

孙学进,李靖,曹绍华,等.多尺度注意力聚合图卷积的高光谱分类[J].湖南科技大学学报(自然科学版),2022,37(4):94-103. doi:10.13582/j.cnki.1672-9102.2022.04.012

SUN X J, LI J, CAO S H, et al. Graph Convolution with Multi-Scale Attention Aggregation for Hyperspectral Classification[J]. Journal of Hunan University of Science and Technology (Natural Science Edition), 2022, 37(4): 94-103. doi:10.13582/j.cnki.1672-9102.2022.04.012

# 多尺度注意力聚合图卷积的高光谱分类

孙学进<sup>1\*</sup>,李靖<sup>1</sup>,曹绍华<sup>2</sup>,王娟娟<sup>1</sup>

(1.青岛黄海学院 大数据学院,山东 青岛 266427;2.中国石油大学(华东) 计算机科学与技术学院,山东 青岛 266580)

**摘要:**在遥感图像分类任务中,考虑到传统卷积滤波器受限于接受域,不能有效地捕获地物丰富的细节信息的缺陷,同时,为加强邻域内像元之间的多尺度时空交互,提出一种多尺度注意力聚合图卷积的高光谱图像分类方法.首先,为了增强遥感图像中地物的时空表征能力,构建不同尺度的拓扑图,以实现时空信息的建模.其次,利用图卷积神经网络加强相邻节点之间的交互性,提高时空信息的汇聚和传递,同时,利用自注意力引导层自适应地细化多尺度信息,以捕捉时空信息在通道之间的相关性和位置信息.此外,相邻节点特征差的范数为节点之间的边权重,而层之间信息的传递采用动态聚合方式.试验结果表明:所提出的分类框架在 Indian Pines 基线数据集上的总体分类精度(OA)、平均分类精度(AA)和卡帕系数(Kappa)分别为(99.80±0.15)%,(98.30±1.28)%和(99.77±0.17)%.

**关键词:**遥感图像分类;多尺度时空交互;图卷积网络;动态聚合;自注意力网络;多尺度信息

中图分类号:TP391.9

文献标志码:A

文章编号:1672-9102(2022)04-0094-10

## Graph Convolution with Multi-Scale Attention Aggregation for Hyperspectral Classification

SUN Xuejin<sup>1</sup>, LI Jing<sup>1</sup>, CAO Shaohua<sup>2</sup>, WANG Juanjuan<sup>1</sup>

(1. School of Data Science, Qingdao Huanghai University, Qingdao 266427, China;

2. College of Computer Science and Technology, China University of Petroleum (East China), Qingdao 266580, China)

**Abstract:** In the remote sensing image classification task, considering that the traditional convolution filter is limited in the receptive domain, it cannot effectively capture the rich detailed information of the ground objects. At the same time, it is to strengthen the multi-scale spatio-temporal interaction between the pixels in the neighborhood. A hyperspectral image classification method based on Multi-scale Attention Aggregation Convolution (MAtt-AGCN) is proposed. First, in order to enhance the space-time representation ability of the ground objects in remote sensing images, topological maps of different scales are constructed to realize the modeling of space-time information. Secondly, the graph convolutional neural network is used to strengthen the interaction between adjacent nodes and improve the convergence and transmission of spatiotemporal information; at the same time, the self-attention guidance layer is used to adaptively refine multi-scale information to capture spatiotemporal information in the channel and the correlation and location information. In addition, the norm of the feature difference between adjacent nodes is the edge weight between nodes, and the transmission of

收稿日期:2021-08-23

基金项目:山东省重点研发计划资助项目(2019GGX105001)

\*通信作者,E-mail:sunxuejin123321@163.com

information between layers adopts a dynamic aggregation method. Experimental results show that the overall classification accuracy (OA), average classification accuracy (AA) and Kappa coefficient (Kappa) of the proposed classification framework on the Indian Pines baseline data set are  $(99.80 \pm 0.15)\%$  and  $(98.30 \pm 1.28)\%$ , and  $(99.77 \pm 0.17)\%$ , respectively.

**Keywords:** remote sensing image classification; Multi-scale Attention Aggregation Convolution (MAtt-AGCN); Graph Convolutional Network; dynamic aggregation; self-attention network; multi-scale information

随着遥感技术的不断发展,高光谱图像被广泛地应用于军事目标检测、灾害防治、植被监测和生态环境监测等众多领域<sup>[1]</sup>.高光谱图像由数百个连续的波段组成,包含大量的光谱特征和空间信息<sup>[2]</sup>,存在空间分布不均匀和大量的高纬度冗余信息以及同类别或不同类别间的光谱差异性,给高光谱图像中地物的准确分类和空间-光谱信息的提取带来了巨大挑战.

近年来,国内外学者设计了多种不同的方法用于高光谱影像像元的划分,即将像元划分为特定的土地覆盖类别.早期的地物分类方法主要使用简单的线性分类器进行分类,如k-最近邻和支持向量机(Support Vector Machine, SVM)等<sup>[3-4]</sup>方法,这些方法因其非常简单,在高维高光谱数据分类任务中获得了广泛应用.除此之外,极限学习<sup>[5]</sup>、稀疏表示<sup>[6]</sup>等分类方法也被用来提高高光谱影像中地物的分类精度.然而,仅利用光谱信息很难有效区分不同的土地覆盖类型.考虑到相邻像元在平滑空域内存在相关信息,马尔可夫随机场(Markov Random Field, MRF)和形态学属性剖面<sup>[7-8]</sup>等方法被用于空间和光谱信息的提取,并取得了较好的分类性能.

然而,上述方法都是基于手工制作的光谱空间特征,严重依赖于丰富的专业知识.为了解决这些缺陷,深度学习依靠强大的自学习能力和特征表示<sup>[9-12]</sup>被广泛应用于高光谱图像的地物分类任务.该技术可以通过逐步聚集低级特征自动获得抽象的高级表示,从而有效地避免人工参与特征设计和阈值设置造成的分类误差.如堆叠自动编码器(Stacked AutoEncoder, SAE)和深度置信网络(Deep Blief Network, DBN)<sup>[13-14]</sup>将训练样本扁平化到一维,在分层训练模式下提取深度特征.二维卷积神经网络(2 Dimensions Convolutional Neural Network, 2D-CNN)首先利用主成分分析法对原始输入影像进行降维,然后利用卷积核提取图像中地物的空间特征.但是SAE, DBN和2D-CNN<sup>[15]</sup>并不能直接利用原始高光谱图像的空间和光谱信息,需要进行一系列复杂的预处理,忽略了相邻像元之间的相关性,并且不能有效应对几何分布(如形状或尺度变化较大)相对复杂的地物对象.Li等<sup>[16]</sup>使用深度卷积神经网络学习图像中像元对的特征并融合来自邻域内不同像元对的分类结果;Qing等<sup>[17]</sup>提出了一个融合高效通道注意网络(MRA-NET)的多尺度残差卷积神经网络模型,该模型适用于高光谱图像分类.尽管以卷积神经网络为基础的高光谱图像分类方法能自动提取光谱空间信息,但这些光谱空间特征的有效性仍受到一些限制,如它们简单地将固定的卷积核应用于高光谱影像的不同区域,并未考虑各个局部区域的几何外观变化,同时,在卷积操作过程中,每个卷积核的权值是相同的,因此,在特征提取过程中可能会丢失地物的边界信息,造成不良的误分类.

高光谱数据经常受到噪声的污染,区域内相邻像元之间的边缘权值并不能有效地表示它们之间的内在相似性.此外,传统卷积神经网络(Convolutional Neural Network, CNN)只能利用影像像元的光谱和局部空间特征,而忽略了全局的上下文信息,缺少时空信息之间的交互,同时,因区域内像元之间也存在较大的同类别或不同类别的差异性,在分类过程中会导致边界像元被误分.为了解决这些问题,提出一种多尺度注意力聚合图卷积网络(Multi-scale Attention Aggregation Convolution, MAtt-AGCN)用于高光谱分类.MAtt-AGCN构建了不同尺度的拓扑图,并使用图卷积网络(Graph Convolutional Networks, GCNs)对这些拓扑图进行优化,汇聚和传递有用的细节信息;其次,利用自注意力导向层逐步细化多尺度特征图,以进一步捕获和优化高光谱数据的时空信息.

综上所述,所提MAtt-AGCN算法的主要创新点有以下3个方面:

- 1) 构建了不同尺度的拓扑图,采用图卷积网络(GCNs)对拓扑图进行优化,实现邻域时空信息的汇聚和传递.值得注意的是,拓扑图中的边权重为各自节点的多层感知机(Multilayer Perceptron, MLP)特征

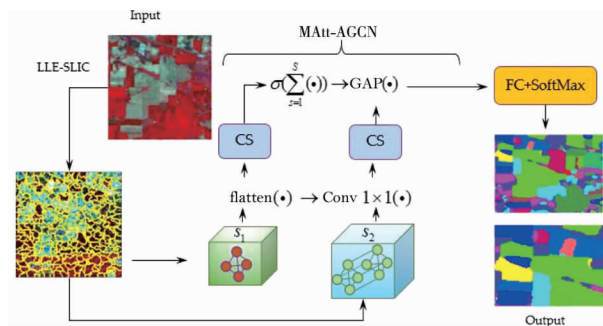
的差值的范数.同时,层间传递过程为动态聚合方式,以增强区域内像元的差异性,提高时空信息的交互.

2) 利用自导注意力网络,广泛地捕捉遥感图像中地物的光谱空间信息,同时,对特征信息进行了逐步细化和聚合,获得了更好的特征表示.

3) 为了降低了算法复杂度,采用局部线性嵌入与线性迭代聚类方法(Locally Linear Embedding-Simple Linear Iterative Clustering, LLE-SLIC)对图像进行预处理,其中局部线性嵌入算法(Locally Linear Embedding, LLE)尽可能地保留高光谱数据中的原始光谱信息,即有用的细节信息.此外,采用定性和定量等方法在 Indian Pines 基线数据集上进行了试验验证.

## 1 MAtt-AGCN 模型

MAtt-AGCN 分类框架的基本组件包括 GCNs、拓扑图的构建、动态聚合策略和双向自导注意力网络(通道注意力和空间注意力),MAtt-AGCN 的网络结构如图 1 所示.



CS 为自注意力引导层;GAP 为全局平均池化操作;flatten(·)为扁平化操作;FC 为全连接层;Conv1×1(·)表示卷积核大小为 1×1 的卷积层

图 1 MAtt-AGCN 的网络结构

### 1.1 多尺度图卷积

图卷积网络(GCNs)是一种能直接在拓扑图上进行卷积操作的神经网络.传统卷积神经网络因接受域的限制,对尺度变化较大的地物的敏感性差,而 GCNs 不仅能有效地捕获非欧几里德数据中的细节信息,同时也能有效汇聚邻域的信息,增强邻域内像元之间的交互,去除不相关的冗余信息,提高特征信息的表征能力.此外,邻域信息在图层间传递时,有助于全局信息的获取.拓扑图的构建和优化步骤如下:

**Step 1** 为了降低计算复杂性,使用 LLE-SLIC 对原始高光谱数据进行非线性降维处理和超像素分割操作,将高光谱数据分割成  $N$  个区域,可表示为  $\mathbf{v} = \{v_1, v_2, \dots, v_N\}$ ,  $v_i = \{x_1, x_2, \dots, x_m\}$ ,  $v_i \in \mathbf{v}$ , 其中  $\mathbf{v}$  为图节点的集合,  $v_i$  为第  $i$  个节点,每个节点  $v_i$  由  $m$  个图像像元组成,即  $x_1, x_2, \dots, x_m$ , 其中  $x_m$  表示该图节点中的第  $m$  个图像像元.

**Step 2** 为了捕获邻域内重要的细节信息以及加强邻域内像元之间的交互,构建了拓扑图  $\zeta$ . 假设  $\mathbf{v}$  为节点,  $\boldsymbol{\varepsilon} = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N\}$  为节点之间的边集合,而拓扑图  $\zeta$  则表示为  $\zeta = (\mathbf{v}, \boldsymbol{\varepsilon})$ .

**Step 3** 为了更好地获取邻域信息,以及对邻域内相邻像元之间的依赖性和相关性建模,使用节点自循环的图卷积网络进行信息汇聚和传递.邻域信息汇聚和传递过程如式(1)所示.

$$\begin{cases} H^{(l)} = \sigma(\tilde{\mathbf{D}}^{\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{\frac{1}{2}} H^{(l-1)} \mathbf{W}^{(l-1)}); \\ \tilde{\mathbf{D}}_{ii} = \sum_j \tilde{\mathbf{A}}_{ij}. \end{cases} \quad (1)$$

式中:  $H^{(l)}$  为第  $l$  层输出的邻域信息;  $\sigma(\cdot)$  为 Leaky Relu 激活函数;  $\tilde{\mathbf{D}}$  为度矩阵(对角矩阵);  $\tilde{\mathbf{A}}$  为使用拉普拉斯重则化后的邻接矩阵;  $\mathbf{W}$  为权重矩阵;  $l$  为层数;  $\tilde{\mathbf{A}}_i, \tilde{\mathbf{D}}_i$  分别为第  $i$  个节点的邻接矩阵和度矩阵.

由于邻域信息中含有大量的冗余信息,同时,进一步加强邻域内节点像元之间的交互性,突出不同类像元特征的差异性,对邻接矩阵进行优化操作,即使用多层感知机(MLP)的范式作为节点之间的边权重,其优化过程如式(2)所示.

$$\mathbf{A}_{ij} = \begin{cases} \|\sigma_{\text{MLP}}(v_i) - \sigma_{\text{MLP}}(v_j)\|, v_i \neq v_j; \\ 1, v_i = v_j; \\ 0, \text{others.} \end{cases} \quad (2)$$

式中:  $\mathbf{A}_{ij}$  为节点  $v_i$  和节点  $v_j$  的邻接矩阵;  $\|\cdot\|$  为一范式操作;  $\sigma_{\text{MLP}}(v_i), \sigma_{\text{MLP}}(v_j)$  分别为第  $i$  个节点  $v_i$  和第  $j$  个节点  $v_j$  的多层感知机(MLP)特征.

**Step 4** 多尺度信息有助于改善特征的表征能力,因此,多尺度条件下拓扑图在 GCNs 中的表示为

$$\mathbf{H}^{(l)} = \sigma(\tilde{\mathbf{D}}_s^{-\frac{1}{2}} \tilde{\mathbf{A}}_s \tilde{\mathbf{D}}_s^{-\frac{1}{2}} \mathbf{H}_s^{(l-1)} \mathbf{W}_s^{(l-1)}). \quad (3)$$

式中:  $s$  为尺度;  $\mathbf{H}_s^{(l-1)}$  为第  $l-1$  层  $s$  尺度的输出信息,值得注意的是,式(1)中的  $\mathbf{H}^{(l)}$  为单尺度情况下输出的特征信息;  $\mathbf{W}_s^{(l-1)}$  为第  $l-1$  层  $s$  尺度的权重矩阵;  $\tilde{\mathbf{A}}_s$  为  $s$  尺度的邻接矩阵.

高光谱图像中像元和光谱波段在空间和光谱域中拥有相同的权重值,但不同光谱波段和空间像元对呈现的特征信息存在巨大差异,因此,为了解决这些问题,设计了双向自导注意力网络,以更多地关注有用的细节信息和弱化非重要信息.采用通道注意力和空间序列注意力模块<sup>[18-21]</sup>逐步细化特征信息,为通道和序列信息分配不同的权重,分配的权重值越大,则说明该特征点越重要,反之,特征点的重要性越小.

**Step 5** 进行注意力操作前,需要将多尺度图卷积网络输出的特征信息进行扁平化操作,其次,使用卷积核大小为  $1 \times 1$  的卷积层<sup>[14,17]</sup>对不同尺度的拓扑图信息进行变换,并将其作为通道和空间序列注意力模块的输入.此外,卷积核大小为  $1 \times 1$  的卷积层的主要作用是对上一层特征图进行集成,提高特征的质量,同时减少下一层的计算参数量.

$$f_{\xi(v,\epsilon)}^j = \text{Conv1} \times 1[\text{flatten}(\mathbf{H}_s)]. \quad (4)$$

式中:  $f_{\xi(v,\epsilon)}^j$  为经过多尺度图神经网络的输出特征;  $\mathbf{H}_s$  为第  $s$  个尺度的图卷积输出.

**Step 6** 使用通道注意力模块,进一步细化特征信息,即获得细化的光谱注意力特征图.操作过程如式(5)所示.

$$\begin{cases} f_{\text{CA}}(j) = \lambda \sum_{i=1}^C (f_{\xi(v,\epsilon)}^{ij} a_i) + a_j; \\ f_{\xi(v,\epsilon)}^{ij} = \frac{\exp(a_i a_j)}{\sum_{i=1}^C \exp(a_i a_j)}. \end{cases} \quad (5)$$

式中:  $f_{\text{CA}}(j)$  为第  $j$  个节点的通道注意力特征;  $\lambda$  为一个初始为 0 的可学习参数;  $C$  为输入特征的通道;  $f_{\xi(v,\epsilon)}^{ij}$  为第  $i$  个节点和第  $j$  个节点之间输出的多尺度图卷积特征;  $a_i, a_j$  为第  $i$  个节点和第  $j$  个节点的注意力系数.

空间序列注意力模块细化了空间信息,强化了邻域内相邻像元之间的相关性,并进一步加强邻域像元之间的交互性.空间序列注意力的操作过程如式(6)所示.

$$\begin{cases} f_{\text{PA}}(j) = \delta \sum_{i=1}^N (f_{\xi(v,\epsilon)}^{ij} D_i) + a_j; \\ f_{\xi(v,\epsilon)}^{ij} = \frac{\exp(B_i + C_j)}{\sum_{i=1}^N \exp(B_i + C_j)}. \end{cases} \quad (6)$$

式中:  $f_{\text{PA}}(j)$  为第  $j$  个节点的空间位置注意力特征;  $\delta$  为一个初始为 0 的可学习参数;  $D_j$  为第  $j$  个节点的重构特征;  $B_i$  为第  $i$  个节点的原始特征;  $C_j$  为第  $j$  个节点的压缩转换特征.

## 1.2 动态聚合策略

为了进一步利用多尺度信息形成不同信息之间的互补,以及强化依赖关系,对动态聚合后的多尺度信息进行全局平均池化(GAP)操作,最后利用全连接层(Fully Connect, FC)实现分类.具体操作如式(7)所示.

$$\begin{cases} f_{\text{DA}} = \sigma \left[ \sum_{s=1}^s (f_s^{\text{PA}} \oplus f_s^{\text{CA}}) \right]; \\ 0 = \text{GAP}(f_{\text{DA}}). \end{cases} \quad (7)$$

式中:  $f_{DA}$  为动态聚合特征;  $f_s^{PA}$ ,  $f_s^{CA}$  分别为第  $s$  尺度的空间位置注意力特征和通道注意力特征;  $\oplus$  为拼接操作;  $O$  为输出;  $GAP(\cdot)$  为全局平均池化操作.

在高光谱分类任务中, 常常采用交叉熵衡量预测结果与真实值之间差异.

$$\tau_{total} = \sum_{m=1} o_m (\log \sum_{n=1} e^{\bar{o}_n} - \bar{o}_m). \quad (8)$$

式中:  $\tau_{total}$  为总损失;  $\bar{o}_m$ ,  $\bar{o}_n$  为第  $m$  和第  $n$  个图像像元的预测标签.

由上可知: 所提 MAtt-AGCN 分类框架不仅有效地捕获了影像的时空信息, 同时通过聚合这些信息强化了时空特征信息之间的依赖关系, 进一步加强了邻域内像元之间的交互, 最大限度地保留了重要的细节信息和边界像元的属性结构.

## 2 试验结果与分析

为验证所提 MAtt-AGCN 分类框架的有效性, 以 Indian Pines 高光谱基线数据为基础, 将 MAtt-AGCN 分类框架与其他先进方法进行比较. 为了证明所提分类框架中不同组件的性能, 使用不同的组件进行试验验证, 并验证该框架处理区域内边界像元的有效性. 试验以总体分类精度 (Overall Accuracy, OA)、平均分类精度 (Average Accuracy, AA) 和卡帕 (Kappa) 系数为评价指标. 此外, 所有数据均经过 LLE-SLIC 进行预处理操作. OA, AA 和 Kappa 系数的计算如式 (9) 所示.

$$\begin{cases} \text{Kappa} = \frac{OA - p}{1 - p}, p = \frac{GT \times TP}{N \times N}; \\ OA = \frac{TP}{T_{total}}; \\ AA = \frac{1}{N} T_c. \end{cases} \quad (9)$$

式中: OA 为总体分类精度;  $p$  为正确样本精度; GT 为真值样本数量; TP 为预测正确的样本数量;  $N$  为样本数量;  $T_{total}$  为总样本数量;  $T_c$  为所有类的总精度.

### 2.1 数据来源

Indian Pines 数据集<sup>[16-17]</sup>由机载可视红外成像光谱仪 (AVIRIS) 于 1992 年对美国印第安纳州一块印度松树进行成像, 影像尺寸为  $145 \times 145$ , 共包含 220 个光谱波段, 其标记像元个数为 10 249, 地物类别 16 种. 详细信息和地物标签见表 1 和图 2 所示.

表 1 详细的地物类型和像元个数

ID	Class	Samples
1	Alfalfa	46
2	Corn-notill	1 428
3	Corn-mintill	830
4	Corn	237
5	Grass-pasture	483
6	Grass-tress	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2 455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1 265
15	Buildings-Grass-Trees-Drives	386
16	Stone-Steel-Towers	93

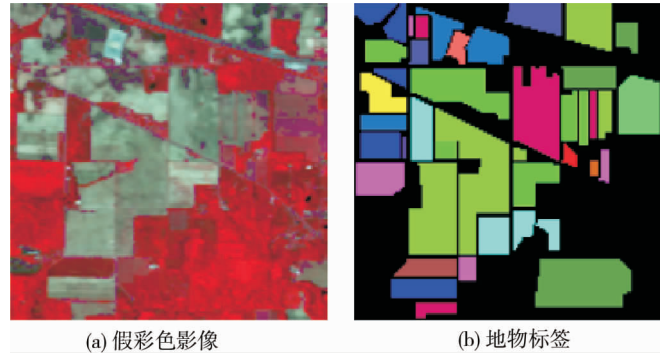


图 2 假彩色合成影像和地物真实影像

## 2.2 试验环境

为了确保试验的有效性,将数据随机划分为 3 个部分,即使用 40% 的训练样本学习网络参数,10% 的验证集进行网络优化,50% 的测试集来验证分类框架的性能.在训练过程中,使用 Adadelta 作为优化器,其中学习率为  $3e-4$ ,权重衰减率为 0.000 1,迭代次数为 800 次,Dropout 为 0.25,图卷积的层数为 2,尺度为 2.全部试验均在 2 个 RTX 3080 的 GPU 显卡上进行,Python 版本为 3.6.10,Torch 版本为 1.7.1+cu110.

## 2.3 与其他先进方法的比较

为了验证所提 MAtt-AGCN 分类框架的有效性,将 MAtt-AGCN 与多个先进方法进行比较,并以定量和定性的方式评估所提 MAtt-AGCN 方法的分类性能.此外,为了进一步保证试验的公平性,所有方法均采用相同的预处理,即采用 LLE-SLIC 表示非线性降维的超像素分割.具体试验和可视化结果见表 2 和图 3 所示.图 3a 为真实的地物分布情况,图 3b 为 SVM 分类,图 3c 为光谱残差网络(Spectral-Spatial Residual Network, S-SRN),图 3d 为光谱密集连接卷积网络(Spectral-Spatial Densely Convolution Network, S-SDCN),图 3e 为双分支多注意力网络(Double-Branch Multi-Attention Mechanism Network, DBMA),图 3f 为本文所提 MAtt-AGCN 分类框架.

表 2 不同方法在 Indian Pines 数据集上的试验结果

Model ID	SVM/%	S-SRN/%	S-SDCN/%	DBMA/%	MAtt-AGCN/%
1	0	2.439±0.048 7	48.80±0.315	52.63±0.416	75.00±0.190
2	72.76±0.257 0	91.02±0.083 0	91.42±0.181	90.92±0.072	100.00±0.040
3	59.49±0.348 6	88.20±0.012 4	96.20±0.113	97.56±0.011	100.00±0.070
4	13.93±0.260 1	47.49±0.377 9	84.22±0.027	96.38±0.049	100.00±0.090
5	44.60±0.363 5	84.89±0.080 3	87.03±0.152	99.62±0.004	97.87±0.140
6	76.80±0.385 2	97.11±0.040 5	96.83±0.038	96.52±0.063	100.0±0.060
7	0	0	0	13.33±0.266	100.00±0.100
8	97.18±0.032 0	99.85±0.001 8	99.01±0.015	97.54±0.031	100.00±0.018
9	0	0	0	0	100.00±0.130
10	85.32±0.142 3	86.03±0.144 2	93.22	95.18±0.033	100.00±0.080
11	72.39±0.285 1	79.67±0.221 2	85.63±0.034	93.99±0.074	100.00±0.070
12	80.35±0.226 9	97.64±0.028 3	98.63±0.008	97.72±0.035	100.0±0.016
13	43.80±0.418 7	62.11±0.341 1	77.58±0.388	99.02±0.013	100.00±0.054
14	99.78±0.001 6	93.14±0.136 6	98.73±0.015	95.60±0.053	100.00±0.092
15	30.95±0.386 6	47.26±0.418 1	69.93±0.314	92.27±0.106	100.00±0.380
16	14.45±0.289 1	71.95±0.231 5	71.58±0.356	95.40±0.014	100.00±0.021
OA	71.97±0.150 8	84.58±0.076 7	90.30±0.077	94.50±0.029	99.80±0.130
AA	49.19±0.151 8	65.57±0.091 1	73.61±0.105	82.11±0.042	98.30±0.095
Kappa	68.16±0.170 8	82.51±0.086 3	89.05±0.086	93.76±0.033	99.77±0.010

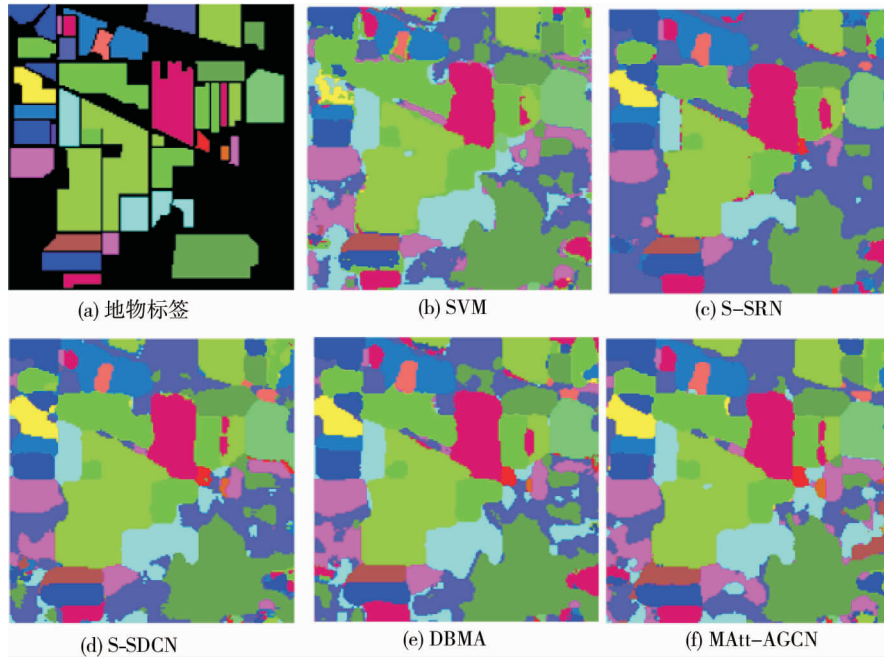


图3 不同模型的分分类效果

由表2可知:相比于其他神经网络,SVM取得了最低的分类效果,其OA,AA和Kappa系数分别为71.97%,49.19%和68.16%,因为SVM忽略了地物的空间邻域信息,而其他神经网络在分类时融合了空间和光谱信息;相比于S-SDCN和S-SRN网络,DBMA由于利用不同分支分别提取地物的空间和光谱特征,并引入注意机制细化特征信息,在3个评价指标上取得了较好的分类性能,而S-SDCN采用密集连接替换剩余连接,其Kappa系数比S-SRN提高了7.92%。

通过表2和图3也可以进一步发现:当训练样本较少时,SVM,S-SDCN,S-SRN和DBMA的分类精度仍然很差,如第9类和第7类地物的分类精度为0,可能出现了过拟合现象,而本文提出的MAtt-AGCN分类框架在训练样本较少的情况下依然取得了最优的分类效果,如第9类和第7类地物的分类精度都为100%。MAtt-AGCN分类框架利用多尺度的图卷积网络获取了更好的邻域信息,同时,通过拓扑图中节点信息的汇聚和传递加强了像元之间的交互,其次,利用自注意力引导层进一步细化特征信息,提高了分类性能。

#### 2.4 不同组件的分类性能

为了分析MAtt-AGCN框架中不同组件对分类性能的影响,在不同组件上进行试验测试,试验结果如图4所示,图4中的数据均为Kappa系数,No-CS表示该框架中不包含自注意力引导组件,No-DA表示不包含动态策略组件。

由图4可知:No-DA和 $l=3, s=2$ 模块对于高光谱遥感影像都有很好的分类效果,但是,在所有情况下,No-DA组件的分类性能明显好于 $l=3, s=2$ 模块,在整个MAtt-AGCN框架中除了使用了多尺度信息之外,No-DA分类较好的主要原因是由于No-DA组件采用了较为合适的图卷积层数,最大限度地捕获了节点之间、不同尺度之间的依赖关系和交互性,而 $l=3, s=2$ 随着图卷积层数的增加,大量的冗余信息被加入。

随着尺度 $s(s=1, 2, 3)$ 和图卷积层 $l(l=1, 2, 3)$ 的增加,分类性能先增加后降低.主要原因是随着图卷积层数的增加,提取的时空信息和物理细节信息(主要从拓扑图中获得)越丰富.值得注意的是,CS和DA组件在提出的MAtt-AGCN分类框架中也产生了性能优势,在不同尺度的拓扑图中其增益是互补的,因此,可以将这些组件有效地结合在一起,以进一步提高网络性能。

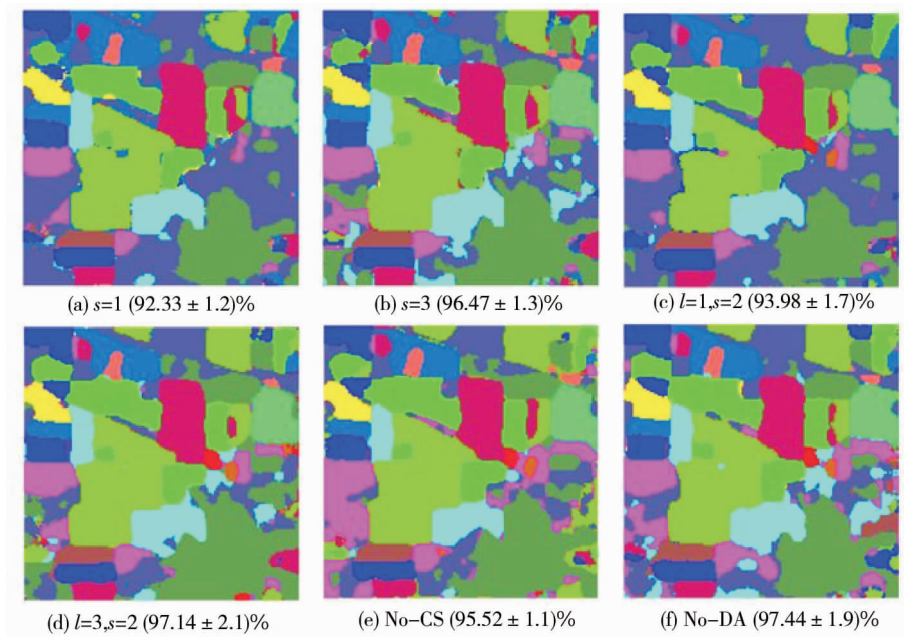


图 4 不同组件的分类性能

## 2.5 消融试验

### 2.5.1 讨论样本数量

为了研究在不同数量的标注样本下,所提 MAtt-AGCN 分类框架和其他分类模型分类精度,因此,将训练样本的数量从 5% 增加到 30%,并在 Indian Pines 基线数据集上进行试验测试.试验结果如图 5 所示.

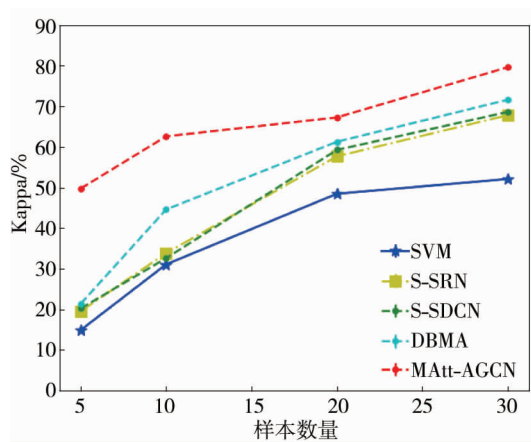


图 5 不同样本数量的 Kappa 系数

由图 5 可知:随着训练样本的增加,所有分类框架的分类性能都在增加,且 SVM 的 Kappa 系数最低,其在训练样本为 5% 时,Kappa 系数为 14.92%.但值得注意的是,DBMA, S-SDCN, S-SRN 在训练样本非常有限的情况下,分类性能也很差,即在训练样本为 5% 时,相应的 Kappa 系数分别为 21.40%, 20.24% 和 19.55%.此外也可以观察到,随着标记样本数量的增加,所提出的 MAtt-AGCN 分类框架的 Kappa 系数始终高于其他分类模型,即当训练样本为 5% 时,MAtt-AGCN 分类框架的 Kappa 系数为 49.77%,进一步表明了所提 MAtt-AGCN 分类框架的有效性和稳定性.

### 2.5.2 不同超像素分割方法的分类性能

不同组件的试验结果如表 3 所示.由表 3 可知:与 SLIC, LDA-SLIC 和 No-LLE-SLIC 方法相比,提出的 MAtt-AGCN 采用非线性超像素分割方法取得了最好的分类性能.相比于 No-LLE-SLIC 方法,SLIC, LDA-SLIC 和 MAtt-AGCN 的 FLOPs 也具有一定的竞争优势.这也表明本文提出的超像素方法能有效降低模型的计算复杂度.



表3 试验结果

Model ID	SLIC	LDA-SLIC	No-LLE-SLIC	No-CA	No-PA	MAtt-AGCN
OA/%	94.16±0.45	96.57±0.23	90.78±0.51	96.77±0.08	97.15±0.18	99.80±0.13
AA/%	91.97±0.840	95.99±0.120	90.25±0.260	95.86±0.150	96.95±0.190	98.30±0.095
Kappa/%	93.85±0.72	96.07±0.26	90.49±0.32	96.42±0.04	97.08±0.24	99.77±0.01
FLOPs/GB	3.7	4.1	11.8	3.9	4.2	4.3

注:LDA-SLIC为线性的超像素分割方法;No-LLE-SLIC为无超像素分割预处理方法;No-CA为无通道注意力模块;No-PA为无空间位置注意力模块;FLOPs为每秒浮点运算次数,主要用于衡量效率和模型复杂度

### 2.5.3 边界区域的分类性能

为了进一步证明所提 MAtt-AGCN 分类框架对边界像元的区分能力,研究并讨论了不同方法在边界区域的分类性能,试验结果如图6所示。

传统卷积神经网络在对图像进行卷积操作时,其卷积核的权值和尺寸大小基本是相同的,这可能会对边界区域内的像元产生误分.与传统卷积网络的固定大小和权重的粗卷积方法不同,所提出的 MAtt-AGCN 分类框架可以灵活、自适应地捕获图像的多尺度和邻域信息,因此,在图节点信息汇聚和传递过程中不会显著地消除待分类地物目标的边界.在 SVM, S-SRN, S-SDCN 和 DBMA 的分类图中(图6b~图6e)可以看出:类边界附近的分类结果比较混乱和不精确,因为这些方法在区分类边界附近像素时,并没有充分利用空间信息,同时也忽略了重要的可辨别信息.相比之下,本文提出的 MAtt-AGCN 分类效果(图6f)比其他方法的分类图更紧凑、更准确。

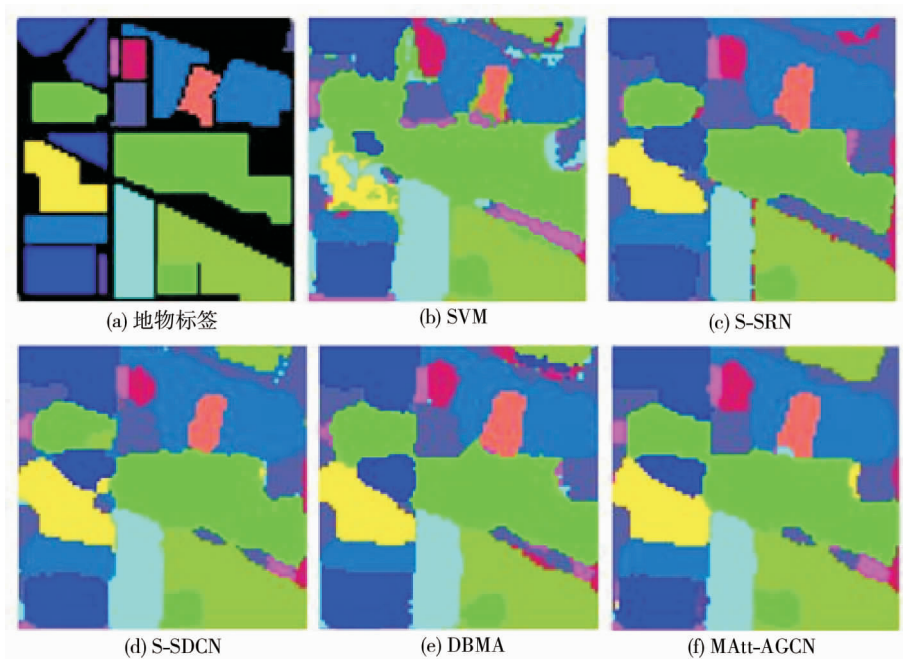


图6 不同方法的边界分类效果

## 3 结论

1) 相邻像元之间存在着较强的关联性,传统卷积方法仅在固定方形区域内实现卷积操作,很难获取较远像元的特征信息,而图卷积方法则通过汇聚邻域内的节点特征,充分建立长距离依赖,加强了特征信息的表示效果。

2) 传统卷积方法很难对边界像元的属性进行正确决策,所提出的图卷积方法能加强邻域像元的交互和依赖,动态聚合了不同尺度所承载的隐藏空间和时序信息,同时引入的双注意力机制有效缓解了冗余信息的使用,进一步提高了特征的表示。

3) 接下来,将使用分层的图卷积网络对遥感影像进行编码,进一步捕获邻域内相邻像元之间的层级结构信息,最终达到改善网络的分类性能的目的。

## 参考文献:

- [1] 刘洪超,董安国.基于非局部方式特征融合的高光谱遥感图像分类算法[J].激光与光电子学进展,2020,57(6):195-201.
- [2] 王凡.基于PCA网络的光谱-空间联合的高光谱图像分类[J].计算机仿真,2018,35(10):350-353.
- [3] Akila D, Bhaumik A, Doss S, et al. Hyperspectral image classification by means of suprepixel representation with KNN[C]//Intelligent Computing and Innovation on Data Science. Springer, Singapore, 2020:369-378.
- [4] Zhao L, Qi J, Ren Z, et al. Shallow water bathymetry retrieving of optical remote sensing combined with SVM bottom classification[C]//2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP). IEEE, 2020: 530-534.
- [5] 何阳,闫德勤,刘德山.局部信息保持极限学习机的遥感图像分类[J].计算机应用与软件,2019,36(1):263-270.
- [6] 梁开阳.基于多尺度变换和稀疏表示的遥感图像融合算法研究[D].银川:北方民族大学,2020.
- [7] Cao X, Wang X, Wang D, et al. Spectral-Spatial Hyperspectral Image Classification Using Cascaded Markov Random Fields[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 12(12):4861-4872.
- [8] 陈志会,卞振奇,赵秀英.基于属性形态学剖面多特征超高分辨率遥感影像分类研究[J].测绘与空间地理信息,2019,42(9):115-116.
- [9] Jiang R. Hyperspectral Remote Sensing Image Classification Based on Deep Learning[J]. Journal of Physics: Conference Series, 2021, 1744(4):042185.
- [10] Chen S, Jin M, Ding J. Hyperspectral remote sensing image classification based on dense residual three-dimensional convolutional neural network[J]. Multimedia Tools and Applications, 2020,80(2):1859-1882.
- [11] Zhang C, Li G, Lei R, et al. Deep Feature Aggregation Network for Hyperspectral Remote Sensing Image Classification[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020,13:5314-5325.
- [12] Man Q, Dong P, Guo H. Pixel and feature-level fusion of hyperspectral and lidar data for urban land-use classification[J]. International Journal of Remote Sensing, 2015,36(6):1618-1644.
- [13] Tong G, Li Y, Cao L, et al. A DBN for hyperspectral remote sensing image classification[C]//2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA). IEEE,2017:1757-1762.
- [14] Mohan A, Venkatesan M. Hybrid CNN based hyperspectral image classification using multiscale spatio-spectral features[J]. Infrared Physics & Technology, 2020, 108:103326.
- [15] Wang L Z, Zhang J B, Liu P, et al. Spectral-spatial multi-feature-based deep learning for hyperspectral remote sensing image classification[J]. Soft Computing, 2017,21(1):213-221.
- [16] Li Z, Huang L, He J. A Multiscale Deep Middle-level Feature Fusion Network for Hyperspectral Classification[J]. Remote Sensing, 2019, 11(6): 695.
- [17] Qing Y, Liu W. Hyperspectral image classification based on multi-scale residual network with attention mechanism[J]. Remote Sensing, 2021, 13(3): 335.
- [18] Liu N, Zhang B, Ma Q, et al. Stack Attention-Pruning Aggregates Multiscale Graph Convolution Networks for Hyperspectral Remote Sensing Image Classification[J]. IEEE Access, 2021, 99:44974-44988.
- [19] Wan S, Gong C, Zhong P, et al. Multiscale Dynamic Graph Convolutional Network for Hyperspectral Image Classification[J]. IEEE Transactions on Geoscience & Remote Sensing, 2019,58(5):3162-3177.
- [20] He N, Fang L, Li Y, et al. High-order self-attention network for remote sensing scene classification[C]//IGARSS 2019 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2019: 3013-3016.
- [21] Fu J, Liu J, Tian H, et al. Dual attention network for scene segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 3146-3154.