

# 一种引入注意力机制的多尺度高光谱图像特征提取方法

许张弛, 郭宝峰\*, 吴文豪, 尤靖云, 苏晓通

杭州电子科技大学自动化学院, 浙江 杭州 310018

**摘要** 近年来,随着深度学习的发展,基于深度学习的特征提取方法在高光谱图像领域表现出良好的发展前景。提出了一种引入注意力机制的多尺度高光谱图像特征提取方法,包括光谱特征提取网络和空间特征提取网络两个部分,并使用一种得分融合策略进行融合。在光谱特征提取网络中,引入注意力机制来缓解因光谱维数过高导致的梯度消失问题,以提取多尺度的光谱特征。在空间特征提取网络中,引入注意力机制作用于网络主干,使其关注邻域内的重要部分,帮助分支网络提取关键信息。将5种光谱特征提取方法、3种空间特征提取方法以及3种空间-光谱联合特征提取方法在3个数据集上进行对比实验,实验结果表明,所提方法能够稳定、有效地提升高光谱图像的分类准确率。

**关键词** 高光谱图像; 长短期记忆网络; 注意力机制; 特征提取; 深度学习

中图分类号 TP751.1

文献标志码 A

DOI: 10.3788/LOP230974

## Multi-Scale Feature Extraction Method of Hyperspectral Image with Attention Mechanism

Xu Zhangchi, Guo Baofeng\*, Wu Wenhao, You Jingyun, Su Xiaotong

School of Automation, Hangzhou Dianzi University, Hangzhou 310018, Zhejiang, China

**Abstract** In recent years, with the development of deep learning, feature extraction methods based on deep learning have shown promising results in hyperspectral data processing. We propose a multi-scale hyperspectral image feature extraction method with an attention mechanism, including two parts that are respectively used to extract spectral features and spatial features. We use a score fusion strategy to combine these features. In the spectral feature extraction network, the attention mechanism is used to alleviate the vanishing gradient problem caused by spectral high-dimension and multi-scale spectral features are extracted. In the spatial feature extraction network, the attention mechanism helps branch networks extract important information by making the network backbone focus on important parts in the neighborhood. Five spectral feature extraction methods, three spatial feature extraction methods and three spatial-spectral joint feature extraction methods are used to perform comparative experiments on three datasets. The experimental results show that the proposed method can steadily and effectively improve the classification accuracy of hyperspectral images.

**Key words** hyperspectral images; long short-term memory; attention mechanism; feature extraction; deep learning

## 1 引言

高光谱成像技术采集到的高光谱图像由包含数百个连续波段的光谱带组成,具有丰富的光谱信息和空间信息。目前,高光谱图像已经被广泛应用于侦察<sup>[1]</sup>、精准农业<sup>[2]</sup>、植被分析<sup>[3]</sup>等领域。高光谱图像提供的丰富光谱信息可用于区分具有略微不同光谱特征的不同

地物,然而,其高维性引起的维数灾难,使得数据处理困难,导致分类精度降低。对于上述问题,使用降维方法可以取得显著的效果,其主要方法分为特征选择和特征提取。特征选择通过从原始光谱波段中选择最具有判别能力的子集来实现降维,该类方法计算量通常较大,并且容易造成人为的波段丢失。特征提取通过将输入数据线性或非线性地转换到一个低维空间来降

收稿日期: 2023-03-28; 修回日期: 2023-04-25; 录用日期: 2023-05-04; 网络首发日期: 2023-05-14

基金项目: 国家自然科学基金(61375011)

通信作者: \*gbf@hdu.edu.cn

低维数,并在低维空间中提取信息的特征。目前,特征提取是高光谱遥感图像处理中的热门方法,在提高分类精度和缓解数据处理压力等方面有重要的研究意义<sup>[4]</sup>。

近年来,深度学习方法发展迅速,并且在高光谱数据处理领域展现出良好的发展前景<sup>[5-6]</sup>。深度学习方法如堆叠式自动编码器(SAE)<sup>[7]</sup>、卷积神经网络(CNN)<sup>[8]</sup>和生成对抗网络(GAN)<sup>[9]</sup>等,主要通过一个端到端的框架来自动完成特征提取和分类任务,其提取的深层特征具有较强的语义信息。高光谱图像具有连续且狭窄的光谱波段,波段之间具有较强的相关性。不同于传统的前馈神经网络,循环神经网络(RNN)将高光谱图像的每个像素视为序列而非高维向量,因此,RNN能够提取光谱波段间的上下文信息。Zhou等<sup>[10]</sup>提出一种基于RNN的空间-光谱(空-谱)联合特征提取方法,使用长短期记忆网络(LSTM)分别提取光谱特征和空间特征,并使用全连接网络进行融合。Liu等<sup>[11]</sup>提出一种基于双向卷积长短期记忆网络(ConvLSTM)的空-谱联合特征提取方法,使用一种将卷积算子并入递归结构的新型LSTM,采取双向递归的方式充分捕获频谱信息。上述方法没有考虑到RNN的梯度消失问题,该问题在处理高光谱图像这类高维数据时尤为明显。Xu等<sup>[12]</sup>提出了一种光谱波段分组策略,采用切分频谱的方式缓解光谱维度的梯度消失现象。Wang等<sup>[13]</sup>在此基础上引入了多尺度策略,提出一种端到端的自适应光谱-空间多尺度网络(ASSMN),该网络包括光谱子网和空间子网两个部分,分别用于提取高光谱图像中光谱与空间上的多尺度上下文信息。然而,ASSMN在使用分组策略的过程中,仍然受到梯度消失现象的限制。此外,在进行空间特征提取的过程中,上述方法均没有考虑到空间邻域中不同部分的重要程度存在差异。注意力机制是一种应对梯度消失问题的有效方法<sup>[14]</sup>,其早期主要用于Seq2Seq(Sequence-to-Sequence)模型,随后研究人员又提出了自注意力机制来处理多对一问题。同时,注意力机制也被广泛应用于图像处理领域,能够有效帮助网络关注到空间信息中的重要部分<sup>[15]</sup>。

本文提出一种引入注意力机制的多尺度高光谱图像特征提取方法,包括光谱子网与空间子网两部分。在光谱子网中,通过一维卷积与分组策略实现感受野与分组多尺度特征提取,并引入光谱注意力机制网络以缓解LSTM的梯度消失现象,扩展多尺度分组策略的可能性。在空间子网中,采用一种卷积-注意力模块,通过空间注意力机制网络辅助网络主体,实现感受野多尺度策略,并通过ConvLSTM进一步提取不同尺度下的空间特征。最后,采用一种得分融合策略,将上述子网进行整合,同时进行训练与融合,得到空-谱联合特征。

## 2 基本原理与网络框架

### 2.1 RNN与LSTM的梯度消失现象

与传统前馈神经网络不同,RNN是一种能够提取数据中上下文信息的网络结构。它通过在隐藏层之间建立连接,以一种递归结构对序列进行建模,使序列当前输出与之前的输出之间建立联系。RNN在单一时间步(time step)中的计算公式为

$$\begin{cases} \hat{y}_i = \sigma_1(\mathbf{V}\mathbf{h}_i) \\ \mathbf{h}_i = \sigma_2(\mathbf{W}\mathbf{h}_{i-1} + \mathbf{U}\mathbf{x}_i) \end{cases}, \quad (1)$$

式中: $\hat{y}_i$ 、 $\mathbf{h}_i$ 、 $\mathbf{x}_i$ 分别为第*t*时间步的预测值、隐藏层状态向量和输入; $\sigma_1(\cdot)$ 、 $\sigma_2(\cdot)$ 为激活函数; $\mathbf{V}$ 、 $\mathbf{U}$ 、 $\mathbf{W}$ 分别为不同全连接层的权重矩阵。目前,RNN主要用于处理序列建模问题,被广泛应用于语音识别、机器翻译等领域。

然而,在处理长序列时,RNN在进行反向传播的过程中容易出现梯度消失现象。不同于普通多层神经网络或CNN,RNN的梯度消失现象主要体现在反向传播梯度被近距离的梯度主导,使得RNN难以解决长期依赖性问题。在反向传播过程中,对式(1)中的 $\mathbf{W}$ 和 $\mathbf{U}$ 进行权重更新时,将会出现时刻*k*到时刻*t*反向传播过程中产生的连乘项:

$$\prod_{i=k+1}^t \left( \frac{\partial \mathbf{h}_i}{\partial \mathbf{h}_{i-1}} \right) = \frac{\partial \mathbf{h}_t}{\partial \mathbf{h}_{t-1}} \frac{\partial \mathbf{h}_{t-1}}{\partial \mathbf{h}_{t-2}} \cdots \frac{\partial \mathbf{h}_{k+1}}{\partial \mathbf{h}_k} = \prod_{i=k+1}^t \left\{ \mathbf{W}^T \text{diag} \left[ \sigma_2'(\mathbf{W}\mathbf{h}_{i-1} + \mathbf{U}\mathbf{x}_i) \right] \right\}, \quad (2)$$

式中: $\text{diag}(\cdot)$ 为提取矩阵对角线元素操作; $\sigma_2'(\cdot)$ 为非线性激活函数。由于RNN常用的标准非线性激活函数容易得到小于1的梯度矩阵,随着序列长度的增加,连乘项的项数随之增多,距离较远的输入对梯度的影响会快速收敛至0。

为了缓解上述现象,LSTM中加入3种门控机制以及一种记忆单元,能够更有效地捕获长序列中的上下文信息。相比于传统RNN,LSTM在反向传播过程中的连乘项由遗忘门主导,因此可以在一定程度上缓解梯度消失现象。然而,当序列过长时,遗忘门无法保证恒为单位矩阵 $\mathbf{I}$ ,LSTM的梯度仍然会出现被近距离的梯度主导的情况。

高光谱图像的光谱维度通常存在数百个波段,直接使用每一个波段作为一个时间步的输入将会导致LSTM的网络过深而难以训练。Xu等<sup>[12]</sup>提出使用光谱波段分组的方式在光谱维度进行切分,使用切分后的光谱波段组作为LSTM各时间步的输入进而缩短序列长度。实验证明,相比前者,使用分组的方式能够更好地取得更好的分类效果。

### 2.2 相关研究

为了获取多尺度的光谱上下文信息,一般采取感受野多尺度与分组多尺度两种策略。Wang等<sup>[13]</sup>提出一种光谱多尺度特征提取网络(SeMN),对不同感受

野尺度下的光谱特征向量进行切分,以此缩短序列长度,在缓解梯度消失问题的同时实现分组多尺度策略。然而,当分组的组数增加时,输入 LSTM 的序列随之变长,其分类准确率与稳定性均明显降低。Xu 等<sup>[12]</sup>同样指出,在使用光谱波段分组策略时,随着组数的增加,将 LSTM 最后的隐藏层状态向量作为特征无法取得令人满意的分类精度。

针对上述问题,引入注意力机制网络能够起到一定效果。注意力机制网络使 RNN 能够再次回顾网络中不同时刻的隐藏层状态,在处理信息瓶颈问题的同时,进一步缓解 RNN(包括 LSTM)梯度消失的问题。Bahdanau 等<sup>[16]</sup>提出一种注意力机制网络,使 Seq2Seq 模型中的译码器能够选择性地回顾编码器中的信息,获取具有更多相关信息的语义编码向量。此外,在文本分类领域,Yang 等<sup>[17]</sup>提出一种注意力机制网络来处理多对一问题,具体公式为

$$\begin{cases} s = \sum_i \alpha_i h_i \\ \alpha_i = \exp(u_i^T u_w) / \sum_i [\exp(u_i^T u_w)], \\ u_i = \tanh(W_w h_i + b_w) \end{cases} \quad (3)$$

式中: $W_w$ 和 $b_w$ 分别为全连接层的权重矩阵和偏置; $u_w$ 为一个随机初始化的上下文信息向量; $u_i$ 为提取的语义特征; $\tanh(\cdot)$ 为 $\tanh$ 激活函数。通过隐藏层状态向量 $h_i$ 与上下文信息向量 $u_w$ 之间的相关性计算权重 $\alpha_i$ ,利用权重对隐藏层状态向量进行加权求和。相比直接

使用 LSTM 最后时刻的隐藏层状态向量 $h_n$ ,这种方式能够得到包含更多相关信息的输出 $s$ 。

注意力机制网络同样可以用于处理空间信息。为实现空间多尺度策略,常见的方法是采用多次卷积的方式,在不同卷积层的输出中得到不同感受野尺度下的空间特征。然而,随着卷积网络的加深,无法保证不同感受野下的分支网络能关注到原始图像中的重点信息。注意力机制能够很自然地加入卷积层之间,将网络聚焦到空间中的重要部分。Woo 等<sup>[18]</sup>提出一种空间注意力机制网络,如图 1 所示。为计算空间中不同像素的权重,他们采用平均池化与最大池化的方式进行信息的聚合,并使用卷积的方式得到空间上的权重 $\alpha$ ,具体计算方法为

$$\begin{cases} Y = X + X' \\ X' = \alpha \circ X \\ \alpha = \sigma[f_{m \times m}(e)] \\ e = \text{Concat}(e_1, e_2) \\ e_1 = \text{Avgpool}(X) \\ e_2 = \text{Maxpool}(X) \end{cases} \quad (4)$$

式中: $Y$ 为空间注意力机制网络的输出; $X$ 为高光谱图像经过主成分分析(PCA)降维后的局部邻域; $\circ$ 表示哈达玛积(hadamard product); $\sigma(\cdot)$ 为 Sigmoid 激活函数; $f_{m \times m}(\cdot)$ 表示应用卷积核大小为 $m \times m$ 的二维卷积层; $\text{Concat}(w, v)$ 表示对 $w$ 和 $v$ 进行拼接; $\text{Avgpool}(\cdot)$ 和 $\text{Maxpool}(\cdot)$ 分别表示通道维度上的平均池化和最大池化操作。

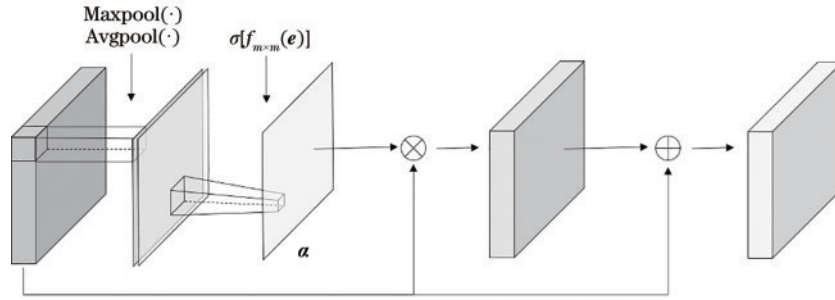


图 1 空间注意力机制网络结构图

Fig. 1 Structure of spatial attention network

### 2.3 网络框架

引入注意力机制的光谱多尺度特征提取网络(SeAMN)具体结构如图 2 所示。SeAMN 利用 3 组一维卷积层和最大池化层来实现感受野多尺度策略。此外,为实现分组多尺度策略,分别对原始光谱向量以及不同感受野下的特征向量进行分组,采用递减的方式来设置组数,感受野越大,组数越小。在不同尺度下,利用一个 LSTM 进行进一步的光谱特征提取。最后,将不同尺度下的光谱特征相加,得到多尺度光谱特征。为充分发挥多尺度分组策略的作用,SeAMN 在不同尺度下的 LSTM 输出处引入一种光谱注意力机制网

络,以此缓解梯度消失现象。

参考图 1 的注意力机制网络结构,所提光谱注意力机制网络如图 3 所示,其主要利用两个全连接层来处理不同光谱波段组在 LSTM 中的隐藏层状态向量进而生成权重,并计算加权求和后的信息 $u$ ,具体公式为

$$\begin{cases} u = O \cdot \alpha \\ \alpha = \text{Softmax}(W_2 \cdot e + b_2), \\ e = \tanh(W_1 \cdot O + b_1) \end{cases} \quad (5)$$

式中: $O = [h_1, h_2, \dots, h_n]$ 为 LSTM 的隐藏层状态矩

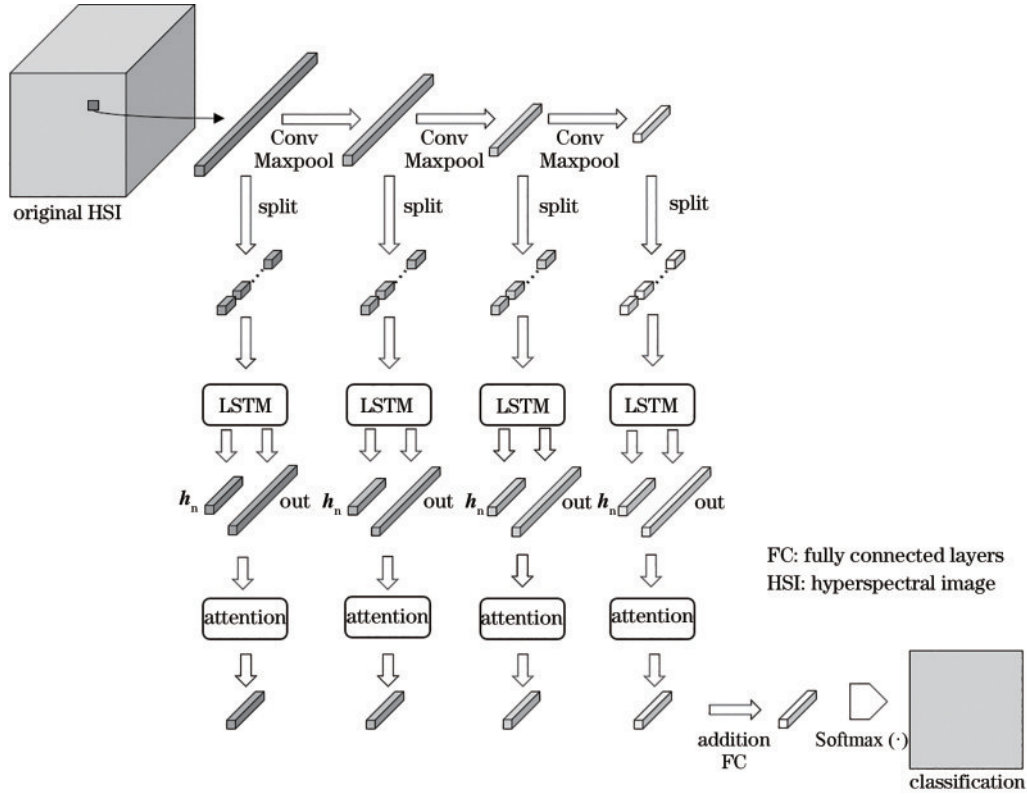


图 2 SeAMN 结构图  
Fig. 2 Structure of SeAMN

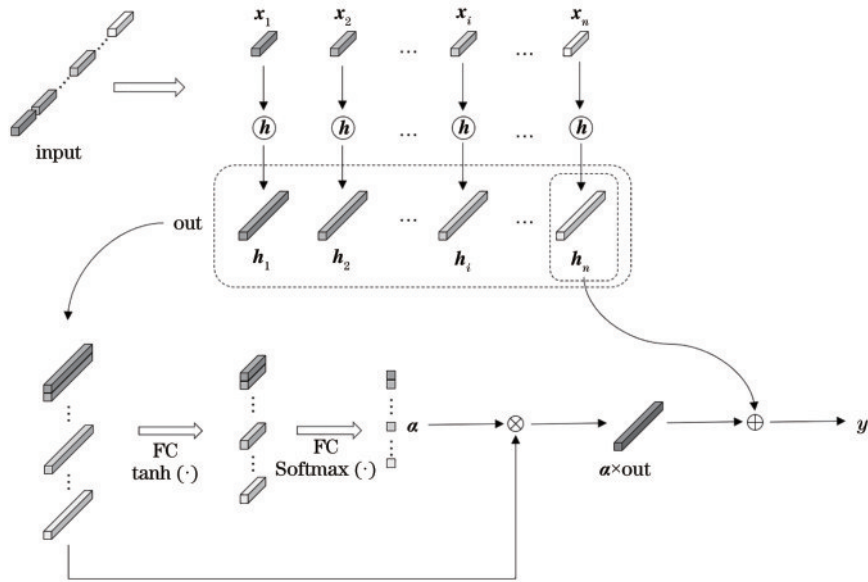


图 3 光谱注意力机制网络结构图  
Fig. 3 Structure of spectral attention network

阵;  $\text{Softmax}(\cdot)$  为 Softmax 激活函数;  $W_1$  和  $W_2$  分别为第一个和第二个全连接层的权重矩阵, 而  $b_1$  和  $b_2$  则分别表示其偏置。首先使  $O$  经过一层全连接神经网络, 并使用  $\tanh(\cdot)$  作为激活层, 得到新的隐藏层表达  $e$ ; 然后使用第二层全连接层来衡量每一组光谱波段组隐藏层状态向量的重要性, 并使用  $\text{Softmax}(\cdot)$  将其映射为概率分布; 再采用软注意力机制, 利用权重  $\alpha$  对  $O$  进

行加权求和; 最后, 将  $h_n$  与  $u$  的和作为光谱注意力机制网络的输出:

$$y = h_n + u. \tag{6}$$

通常, 注意力机制网络将  $u$  作为网络的输出。然而, 在 LSTM 中,  $h_n$  捕获了大部分光谱信息, 因此, 与其他时刻的输出相比,  $h_n$  更重要。所提方法从残差网络的结构中得到启发, 选择使用  $y$  作为输出, 使光谱注

注意力机制网络的学习目标发生变化。假设光谱信息的真实分布为  $H(x)$ , 当使用  $u$  作为输出时, 注意力机制网络的目标在于使输出  $u$  尽量拟合光谱信息的真实分布  $H(x)$ 。然而, 当使用  $y$  作为输出时, 注意力机制网络的目标在于使输出  $u$  尽量拟合  $H(x) - h_n$ , 即使用注意力机制网络从 LSTM 的数量为  $n$  的输出中收集丢失的信息对  $h_n$  进行补充, 这种方式保证了  $h_n$  的重要性。

引入注意力机制的空间多尺度特征提取网络

(SaAMN) 如图 4 所示。首先使用 PCA 对原始高光谱图像进行降维。随后, 使用 3 组卷积-注意力模块处理像素的局部邻域, 以此来实现多尺度的空间特征提取。在不同尺度下, 需要将卷积-注意力模块得到的空间特征在空间维度上切分成局部块, 以方便后续处理。对于一个尺寸为  $H \times W$  的邻域信息  $X$ , 将其切分为  $s \times s$  的局部块  $Z_{ij}$  (其中:  $i < H/s; j < W/s$ ), 其计算公式为

$$Z_{ij} = \begin{bmatrix} X_{ij} & X_{i(j+w/s)} & \cdots & X_{i[j+w/s \times (s-1)]} \\ X_{(i+H/s)j} & X_{(i+H/s)(j+w/s)} & \cdots & X_{(i+H/s)[j+w/s \times (s-1)]} \\ \vdots & \vdots & \ddots & \vdots \\ X_{[i+H/s \times (s-1)]j} & X_{[i+H/s \times (s-1)](j+w/s)} & \cdots & X_{[i+H/s \times (s-1)][j+w/s \times (s-1)]} \end{bmatrix} \quad (7)$$

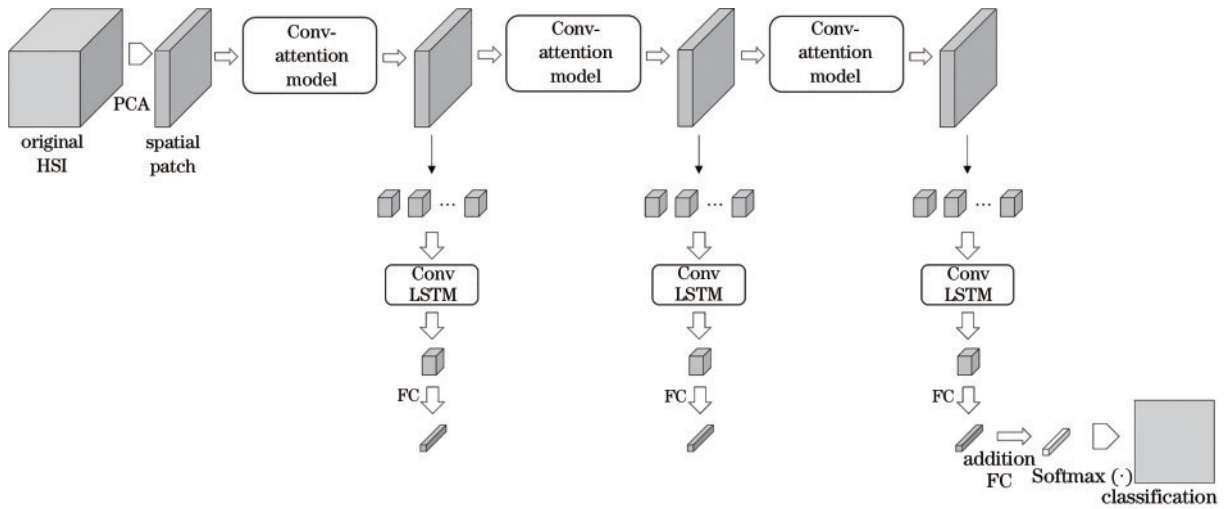


图 4 SaAMN 结构图

Fig. 4 Structure of SaAMN

不同尺度下的切分结果作为不同 ConvLSTM 的输入, 进一步提取空间邻域中的上下文信息。最后, 对多种尺度下的结果求和, 获取多尺度空间特征。其中, ConvLSTM 是 LSTM 的一种改进方案, 是在 LSTM 中引入卷积算子, 以保留邻域像素之间的空间关系, 具体计算公式为

$$\begin{cases} f_t = \sigma(W_{hf} * H_{t-1} + W_{xf} * X_t + W_{cf} \circ C_{t-1} + b_f) \\ i_t = \sigma(W_{hi} * H_{t-1} + W_{xi} * X_t + W_{ci} \circ C_{t-1} + b_i) \\ o_t = \sigma(W_{ho} * H_{t-1} + W_{xo} * X_t + W_{co} \circ C_{t-1} + b_o) \\ H_t = o_t \circ \tanh(C_t) \\ C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_t \\ \tilde{C}_t = \tanh(W_{hc} * H_{t-1} + W_{xc} * X_t + b_c) \end{cases}, \quad (8)$$

式中:  $f_t, i_t, o_t$  分别为遗忘门、输入门、输出门;  $X_t, H_t, C_t$  分别为第  $t$  时间步的输入、隐藏层状态和记忆单元;  $[(W_{\odot i}, b_i), (W_{\odot f}, b_f), (W_{\odot o}, b_o), \odot = x, h, c]$  为权重与偏置矩阵;  $*$  表示卷积操作。

上述卷积-注意力模块由两个空洞卷积层和一个空间注意力机制网络构成, 如图 5 所示。首先通过两层空洞卷积层, 在图像尺寸保持不变的情况下增加感受野, 进而实现多尺度特征提取; 然后, 利用式(4)计算空间注意力图并进行加权处理, 使网络关注到空间上的重要信息。通过在网络主干部分引入空间注意力机制网络, 能够有效帮助主干上的卷积网络突出空间中的关键信息。空间注意力机制网络的输出结果将作为当前尺度下分支网络中 ConvLSTM 的输入, 有利于 ConvLSTM 进行空间维度的特征提取。

为进行空-谱特征的联合提取, 采取一种得分融合策略对上文中两部分网络进行融合, 如图 6 所示。该策略能够将光谱与空间网络合并为一个整体, 同时进行训练与融合, 得到空-谱联合特征, 具体计算方法为

$$S = \sigma(m) S_{se} + [1 - \sigma(m)] S_{sa} = \sigma(m) W_{se} F_{se} + [1 - \sigma(m)] W_{sa} F_{sa} + \sigma(m) b_{se} + [1 - \sigma(m)] b_{sa}, \quad (9)$$

式中:  $F_{se}$  和  $F_{sa}$  分别为提取到的多尺度光谱特征和多尺度空间特征;  $m$  为一个可通过神经网络优化的权重

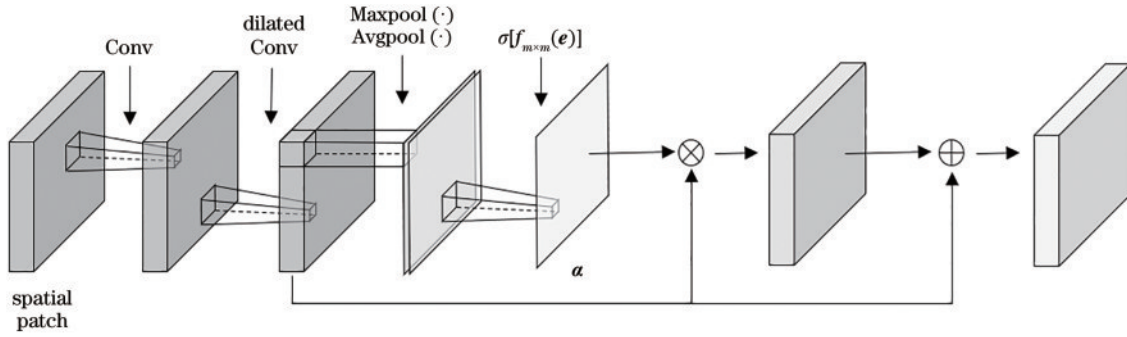


图 5 卷积-注意力模块网络结构图  
Fig. 5 Structure of Conv-attention model

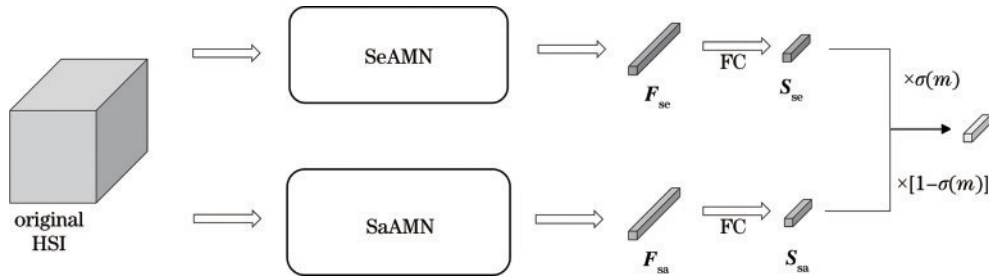


图 6 引入注意力机制的多尺度特征提取网络框架  
Fig. 6 Structure of multiscale feature extraction method with attention mechanism

参数;  $W_{se}$ 、 $b_{se}$ 、 $W_{sa}$ 和 $b_{sa}$ 分别为用于计算光谱与空间得分的权重和偏置。首先,利用两个全连接层分别处理  $F_{se}$ 和 $F_{sa}$ ,进而得到光谱得分向量  $S_{se}$ 与空间得分向量  $S_{sa}$ 。随后,通过一个可迭代的参数  $m$ 作为权重值,将得到的光谱得分向量和空间得分向量进行加权求和,得到最终的得分向量  $S$ 。

引入注意力机制的多尺度高光谱图像特征提取方法包括光谱和空间两个部分。SeAMN考虑到梯度消失现象对分组策略的影响,通过光谱注意力机制网络,利用LSTM每一时间步的隐藏层状态向量来补充丢失的光谱信息,进而缓解光谱波段分组策略中的梯度消失现象带来的分组组数限制,充分发挥分组策略的作用。在SaAMN中使用一种卷积-注意力模块,将空间注意力机制网络自然地加入到多尺度卷积层之间,能够有效地帮助主干上的卷积网络突出空间中的关键信息,有利于ConvLSTM进行空间维度的特征提取。最后,使用一种得分融合策略有效地将两部分网络进行融合并提取空-谱联合特征,提高图像分类的准确度。

### 3 实验结果与分析

#### 3.1 数据集

为检验所提方法的有效性,在3个高光谱公开数据集上进行了实验,数据集分别为Pavia University、

KSC(Kennedy Space Center)以及Indian Pines。

Pavia University数据集由ROSIS(reflective optics spectrographic imaging system)传感器(ROSIS-03)对意大利的帕维亚大学成像获得,该数据集包含9类地物,尺寸为610 pixel × 340 pixel、空间分辨率为1.3 m/pixel。数据集包含115个波段,去除其中受到噪声影响的12个波段后,使用剩余103个波段进行实验。

KSC数据集由机载可见/红外成像光谱仪(AVIRIS)于佛罗里达州肯尼迪航天中心采集得到,该数据集包含13类地物,尺寸为512 pixel × 614 pixel、空间分辨率为18 m/pixel。数据集包含224个波段,在预处理中去除48个波段后,保留剩余176个波段作为实验数据。

Indian Pines数据集是由AVIRIS对位于美国印第安纳州西北部的印第安试验场进行成像捕获到的植被影像。该数据集包含16个类别,尺寸为145 pixel × 145 pixel,由220个波段组成,去除20个受到影响的波段后,保留剩下的200个波段作为实验数据。

在本实验中,为了对网络进行训练,从上述3种数据集的标记样本中随机抽取一定量的样本作为训练样本,其他样本则作为测试样本,3种数据集的训练样本设置如表1、表2、表3所示。

表1 Pavia University 数据集类别及训练样本设置

Table 1 Categories and settings of Pavia University dataset

| Class | Class Name           | Training number | Total number |
|-------|----------------------|-----------------|--------------|
| 1     | Asphalt              | 100             | 6631         |
| 2     | Meadows              | 100             | 18649        |
| 3     | Gravel               | 100             | 2099         |
| 4     | Trees                | 100             | 3064         |
| 5     | Painted metal sheets | 100             | 1345         |
| 6     | Bare soil            | 100             | 5029         |
| 7     | Bitumen              | 100             | 1330         |
| 8     | Self-blocking bricks | 100             | 3682         |
| 9     | Shadows              | 100             | 947          |
| Total |                      | 900             | 42776        |

表2 KSC 数据集类别及训练样本设置

Table 2 Categories and settings of KSC dataset

| Class | Class Name      | Training number | Total number |
|-------|-----------------|-----------------|--------------|
| 1     | Scrub           | 33              | 761          |
| 2     | Willow swamp    | 23              | 243          |
| 3     | CP hammock      | 24              | 256          |
| 4     | CP/Oak          | 24              | 252          |
| 5     | Slash pine      | 15              | 161          |
| 6     | Oak/Broadleaf   | 22              | 229          |
| 7     | Hardwood swamp  | 9               | 105          |
| 8     | Graminoid marsh | 38              | 431          |
| 9     | Spartina marsh  | 51              | 520          |
| 10    | Catiai marsh    | 39              | 404          |
| 11    | Salt marsh      | 41              | 419          |
| 12    | Mud flats       | 49              | 503          |
| 13    | Water           | 91              | 927          |
| Total |                 | 459             | 5211         |

表3 Indian Pines 数据集类别及训练样本设置

Table 3 Categories and settings of Indian Pines dataset

| Class | Class Name                   | Training number | Total number |
|-------|------------------------------|-----------------|--------------|
| 1     | Alfalfa                      | 33              | 46           |
| 2     | Corn-notill                  | 100             | 1428         |
| 3     | Corn-mintill                 | 100             | 830          |
| 4     | Corn                         | 100             | 237          |
| 5     | Grass-pasture                | 100             | 483          |
| 6     | Grass-trees                  | 100             | 730          |
| 7     | Grass-pasture-mowed          | 20              | 28           |
| 8     | Hay-windrowed                | 100             | 478          |
| 9     | Oats                         | 14              | 20           |
| 10    | Soybean-notill               | 100             | 972          |
| 11    | Soybean-mintill              | 100             | 2455         |
| 12    | Soybean-clean                | 100             | 593          |
| 13    | Wheat                        | 100             | 205          |
| 14    | Woods                        | 100             | 1265         |
| 15    | Buildings-Grass-Trees-Drives | 100             | 386          |
| 16    | Stone-Steel-Towers           | 75              | 93           |
| Total |                              | 1342            | 10249        |

### 3.2 实验设置

为验证所提方法的有效性,通过实验分别对比不同的

光谱特征提取、空间特征提取以及空-谱联合特征提取方法。在光谱特征提取方法中,选择3种具有代表性的RNN高光谱特征提取方法进行对比实验,分别为以单一光谱波段为输入的LSTM方法(LSTM-byb)<sup>[12]</sup>、采用光谱波段分组策略的LSTM方法(LSTM-split)<sup>[12]</sup>、ASSMN中提出的多尺度光谱特征提取方法(SeMN)<sup>[13]</sup>。此外,为对比RNN以外的深度学习特征提取方法,实验中还对比了基于一维CNN的光谱特征提取方法(1DCNN)<sup>[19]</sup>。在空间特征提取方法中,将ASSMN中提出的多尺度空间特征提取方法(SaMN)<sup>[13]</sup>以及基于二维CNN的空间特征提取方法(2DCNN)<sup>[20]</sup>作为对比方法。在空-谱联合特征提取方法中,将ASSMN<sup>[13]</sup>与光谱-空间统一网络(SSUN)<sup>[12]</sup>两种联合特征提取方法作为对比方法。使用总体准确度(OA)、平均准确度(AA)、卡帕系数(KC)以及训练时长(TrnTime)4种指标对所有方法的分类性能进行评价。

实验中,所提方法需要调节的超参数如表4所示。需要调节的超参数为:SeAMN中的光谱波段分组组数、SeAMN的隐藏层维数以及SaAMN的卷积核大小。光谱波段的分组组数直接影响到不同类别分类的准确性,应当依据实验数据集的光谱维数进行设置。经过实验,分组组数依据卷积网络的从浅到深依次减半,分别设置为:Indian Pines数据集中8-4-2-1;Pavia University数据集中8-4-2-1;KSC数据集中16-8-4-2。SeAMN中不同隐藏层大小与SaAMN不同卷积核大小的分类效果对比结果如图7(a)(b)所示,图7(c)为不同分组方式的分类效果。实验证明,当SeAMN中的隐藏层大小在75左右时,网络可以达到理想的性能,因此,在3组数据集中,均将其设置为75。此外,当SaAMN中的卷积核大小达到 $5 \times 5$ 后,网络性能将出现明显的下降,在3种数据集中该参数大小均设为 $3 \times 3$ 。在将SeAMN和SaAMN进行联合提取时,除在KSC数据集中的光谱波段分组组数修改为32-16-8-4外,其余参数均与上述参数一致。

表4 所提方法的参数设置

Table 4 Settings of proposed method

| Method          | Dataset          | Split     | Hidden size | Kernel size |
|-----------------|------------------|-----------|-------------|-------------|
| SeAMN           | Indian Pines     | 8-4-2-1   | 75          | —           |
|                 | Pavia University | 8-4-2-1   | 75          | —           |
|                 | KSC              | 16-8-4-2  | 75          | —           |
| SaAMN           | Indian Pines     | —         | —           | 3           |
|                 | Pavia University | —         | —           | 3           |
|                 | KSC              | —         | —           | 3           |
| Proposed method | Indian Pines     | 8-4-2-1   | 75          | 3           |
|                 | Pavia University | 8-4-2-1   | 75          | 3           |
|                 | KSC              | 32-16-8-4 | 75          | 3           |

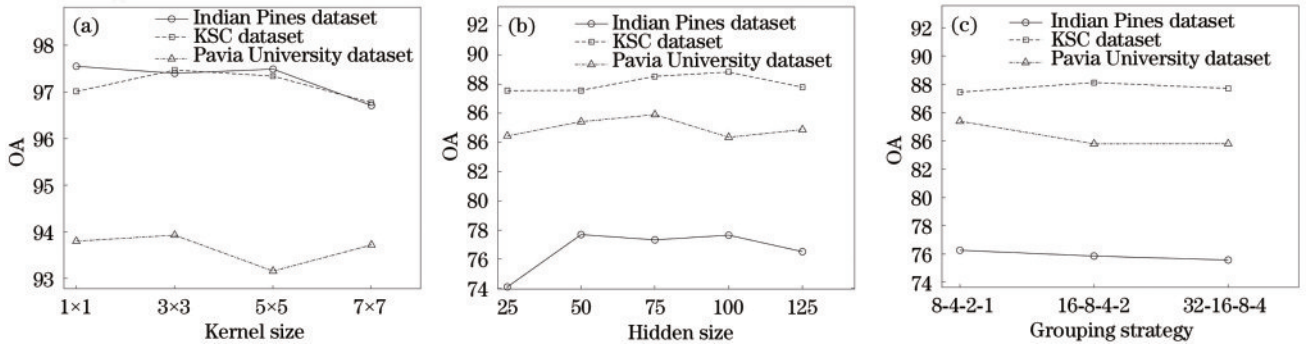


图 7 在不同超参数下,独立运行 10 次实验的网络总体准确度。(a) SaAMN 在不同卷积核大小下的表现;(b) SeAMN 在不同隐藏层大小下的表现;(c) SeAMN 在不同分组策略下的表现

Fig. 7 Overall accuracy of proposed method by running 10 experiments independently with different hyperparameters. (a) Performance of SaAMN with different kernel sizes; (b) performance of SeAMN with different hidden sizes; (c) performance of SeAMN with different grouping strategies

### 3.3 实验结果

表 5 给出了 5 种光谱特征提取方法、3 种空间特征提取方法以及 3 种空-谱联合特征提取方法在 KSC 数据集上经过 10 次独立运行后的平均分类精度 (AP) 与其对应的标准差 (SD)。表 5 中, LSTM-byb 表现最差, 这是由于输入序列的长度为高光谱图像的光谱维数,

过长的输入序列导致网络出现梯度消失现象, 难以准确地通过得到的光谱特征对高光谱图像进行分类。相比 LSTM-byb, LSTM-split 使用了光谱波段分组的手段, 有效缓解了梯度消失现象, 因此其分类准确率得到了明显提升。与前两种方法相比, SeMN 引入多尺度思想, 提取不同尺度下的光谱上下文信息, 进一步提高

表 5 不同特征提取方法在 KSC 数据集上的分类结果

Table 5 Classification results of different feature extraction methods on KSC dataset

| Method                      | Spectral feature extraction method |              |             |       |       | Spatial feature extraction method |              |        | Joint feature extraction method |              |        | Proposed method |
|-----------------------------|------------------------------------|--------------|-------------|-------|-------|-----------------------------------|--------------|--------|---------------------------------|--------------|--------|-----------------|
|                             | 1DCNN                              | LSTM -byb    | LSTM -split | SeMN  | SeAMN | 2DCNN                             | SaMN         | SaAMN  | SSUN                            | ASSMN        |        |                 |
| OA / %                      | AP                                 | 85.81        | 68.11       | 73.00 | 88.03 | <b>88.55</b>                      | 96.58        | 96.80  | <b>97.47</b>                    | 97.43        | 98.01  | <b>98.46</b>    |
|                             | SD                                 | ±1.26        | ±1.90       | ±1.64 | ±1.06 | <b>±0.82</b>                      | ±0.51        | ±1.54  | <b>±1.65</b>                    | ±0.65        | ±0.77  | <b>±0.50</b>    |
| AA / %                      | AP                                 | 78.90        | 55.56       | 61.50 | 82.71 | <b>83.23</b>                      | 95.85        | 95.91  | <b>96.94</b>                    | 96.84        | 97.32  | <b>98.00</b>    |
|                             | SD                                 | ±2.49        | ±3.12       | ±1.89 | ±1.36 | <b>±1.40</b>                      | ±0.73        | ±1.98  | <b>±2.21</b>                    | ±0.61        | ±1.15  | <b>±0.80</b>    |
| KC × 100%                   | AP                                 | 84.21        | 64.40       | 69.93 | 86.68 | <b>87.25</b>                      | 96.19        | 96.44  | <b>97.18</b>                    | 97.14        | 97.78  | <b>98.29</b>    |
|                             | SD                                 | ±1.40        | ±2.18       | ±1.83 | ±1.18 | <b>±0.91</b>                      | ±0.56        | ±1.71  | <b>±1.84</b>                    | ±0.72        | ±0.86  | <b>±0.56</b>    |
| TrnTime / s                 | AP                                 | <b>7.14</b>  | 46.84       | 10.93 | 30.02 | 50.54                             | <b>13.28</b> | 225.88 | 243.11                          | <b>96.46</b> | 246.37 | 313.15          |
|                             | SD                                 | <b>±0.66</b> | ±0.43       | ±0.22 | ±0.88 | ±1.10                             | <b>±0.32</b> | ±9.46  | ±6.01                           | <b>±4.63</b> | ±5.83  | ±1.20           |
| Classification accuracy / % | 1                                  | 87.43        | 74.77       | 75.87 | 85.96 | 88.13                             | 92.12        | 94.16  | 94.11                           | 94.74        | 96.35  | 96.91           |
|                             | 2                                  | 87.64        | 69.18       | 70.59 | 87.82 | 84.82                             | 94.55        | 93.64  | 96.73                           | 96.86        | 96.82  | 98.41           |
|                             | 3                                  | 75.65        | 32.89       | 52.16 | 86.68 | 87.54                             | 96.64        | 94.35  | 94.22                           | 97.76        | 96.85  | 97.97           |
|                             | 4                                  | 48.68        | 20.83       | 39.34 | 61.54 | 64.96                             | 88.42        | 89.61  | 91.14                           | 90.39        | 91.45  | 94.04           |
|                             | 5                                  | 50.82        | 40.27       | 41.23 | 60.07 | 55.34                             | 99.04        | 96.30  | 97.12                           | 92.74        | 97.05  | 97.26           |
|                             | 6                                  | 49.23        | 14.49       | 36.81 | 56.81 | 55.99                             | 88.84        | 95.31  | 96.43                           | 93.14        | 93.19  | 94.59           |
|                             | 7                                  | 68.02        | 13.13       | 2.81  | 71.98 | 78.96                             | 93.23        | 92.50  | 95.52                           | 99.48        | 96.56  | 97.60           |
|                             | 8                                  | 84.86        | 32.65       | 40.97 | 84.35 | 86.39                             | 98.19        | 92.85  | 96.26                           | 96.31        | 98.04  | 98.02           |
|                             | 9                                  | 92.28        | 82.43       | 85.61 | 94.56 | 95.03                             | 99.04        | 99.72  | 99.87                           | 99.85        | 99.30  | 99.62           |
|                             | 10                                 | 94.03        | 69.34       | 77.18 | 96.47 | 94.93                             | 98.33        | 99.75  | 100.00                          | 99.42        | 99.78  | 99.92           |
|                             | 11                                 | 96.64        | 92.35       | 94.42 | 96.03 | 96.80                             | 99.87        | 100.00 | 100.00                          | 98.60        | 99.81  | 99.74           |
|                             | 12                                 | 90.48        | 80.29       | 82.69 | 93.04 | 93.06                             | 97.82        | 98.63  | 98.83                           | 99.63        | 100.00 | 99.93           |
|                             | 13                                 | 99.89        | 99.68       | 99.89 | 99.99 | 99.98                             | 100.00       | 100.00 | 100.00                          | 100.00       | 100.00 | 100.00          |

Note: "1" to "13" are categories of KCS dataset, OA, AA and KC are calculated from them.



了分类性能。在基于 RNN 的光谱特征提取方法中, SeAMN 引入注意力机制, 缓解梯度消失现象的同时, 充分发挥了光谱波段分组的多尺度策略, 取得了较高的分类准确率。在空间注意力网络中, SaMN 引入多尺度思想, 在不同尺度下提取空间上下文信息, 相比 2DCNN, SaMN 具有更优的分类性能。SaAMN 在卷积层之间引入空间注意力机制网络, 使网络在学习的过程中能够尽量保留邻域中的重要信息, 进而得到了比 SaMN 更好的分类效果。在空-谱联合特征提取方法中, 所提方法同样得到了具有竞争力的分类效果。

在 Pavia University 数据集中, 11 种特征提取方法经过 10 次独立运行后的 AP 与其对应的 SD 如表 6 所示。相较于 KSC 数据集, Pavia University 数据集的光谱波段长度相对较短, 在该数据集中, LSTM-split 并没有发挥出明显的优势, 然而, 由于序列较短, 梯度消

失的现象得到进一步缓解, 相较于 1DCNN, SeAMN 的分类准确率存在明显提升。此外, 在 3 种空间特征提取网络的对比中, SaAMN 能够达到了较高的分类准确率, 同时相比于 SaMN, 其具有更低的标准差。相较于 ASSMN, 所提方法达到了较高的总体准确率, 但某些特定类别地物的分类精度较低, 具体集中于 Asphalt、Gravel、Bare soil、Self-blocking bricks 4 类地物上。通过观察这 4 种地物的光谱曲线, 不难看出它们具有相似的光谱特性, 即通常都在可见光范围内表现为较低的反光率, 而在近红外范围内表现为较高的反光率。与 ASSMN 不同, 所提方法引入注意力机制网络后使得分组策略更加自由, 使用不同的分组策略可以得到想要的结果。然而, 不同的分组策略可能并不适用于区分一些特定类别的地物, 导致这些地物在通过 LSTM 提取特征时出现较高的相似度。

表 6 不同特征提取方法在 Pavia University 数据集上的分类结果  
Table 6 Classification results of different feature extraction methods on Pavia University dataset

| Method                      | Spectral feature extraction method |              |             |       |       | Spatial feature extraction method |              |        | Joint feature extraction method |               |                 |              |
|-----------------------------|------------------------------------|--------------|-------------|-------|-------|-----------------------------------|--------------|--------|---------------------------------|---------------|-----------------|--------------|
|                             | 1DCNN                              | LSTM -byb    | LSTM -split | SeMN  | SeAMN | 2DCNN                             | SaMN         | SaAMN  | SSUN                            | ASSMN         | Proposed method |              |
| OA / %                      | AP                                 | 78.37        | 71.28       | 69.36 | 82.94 | <b>85.91</b>                      | 90.26        | 92.97  | <b>93.93</b>                    | 94.71         | 95.95           | <b>96.61</b> |
|                             | SD                                 | ±2.29        | ±6.20       | ±1.86 | ±2.47 | <b>±1.14</b>                      | ±1.10        | ±2.40  | <b>±1.26</b>                    | ±1.03         | ±1.38           | <b>±1.04</b> |
| AA / %                      | AP                                 | 86.75        | 80.85       | 81.34 | 87.72 | <b>89.49</b>                      | 93.68        | 95.61  | <b>96.02</b>                    | 96.95         | <b>97.77</b>    | 97.72        |
|                             | SD                                 | ±0.40        | ±3.85       | ±0.91 | ±1.07 | <b>±0.74</b>                      | ±0.66        | ±1.18  | <b>±0.73</b>                    | ±0.32         | <b>±0.48</b>    | ±0.51        |
| KC × 100%                   | AP                                 | 72.56        | 64.36       | 61.96 | 77.96 | <b>81.69</b>                      | 87.27        | 90.80  | <b>92.02</b>                    | 93.04         | 94.66           | <b>95.51</b> |
|                             | SD                                 | ±2.54        | ±6.93       | ±1.82 | ±3.06 | <b>±1.45</b>                      | ±1.40        | ±3.08  | <b>±1.62</b>                    | ±1.33         | ±1.80           | <b>±1.35</b> |
| TrnTime / s                 | AP                                 | <b>12.89</b> | 49.43       | 21.52 | 61.85 | 78.80                             | <b>25.90</b> | 387.25 | 416.77                          | <b>178.83</b> | 442.47          | 569.15       |
|                             | SD                                 | <b>±0.73</b> | ±0.42       | ±0.50 | ±0.28 | ±0.59                             | <b>±0.32</b> | ±4.32  | ±15.24                          | <b>±11.70</b> | ±23.50          | ±2.32        |
| Classification accuracy / % | 1                                  | 77.47        | 70.28       | 71.61 | 82.34 | 82.76                             | 88.31        | 90.90  | 92.65                           | 93.66         | 95.95           | 95.32        |
|                             | 2                                  | 70.94        | 62.28       | 57.89 | 79.94 | 84.22                             | 87.47        | 90.44  | 92.10                           | 92.41         | 94.15           | 96.10        |
|                             | 3                                  | 82.03        | 59.69       | 70.73 | 79.34 | 81.60                             | 88.29        | 89.47  | 90.50                           | 95.57         | 97.19           | 95.85        |
|                             | 4                                  | 94.43        | 93.44       | 95.76 | 93.49 | 94.53                             | 94.84        | 97.20  | 97.82                           | 99.42         | 99.17           | 99.16        |
|                             | 5                                  | 99.66        | 97.87       | 98.80 | 99.15 | 99.61                             | 99.94        | 99.97  | 99.98                           | 99.82         | 99.99           | 99.98        |
|                             | 6                                  | 79.84        | 78.60       | 70.81 | 81.88 | 87.19                             | 90.34        | 97.09  | 96.25                           | 96.56         | 96.74           | 96.52        |
|                             | 7                                  | 92.71        | 90.37       | 91.59 | 91.93 | 91.74                             | 97.27        | 99.11  | 99.01                           | 97.76         | 99.48           | 99.45        |
|                             | 8                                  | 83.67        | 75.24       | 74.94 | 81.52 | 83.90                             | 97.57        | 96.43  | 95.98                           | 97.59         | 97.29           | 97.08        |
|                             | 9                                  | 100.00       | 99.85       | 99.95 | 99.89 | 99.83                             | 99.10        | 99.87  | 99.89                           | 99.80         | 100.00          | 99.99        |

Note: "1" to "9" are categories of Pavia University dataset, OA, AA and KC are calculated from them.

在 Indian Pines 数据集中, 11 种特征提取方法经过 10 次独立运行后的 AP 与其对应的 SD 如表 7 所示。Indian Pines 数据集具有较长的光谱维度, 这给基于 RNN 的光谱特征提取方法带来了较为不利的影响。由于 1DCNN 不受光谱维数的影响, 相较于 RNN 方法, 其分类准确率更高。由表 7 可以看出, 相较于其他

RNN 方法, 所提 SeAMN 仍然能够有效缓解梯度消失现象, 提取更多有利于分类的光谱信息。在该数据集下, 所提方法依然能够有效提升分类准确率、降低标准差, 这是由于注意力机制能够有效减少关键信息的流失, 使得多次实验结果的差异相对较小。

表 7 不同特征提取方法在 Indian Pines 数据集上的分类结果  
Table 7 Classification results of different feature extraction methods on Indian Pines dataset

| Method                      | Spectral feature extraction method |              |                |       |       | Spatial feature extraction method |              |        | Joint feature extraction method |               |                    |              |
|-----------------------------|------------------------------------|--------------|----------------|-------|-------|-----------------------------------|--------------|--------|---------------------------------|---------------|--------------------|--------------|
|                             | 1DCNN                              | LSTM<br>-byb | LSTM<br>-split | SeMN  | SeAMN | 2DCNN                             | SaMN         | SaAMN  | SSUN                            | ASSMN         | Proposed<br>method |              |
| OA / %                      | AP                                 | <b>80.46</b> | 64.26          | 71.49 | 75.28 | 77.78                             | 95.87        | 97.36  | <b>97.71</b>                    | 96.28         | 96.79              | <b>98.13</b> |
|                             | SD                                 | <b>±1.41</b> | ±4.74          | ±1.54 | ±2.70 | ±2.40                             | ±0.72        | ±0.72  | <b>±0.38</b>                    | ±0.78         | ±0.72              | <b>±0.37</b> |
| AA / %                      | AP                                 | <b>87.44</b> | 72.31          | 79.06 | 81.97 | 84.16                             | 98.29        | 98.69  | <b>98.82</b>                    | 98.50         | 98.60              | <b>99.03</b> |
|                             | SD                                 | <b>±1.84</b> | ±4.99          | ±2.07 | ±3.07 | ±2.69                             | ±0.28        | ±0.40  | <b>±0.25</b>                    | ±0.01         | ±0.48              | <b>±0.29</b> |
| KC × 100%                   | AP                                 | <b>77.66</b> | 59.51          | 67.58 | 71.80 | 74.64                             | 95.22        | 96.94  | <b>97.34</b>                    | 95.69         | 96.28              | <b>97.84</b> |
|                             | SD                                 | <b>±1.57</b> | ±5.22          | ±1.70 | ±3.03 | ±2.68                             | ±0.83        | ±0.83  | <b>±0.44</b>                    | ±0.90         | ±0.83              | <b>±0.43</b> |
| TmTime / s                  | AP                                 | <b>20.34</b> | 203.95         | 32.92 | 79.25 | 110.56                            | <b>38.20</b> | 667.89 | 639.32                          | <b>273.34</b> | 626.90             | 745.69       |
|                             | SD                                 | <b>±0.68</b> | ±0.66          | ±0.63 | ±0.59 | ±0.40                             | <b>±0.33</b> | ±3.75  | ±38.31                          | <b>±4.69</b>  | ±4.94              | ±47.65       |
| Classification accuracy / % | 1                                  | 95.38        | 81.54          | 86.15 | 85.38 | 82.31                             | 100.00       | 100.00 | 99.23                           | 100.00        | 100.00             | 100.00       |
|                             | 2                                  | 78.52        | 57.06          | 63.31 | 69.72 | 70.45                             | 94.10        | 93.35  | 95.38                           | 93.52         | 93.35              | 96.31        |
|                             | 3                                  | 77.07        | 56.08          | 70.26 | 70.04 | 72.15                             | 97.51        | 98.21  | 97.90                           | 98.85         | 98.05              | 98.85        |
|                             | 4                                  | 90.58        | 65.55          | 84.01 | 87.23 | 84.89                             | 100.00       | 100.00 | 100.00                          | 99.85         | 99.78              | 99.64        |
|                             | 5                                  | 82.92        | 84.93          | 78.28 | 85.33 | 91.70                             | 99.03        | 98.36  | 98.96                           | 99.16         | 98.72              | 98.04        |
|                             | 6                                  | 93.32        | 80.21          | 89.10 | 88.51 | 95.35                             | 99.78        | 99.13  | 99.08                           | 99.65         | 99.86              | 99.86        |
|                             | 7                                  | 95.00        | 78.75          | 88.75 | 86.25 | 81.25                             | 100.00       | 100.00 | 100.00                          | 100.00        | 98.75              | 98.75        |
|                             | 8                                  | 97.99        | 92.78          | 96.08 | 95.87 | 97.25                             | 100.00       | 99.97  | 100.00                          | 100.00        | 100.00             | 100.00       |
|                             | 9                                  | 88.33        | 58.33          | 70.00 | 75.00 | 85.00                             | 100.00       | 100.00 | 100.00                          | 100.00        | 100.00             | 100.00       |
|                             | 10                                 | 86.16        | 68.30          | 83.13 | 80.36 | 83.57                             | 94.66        | 96.11  | 97.32                           | 95.03         | 97.21              | 97.80        |
|                             | 11                                 | 70.53        | 48.15          | 62.65 | 64.46 | 65.53                             | 91.77        | 97.24  | 97.01                           | 93.12         | 94.37              | 97.05        |
|                             | 12                                 | 87.61        | 60.12          | 69.94 | 77.63 | 79.90                             | 97.71        | 98.01  | 97.63                           | 99.27         | 98.88              | 98.80        |
|                             | 13                                 | 98.76        | 96.38          | 97.05 | 97.90 | 98.48                             | 100.00       | 100.00 | 99.90                           | 99.43         | 99.81              | 99.81        |
|                             | 14                                 | 81.18        | 82.24          | 70.23 | 83.16 | 86.31                             | 98.76        | 98.94  | 99.08                           | 98.11         | 98.75              | 99.56        |
|                             | 15                                 | 78.50        | 57.06          | 59.41 | 65.87 | 75.21                             | 99.97        | 99.76  | 99.58                           | 99.97         | 100.00             | 99.97        |
|                             | 16                                 | 97.22        | 89.44          | 96.67 | 98.89 | 97.22                             | 99.44        | 100.00 | 100.00                          | 100.00        | 100.00             | 100.00       |

Note: "1" to "16" are categories of Indian Pines dataset, OA, AA and KC are calculated from them.

## 4 结 论

提出了一种引入注意力机制的多尺度高光谱图像特征提取方法,其中包括SeAMN和SaAMN两个部分,通过一种得分融合策略进行融合。在SeAMN中,通过引入注意力机制能够缓解LSTM中的梯度消失现象,充分发挥多尺度分组策略的作用,更好地提取多尺度光谱特征。在SaAMN中,引入注意力机制能够帮助不同尺度下的分支网络获取关键信息,使网络关注到空间信息中的重要部分。实验结果表明,与其他几种特征提取方法相比,所提方法提取的光谱与空间特征能够稳定、有效地提升高光谱图像分类的准确率。然而,高光谱图像内部的三维结构同样具有重要的参考意义,分别提取光谱与空间特征的方法没有考虑到这方面的信息。在以后的工作中,将考虑引入光谱波段分组策略与注意力机制网络,在三维的高光谱数据块中应用ConvLSTM,提取高光谱图像内部的三维结构信息。

## 参 考 文 献

- [1] Zhao J L, Zhou B, Wang G L, et al. Camouflage target recognition based on dimension reduction analysis of hyperspectral image regions[J]. Photonics, 2022, 9(9): 640.
- [2] Lu B, Dao P, Liu J G, et al. Recent advances of hyperspectral imaging technology and applications in agriculture[J]. Remote Sensing, 2020, 12(16): 2659.
- [3] Zhu C M, Ding J L, Zhang Z P, et al. SPAD monitoring of saline vegetation based on Gaussian mixture model and UAV hyperspectral image feature classification[J]. Computers and Electronics in Agriculture, 2022, 200: 107236.
- [4] 李丹, 孔繁镛, 朱德燕. 基于局部高斯混合特征提取的高光谱图像分类[J]. 光学学报, 2021, 41(6): 0610001.  
Li D, Kong F Q, Zhu D Y. Hyperspectral image classification based on local Gaussian mixture feature extraction[J]. Acta Optica Sinica, 2021, 41(6): 0610001.
- [5] 陈明, 席祥云, 王洋. 利用残差生成对抗网络的高光谱

- 图像分类[J]. 激光与光电子学进展, 2022, 59(22): 2210008.
- Chen M, Xi X Y, Wang Y. Hyperspectral image classification using residual generation against network[J]. *Laser & Optoelectronics Progress*, 2022, 59(22): 2210008.
- [6] 邓子青, 王阳, 张兵, 等. 多尺度特征融合残差网络的高光谱地物分类[J]. 激光与光电子学进展, 2022, 59(18): 1810014.
- Deng Z Q, Wang Y, Zhang B, et al. Classification of hyperspectral ground objects based on multi-scale feature fusion residual network[J]. *Laser & Optoelectronics Progress*, 2022, 59(18): 1810014.
- [7] Zhou P C, Han J W, Cheng G, et al. Learning compact and discriminative stacked autoencoder for hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(7): 4823-4833.
- [8] Guo Y H, Cao H, Bai J J, et al. High efficient deep feature extraction and classification of spectral-spatial hyperspectral image using cross domain convolutional neural networks[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12(1): 345-356.
- [9] 丁繁昌, 郭宝峰, 贾响响, 等. 基于生成对抗网络的高光谱图像特征提取[J]. 光电子·激光, 2021, 32(8): 852-861.
- Ding F C, Guo B F, Jia X X, et al. Feature extraction of hyperspectral images based on generative adversarial networks[J]. *Journal of Optoelectronics·Laser*, 2021, 32(8): 852-861.
- [10] Zhou F, Hang R L, Liu Q S, et al. Hyperspectral image classification using spectral-spatial LSTMs[J]. *Neurocomputing*, 2019, 328: 39-47.
- [11] Liu Q S, Zhou F, Hang R L, et al. Bidirectional-convolutional LSTM based spectral-spatial feature learning for hyperspectral image classification[J]. *Remote Sensing*, 2017, 9(12): 1330.
- [12] Xu Y H, Zhang L P, Du B, et al. Spectral-spatial unified networks for hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(10): 5893-5909.
- [13] Wang D, Du B, Zhang L P, et al. Adaptive spectral-spatial multiscale contextual feature extraction for hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(3): 2461-2477.
- [14] Ismail A, Gunady M K, Pessoa L, et al. Input-cell attention reduces vanishing saliency of recurrent neural networks[EB/OL]. (2019-10-27) [2023-02-01]. <https://arxiv.org/abs/1910.12370>.
- [15] Wang F, Jiang M Q, Qian C, et al. Residual attention network for image classification[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6450-6458.
- [16] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate[EB/OL]. (2014-09-01) [2023-02-01]. <https://arxiv.org/abs/1409.0473>.
- [17] Yang Z C, Yang D Y, Dyer C, et al. Hierarchical attention networks for document classification[C]//Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, June 12-17, 2016, San Diego, California. Stroudsburg: Association for Computational Linguistics, 2016: 1480-1489.
- [18] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018*. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [19] Hu W, Huang Y Y, Wei L, et al. Deep convolutional neural networks for hyperspectral image classification[J]. *Journal of Sensors*, 2015, 2015: 1-12.
- [20] Chen Y S, Jiang H L, Li C Y, 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.