(1 \times 1 \times 7)$	$(p \times p \times c, 12)$
Concatenate	—	$(p \times p \times c, 48)$
BN-Conv	$(1 \times 1 \times 7)$	$(p \times p \times c, 12)$
Concatenate	—	$(p \times p \times c, 60)$
BN-Conv	$(1 \times 1 \times c)$	$(p \times p \times 1, 60)$

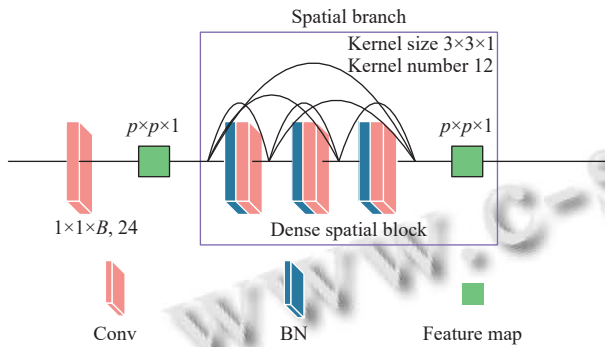


图3 空间分支结构图

表2 空间分支实现细节

Layer name	Kernel size	Output size
Input	—	$(p \times p \times B)$
Conv	$(1 \times 1 \times B)$	$(p \times p \times 1, 24)$
BN-Conv	$(3 \times 3 \times 1)$	$(p \times p \times 1, 12)$
Concatenate	—	$(p \times p \times 1, 36)$
BN-Conv	$(3 \times 3 \times 1)$	$(p \times p \times 1, 12)$
Concatenate	—	$(p \times p \times 1, 48)$
BN-Conv	$(3 \times 3 \times 1)$	$(p \times p \times 1, 12)$
Concatenate	—	$(p \times p \times 1, 60)$

GAM 可在减少信息弥散的同时放大跨维度特征的交互作用, 其整体结构与 CBAM 相似, 包含了通道注意力模块和空间注意力模块两个子模块, 这两个模块分别沿通道和空间两个独立的维度依次进行注意力

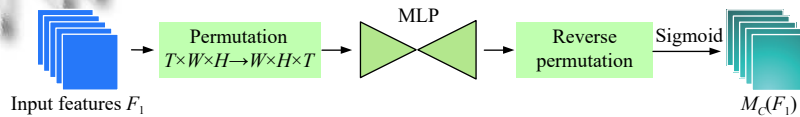


图5 GAM 的通道注意力子模块

### 2.4 光谱-空间信息显式交互模块

本方法中引入了一种挖掘光谱特征和空间特征之间信息的显式交互方法. 该方法受八度卷积模型的启发, 分别在光谱通道融入空间信息以及在空间通道融入光谱信息, 在此卷积层的输出中, 可以得到光谱特征和空间特征的混合信息, 在一定程度上正确地融合光

操作, 最终得到跨维度交互信息的全局特征图.

GAM 的总体结构如图4所示, 给定输入的特征映射  $I_1 \in R^{T \times H \times W}$ , 中间状态  $I_2$  和输出  $I_3$  的定义如式(2)和式(3)所示:

$$I_2 = M_C(I_1) \otimes I_1 \quad (2)$$

$$I_3 = M_S(I_2) \otimes I_2 \quad (3)$$

其中,  $M_C$  和  $M_S$  分别表示通道注意力图和空间注意力图,  $\otimes$  表示按元素进行乘操作.

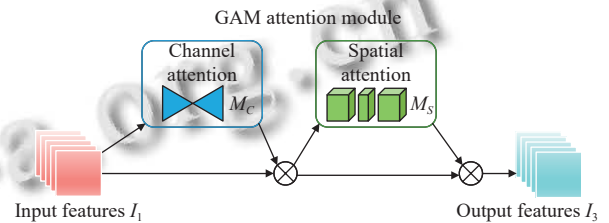


图4 GAM 总体结构图

通道注意力子模块中如图5所示, 其中的  $T$ 、 $W$ 、 $H$  分别为图像块的通道数、空间宽度及长度. 为了尽量保留图像通道信息, 该子模块移除了 CBAM 通道注意力子模块中的池化操作, 首先进行了维度转换, 再经过维度转换的特征图输入到 MLP 来放大跨维通道信息, 然后再转换为原来的维度, 最后用 Sigmoid 处理得到通道注意特征图.

空间注意力子模块如图6所示, 其中  $T$ 、 $W$ 、 $H$  分别为图像块的通道数、空间宽度及长度,  $r$  为缩减比, 我们在该模块中采取了与 BAM 机制相同的缩减比取值, 即  $r$  取值为 4. 主要使用了卷积处理进行空间信息融合, 首先通过卷积核为 7 的卷积缩减通道数量, 缩小计算量, 在经过一个卷积核为 7 的卷积操作, 增加通道数量, 保持通道数量的一致, 最后用 Sigmoid 处理得到空间注意特征图.

谱特征和空间特征, 从而更有效地进行最终的分类, 其结构如图7所示.

图7中输入的  $S_1$  和  $S_2$  分别是经过通道全局注意力和空间全局注意力处理后得到的特征图,  $S_1'$  和  $S_2'$  为经过信息交互后得到的特征图, 最终两特征图级联后得到输出特征图  $S_3$ .

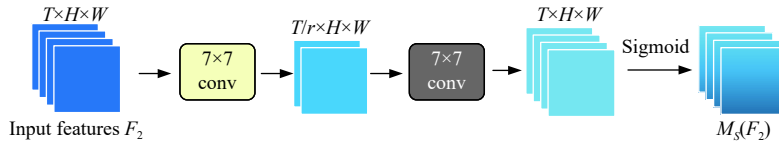


图6 GAM的空间注意力子模块

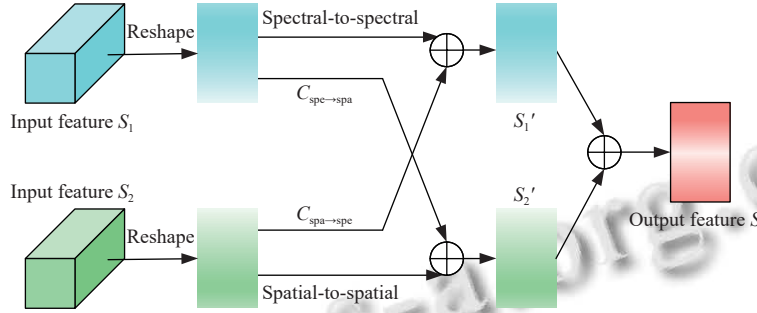


图7 光谱与空间信息交互模块

针对传统卷积模型中存在的空间冗余问题, 将一个图像分解为空间特征和光谱特征, 分别表示全局结构和局部细节. 通过在空间特征和光谱特征之间建立有效的通信, 增大接收域的大小, 使重要的信息流能在光谱和空间特征之间传输, 以便它们的贡献可以更好地融合, 从而获得更多的全局信息. 对于建立光谱与空间特征之间的信息流来说, 最关键的任务是构造一个可学习且可以融合光谱和空间特征的信息交互矩阵, 首先, 将光谱分支输出的光谱特征记作:  $A^{spe} \in \mathbb{R}^{c \times p \times p}$ , 空间分支输出的空间特征记作:  $A^{spa} \in \mathbb{R}^{c \times p \times p}$ , 两个特征都经过变形操作, 分别得到光谱特征  $A^{speW} \in \mathbb{R}^{c \times n}$  和空间特征  $A^{spaW} \in \mathbb{R}^{c \times n}$ , 其中  $n = p \times p$ , 然后, 这两个特征再经过转置操作, 分别得到  $A^{speT} \in \mathbb{R}^{n \times c}$  和  $A^{spaT} \in \mathbb{R}^{n \times c}$ , 接下来, 光谱特征与空间特征之间开始信息传递, 如式 (4) 和式 (5) 所示:

$$C_{spe \rightarrow spa} = [Softmax(A^{spaW} \otimes A^{speT})] \otimes A^{speW} \quad (4)$$

$$C_{spa \rightarrow spe} = [Softmax(A^{speW} \otimes A^{spaT})] \otimes A^{spaW} \quad (5)$$

其中,  $C_{spe \rightarrow spa}$  表示从光谱特征到空间特征的信息流,  $C_{spa \rightarrow spe}$  表示从空间特征到光谱特征的信息流,  $Softmax(A^{spaW} \otimes A^{speT})$  表示从空间特征到光谱特征的空间补充权重, 可以强调空间特征的有用位置, 通过与  $A^{speW}$  做相乘操作, 就可以完成从光谱到空间的信息传输. 同样,  $Softmax(A^{speW} \otimes A^{spaT})$  表示从光谱特征到空间特征的光谱补充权重, 可以突出详细的光谱波段, 通过与  $A^{spaW}$  做相乘操作, 可以获得从空间特征到光谱特征传输的信息. 为了将空间信息整合到光谱特征中, 将  $C_{spa \rightarrow spe}$

与  $A^{speW}$  做相加操作, 以相同的方式, 将  $C_{spe \rightarrow spa}$  与  $A^{spaW}$  做相加操作将光谱信息整合到空间特征中. 如式 (6) 和式 (7) 所示:

$$T^{spe} = C_{spa \rightarrow spe} + A^{speW} \quad (6)$$

$$T^{spa} = C_{spe \rightarrow spa} + A^{spaW} \quad (7)$$

其中,  $T^{spe}$  表示整合后的具有空间信息的光谱特征,  $T^{spa}$  表示整合后具有光谱信息的空间特征. 为了增强这两个特征, 分别对  $T^{spe}$  和  $T^{spa}$  使用交叉熵函数来优化两个分支. 为了使所有的光谱特征和空间特征都参与到分类任务中, 将  $T^{spe}$  和  $T^{spa}$  做相加操作, 得到融合特征  $T^{all}$ , 该特征既包含了所有的重要位置信息, 也强调了光谱波段信息. 最后再将该融合特征图输入到一个全连接层中实现最终的分类.

### 3 实验

#### 3.1 数据集划分

本次实验在 Pavia University (PU)、Salinas Valley (SV) 两个公开的高光谱数据集上进行以验证所提出算法的性能. PU 数据集由  $610 \times 340$  像素组成, 有 115 个光谱反射波段, 光谱分辨率为 4 nm, 去除 12 个受水蒸气吸收影响的噪声通道, 在实验中使用了最后的 103 个波段, 包含 9 类不同生长阶段的不同作物. SV 数据集是由  $512 \times 217$  像素组成, 在 400–2500 nm 之间有 224 个光谱反射波段, 去除 20 个波段后, 实验中保留了最后 204 个波段, 可用的土地覆盖类别涵盖了 16 类. 在数据处理阶段, 将数据集划分为训练集、验证集和

测试集 3 部分, 训练集及其真实标签的作用是更新网络参数, 验证集及其真实标签用于验证在训练阶段生成的临时模型的性能, 测试集用于评估最优模型. 由于 PU 和 SV 数据集样本数量较多, 我们使用 0.5% 的数据作为训练集, 0.5% 的数据作为验证集, 剩下 99% 的数据集作为测试集, PU 和 SV 数据集数据划分详情分别如表 3 和表 4 所示.

表 3 PU 数据集数据划分详情

Order	Class	Total	Train	Validation	Test
1	Asphalt	6631	33	33	6565
2	Meadows	18649	93	93	18463
3	Gravel	2099	10	10	2079
4	Trees	3064	15	15	3034
5	Painted metal sheets	1345	6	6	1333
6	Bare soil	5029	25	25	4979
7	Bitumen	1330	6	6	1318
8	Self-blocking bricks	3682	18	18	3646
9	Shadows	947	4	4	939
Total		42776	210	210	42356

表 4 SV 数据集数据划分详情

Order	Class	Total	Train	Validation	Test
1	Broccoli-green-weeds-1	2009	10	10	1989
2	Broccoli-green-weeds-2	3726	18	18	3690
3	Fallow	1976	9	9	1958
4	Fallow-rough-plow	1394	6	6	1382
5	Fallow-smooth	2678	13	13	2652
6	Stubble	3959	19	19	3921
7	Celery	3579	17	17	3545
8	Grapes-untrained	11271	56	56	11159
9	Soil-vinyard-develop	6203	31	31	6141
10	Corn-senesced-green-weeds	3278	16	16	3246
11	Lettuce-romaine-4wk	1068	5	5	1058
12	Lettuce-romaine-5wk	1927	9	94	1824
13	Lettuce-romaine-6wk	916	4	4	908
14	Lettuce-romaine-7wk	1070	5	5	1060
15	Vinyard-untrained	7268	36	36	7196
16	Vinyard-vertical-trellis	1807	9	9	1789
Total		54129	263	348	53603

### 3.2 实验配置与评价指标

本文的实验都是使用 Python 3.6 版本搭载 PyTorch 学习框架实现的, 实验设备是一台基于 Linux 系统并搭载 NVIDIA GeForce RTX-2070 GPU 的计算机. 在实验中, 我们采用了交叉熵损失作为损失函数, 通过分析比较总体分类精度(OA)、平均分类精度(AA)和 Kappa 系数( $\kappa$ )来评定所提出算法的性能. OA代表正确预测的像素数与总像素数之比. AA指的是在不同的类别中得到的精度的平均值.  $\kappa$ 是指地面真实值和分类结果之间的一致性. 3 个度量值越高代表分类结果越好. 对于本

方法提出的分类网络和进行比较的分类网络, 我们都把批量大小设置为 16, 优化器设置为 Adam 优化器, 学习率设置为 0.0005, 网络训练的最大迭代次数设置为 200 次.

### 3.3 实验结果及分析

为了验证所提算法的性能, 本文把提出的方法与 SVM<sup>[1]</sup>、SSRN<sup>[8]</sup>、DBMA<sup>[16]</sup>、DBDA<sup>[17]</sup> 四种方法在 PU 和 SV 数据集上进行实验并比较分析. 其中, SVM 是基于机器学习的传统分类算法, 其他 3 种方法是基于深度学习的双分支结构的 HSI 分类算法, 经典的双分支结构如图 8 所示.

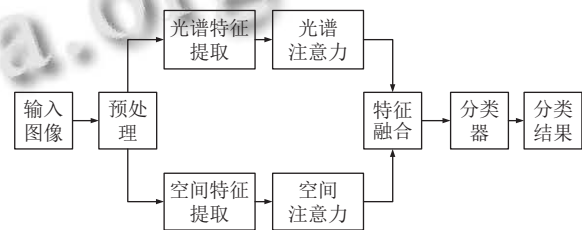


图 8 双分支结构图

对于 PU 数据集和 SV 数据集使用不同方法的分类结果分别如表 5 和表 6 所示, 其中后 3 行分别表示各方法的 OA 值, AA 值以及 Kappa 系数. 不同方法的分类图和地面真实值的分类图分别如图 9 和图 10 所示, 子图题括号内数值表示该分类结果图对应的 OA 值. 可以看出, 本文提出的方法在 PU 数据集和 SV 数据集上的实验结果优于其他 4 种方法.

表 5 PU 数据集上的分类结果 (%)

Order	Class	SVM	SSRN	DBMA	DBDA	Proposed
1	Asphalt	81.98	<b>99.15</b>	96.09	89.03	96.87
2	Meadows	90.92	98.06	98.51	98.32	<b>98.77</b>
3	Gravel	52.64	96.64	89.46	<b>98.70</b>	93.01
4	Trees	94.00	<b>99.86</b>	96.99	98.42	99.65
5	Painted metal sheets	92.69	<b>99.85</b>	97.84	99.78	91.79
6	Bare soil	82.78	96.88	98.11	98.57	<b>99.32</b>
7	Bitumen	57.20	73.24	95.61	<b>95.84</b>	95.15
8	Self-blocking bricks	80.93	82.36	85.51	89.47	<b>91.60</b>
9	Shadows	99.78	<b>100.0</b>	90.82	99.89	88.29
OA		84.86	95.59	95.87	96.00	<b>97.03</b>
AA		81.44	94.01	94.33	<b>96.45</b>	95.59
$\kappa$		79.75	94.15	94.65	94.67	<b>96.05</b>

从实验结果可以看出, 本文提出的方法及其他基于双分支的方法 SSRN、DBMA、DBDA 各项评价指标都要高出 SVM 许多, 从而证明了基于深度学习的分类算法较传统分类方法更适用于高光谱图像分类. 在 PU 数据集的实验结果中, 本文提出的方法得到的 AA

指标值稍低于 DBDA, 分析原因如下: 由于特征信息的交互导致了光谱混合效应, 也就是同种类物体出现了不同的光谱或者是不同种类物体出现了相同的光谱<sup>[21]</sup>, 而个别地物类型在光谱混合效应中表现出的问题更为突出, 因此造成了部分类别的平均分类精度较其他方

法的低. 本文提出的方法在两个数据集上得到的 *OA* 指标值和 *Kappa* 系数都较其他 4 种方法高, 证明了本文方法的有效性, 也展现了在基于双分支结构的高光谱图像分类算法中, 光谱-空间特征信息的交互对高光谱图像分类的重要影响.

表 6 SV 数据集上的分类结果 (%)

Order	Class	SVM	SSRN	DBMA	DBDA	Proposed
1	Brocoli-green-weeds-1	99.69	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	99.98
2	Brocoli-green-weeds-2	99.06	<b>100.0</b>	99.99	99.17	<b>100</b>
3	Fallow	88.22	89.72	97.57	97.74	<b>99.84</b>
4	Fallow-rough-plow	<b>97.68</b>	94.85	92.36	95.95	91.20
5	Fallow-smooth	97.86	99.39	98.51	96.39	<b>99.89</b>
6	Stubble	<b>100.0</b>	99.95	98.98	99.17	99.96
7	Celery	99.29	99.75	98.38	99.83	<b>99.98</b>
8	Grapes-untrained	69.63	88.60	93.94	95.97	<b>96.32</b>
9	Soil-vinyard-develop	96.79	98.48	<b>99.54</b>	99.37	99.32
10	Corn-senesced-green-weeds	83.00	<b>98.81</b>	97.28	96.72	96.56
11	Lettuce-romaine-4wk	92.14	93.30	95.32	93.72	<b>99.78</b>
12	Lettuce-romaine-5wk	78.46	<b>99.95</b>	98.58	99.17	99.27
13	Lettuce-romaine-6wk	91.45	<b>100.0</b>	97.84	<b>100.0</b>	99.59
14	Lettuce-romaine-7wk	93.77	97.86	<b>97.98</b>	96.89	91.77
15	Vinyard-untrained	62.73	89.96	90.51	93.42	<b>93.59</b>
16	Vinyard-vertical-trellis	98.66	<b>100.0</b>	90.72	<b>100.0</b>	99.40
	<i>OA</i>	85.16	94.72	96.31	95.44	<b>97.54</b>
	<i>AA</i>	90.53	96.66	97.11	96.34	<b>97.90</b>
	$\kappa$	83.41	94.12	95.89	94.93	<b>97.27</b>

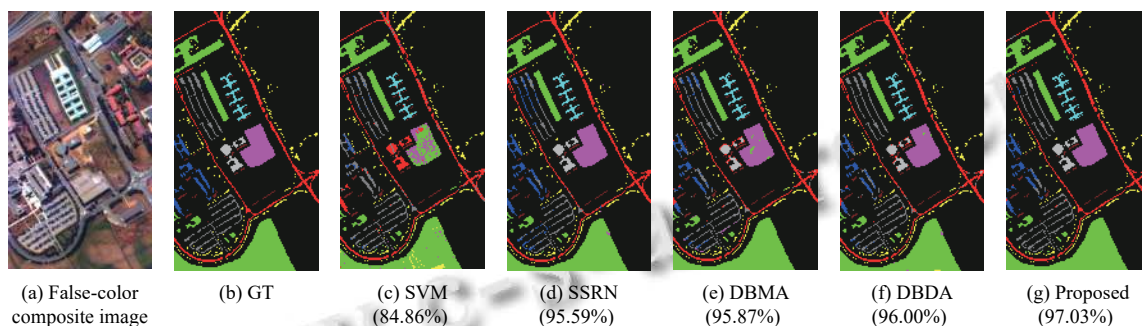


图 9 PU 数据集的分类图

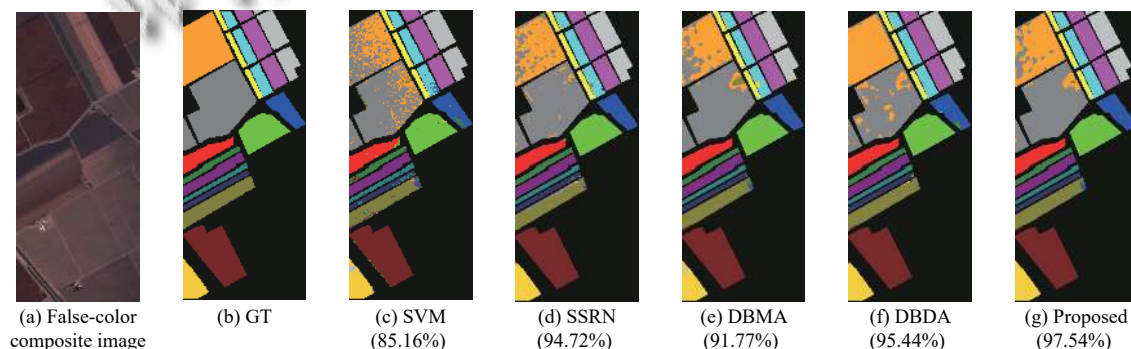


图 10 SV 数据集的分类图

## 4 结论与展望

本文提出了一种以 3D-DenseNet 为特征提取网络, 基于全局注意力信息交互的高光谱图像分类算法。采用密集连接的 3D 卷积来分别细化图像的光谱特征和空间特征, 更充分实现了特征提取, 并减少了参数量; 同时引入 GAM 机制来放大所提取特征中的全局维的交互信息, 有效减少了特征信息的丢失; 最终通过利用光谱特征和空间特征的信息流传输, 实现了光谱-空间信息的交互, 更充分地利用了光谱信息和空间信息, 提高了模型的鲁棒性和泛化性。在两个公共高光谱数据集上对不同的算法进行实验, 结果表明了所提出的算法在 HSI 分类上能获得更为准确的精度。本文所提出的方法虽然在精度上取得了不错的结果, 但是依然存在不足, 且复杂网络对训练样本也提出了更高的要求, 因此我们下一步将尝试引入轻量化的特征提取网络来减少模型的复杂度。而针对实验中出现的不足, 我们将研究如何利用半监督的方法实现对未标记样本的利用, 从而实现样本量的扩增, 减小因样本量少及光谱混合效应对分类效果的负面影响。

### 参考文献

- 1 Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, 42(8): 1778–1790. [doi: 10.1109/TGRS.2004.831865]
- 2 Wang AL, Wang Y, Chen YS. Hyperspectral image classification based on convolutional neural network and random forest. *Remote Sensing Letters*, 2019, 10(11): 1086–1094. [doi: 10.1080/2150704X.2019.1649736]
- 3 Bo CJ, Lu HC, Wang D. Spectral-spatial K-nearest neighbor approach for hyperspectral image classification. *Multimedia Tools and Applications*, 2018, 77(9): 10419–10436. [doi: 10.1007/s11042-017-4403-9]
- 4 Chen YS, Zhao X, Jia XP. Spectral-spatial classification of hyperspectral data based on deep belief network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, 8(6): 2381–2392. [doi: 10.1109/JSTARS.2015.2388577]
- 5 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas: IEEE, 2016. 770–778.
- 6 Huang G, Liu Z, Van Der Maaten L, *et al.* Densely connected convolutional networks. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu: IEEE, 2017. 2261–2269.
- 7 Wang WJ, Dou SG, Jiang ZM, *et al.* A fast dense spectral-spatial convolution network framework for hyperspectral images classification. *Remote Sensing*, 2018, 10(7): 1068. [doi: 10.3390/rs10071068]
- 8 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]
- 9 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.
- 10 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 3–19.
- 11 Park J, Woo S, Lee JY, *et al.* BAM: Bottleneck attention module. *Proceedings of the 2018 British Machine Vision Conference*. Newcastle: BMVC, 2018. 147.
- 12 Misra D, Nalamada T, Arasanipalai AU, *et al.* Rotate to attend: Convolutional triplet attention module. *Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision*. Waikoloa: IEEE, 2021. 3138–3147.
- 13 Anderson P, He XD, Buehler C, *et al.* Bottom-up and top-down attention for image captioning and visual question answering. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 6077–6086.
- 14 Lipton AJ, Fujiyoshi H, Patil RS. Moving target classification and tracking from real-time video. *Proceedings of the 4th IEEE Workshop on Applications of Computer Vision*. Princeton: IEEE, 1998. 8–14.
- 15 Zoran D, Ch�ranowski M, Huang PS, *et al.* Towards robust image classification using sequential attention models. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 9480–9489.
- 16 Ma WP, Yang QF, Wu Y, *et al.* Double-branch multi-attention mechanism network for hyperspectral image classification. *Remote Sensing*, 2019, 11(11): 1307. [doi: 10.3390/rs11111307]
- 17 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]
- 18 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.
- 19 杨国亮, 龚家仁, 刁浩, 等. 基于快速双分支密集连接网络和双注意力机制的高光谱图像分类. *激光与光电子学进展*, 2022, 59(8): 0810002.
- 20 Liu YC, Shao ZR, Hoffmann N. Global attention mechanism: Retain information to enhance channel-spatial interactions. *arXiv:2112.05561*, 2021.
- 21 Ahmad M, Shabbir S, Roy SK, *et al.* Hyperspectral image classification-traditional to deep models: A survey for future prospects. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022, 15: 968–999. [doi: 10.1109/JSTARS.2021.3133021]

(校对责编: 孙君艳)