

# 基于改进阶梯网络的高光谱半监督分类算法

关世豪<sup>1</sup>, 杨光<sup>1\*</sup>, 卢珊<sup>2</sup>, 金椿柏<sup>1</sup>, 李豪<sup>3</sup>, 徐昭洪<sup>4</sup>

<sup>1</sup>空军航空大学, 吉林 长春 130022;

<sup>2</sup>东北师范大学地理科学学院, 吉林 长春 130024;

<sup>3</sup>中国人民解放军 95910 部队, 甘肃 酒泉 735000;

<sup>4</sup>中国人民解放军 95795 部队, 广西 桂林 541000

**摘要** 针对现有基于阶梯网络(LN)的高光谱图像分类算法无法充分提取图像的空谱特征而导致分类精度降低的问题,提出一种基于改进阶梯网络的高光谱半监督分类算法。首先将三维卷积神经网络(3D-CNN)与长短时记忆(LSTM)网络结合,提出一种新的空谱特征提取(3D-CNN-LSTM)网络,使用该网络分步提取局部空间特征与光谱特征。然后使用 3D-CNN-LSTM 网络对阶梯网络的编码器与解码器进行改进,提出一种 3D-CNN-LSTM-LN 半监督分类算法,增强阶梯网络的特征提取能力。最后在 Pavia University 和 Indian Pines 两个数据集上对不同算法进行实验。实验结果表明,在小样本条件下,所提算法取得了最佳的分类效果,验证了所提算法具有优越性。

**关键词** 遥感; 高光谱图像; 半监督分类; 阶梯网络; 空谱特征提取

中图分类号 TP751.1

文献标志码 A

doi: 10.3788/LOP202158.2428008

## Hyperspectral Semi-Supervised Classification Algorithm Based on Improved Ladder Network

Guan Shihao<sup>1</sup>, Yang Guang<sup>1\*</sup>, Lu Shan<sup>2</sup>, Jin Chunbai<sup>1</sup>, Li Hao<sup>3</sup>, Xu Zhaohong<sup>4</sup>

<sup>1</sup>Aviation University of Air Force, Changchun, Jilin 130022, China;

<sup>2</sup>School of Geographic Science, Northeast Normal University, Changchun, Jilin 130024, China;

<sup>3</sup>95910 Troop of PLA, Jiuquan, Gansu 735000, China;

<sup>4</sup>95795 Troop of PLA, Guilin, Guangxi 541000, China

**Abstract** Aiming at the problem that the existing hyperspectral image classification algorithm based on ladder network (LN) cannot fully extract the spatial-spectral features of the image, which leads to the reduction of classification accuracy, a hyperspectral semi-supervised classification algorithm based on improved ladder network is proposed. First, the three-dimensional convolutional neural network (3D-CNN) and the long-short-term memory (LSTM) network are combined to propose a new spatial-spectral feature extraction (3D-CNN-LSTM) network, which is used to extract local spatial features step by step. Then, the 3D-CNN-LSTM network is used to improve the encoder and decoder of the ladder network, and a 3D-CNN-LSTM-LN semi-supervised classification algorithm is proposed to enhance the feature extraction ability of the ladder network. Finally, different algorithms are tested on Pavia University and Indian Pines datasets. The experimental results show that the proposed algorithm achieves the best classification effect under the condition of small samples, which verifies the superiority of the proposed algorithm.

**Key words** remote sensing; hyperspectral image; semi-supervised classification; ladder network; spatial-spectral feature extraction

**OCIS codes** 280.4991; 100.4145; 100.2960

收稿日期: 2020-10-29; 修回日期: 2020-12-17; 录用日期: 2020-12-27

基金项目: 国家自然科学基金(41971290)

通信作者: \*yg2599@126.com

## 1 引言

随着光谱遥感技术的发展,高光谱图像的分辨率呈大幅度提升,数据量也急速增长。但是对于高光谱图像的标注仍比较困难,需要逐像素进行,并根据实地考察情况才能确定各像素的类别,从而导致高光谱图像的标注样本相对较少<sup>[1]</sup>。目前,基于深度学习的高光谱图像分类算法必须依靠大量有标注的样本来训练网络,才能得到较好的分类结果。在实际应用中,高光谱图像的标注样本较少且维度较高,容易出现“维数灾难”<sup>[2]</sup>的现象。因此,当面对海量的高光谱图像数据时,如何通过少量的标注样本训练来获得泛化性能更佳的深度学习网络是实际应用中的关键。

在深度学习领域中,半监督学习算法一般通过探索无标注样本在网络训练中的作用来提升模型的泛化性能,可以在极少量标注样本的条件下得到较好的分类结果。通常构建一个深度学习网络对无标注样本数据进行非线性变换,然后优化重构损失函数对网络进行预训练,最后使用少量有标注的样本进行网络微调,例如堆栈自编码器(Stacked Autoencoder, SAE)<sup>[3]</sup>和深度置信网络(Deep Belief Network, DBN)<sup>[4]</sup>。这类处理方式虽然同时使用了无标注样本和有标注样本,但由于无监督学习和监督学习的过程是分离的,并不能算是严格的半监督学习算法。除此之外,无监督特征的重构一般是为了保留原数据中更多的信息,而监督学习的目的是提取与当前任务有关的信息并过滤其他信息,两种学习过程之间存在一定的冲突,融合较为困难。

阶梯网络可以将无监督学习和监督学习两种任务中的信息进行分离,同时优化重构损失和分类损失,能够避免两类学习任务之间的冲突,较好地实现了半监督学习的过程。阶梯网络可以在少量标注样本的条件下,结合大量无标注本来训练网络,从而实现半监督学习,目前在工业<sup>[5]</sup>、医学<sup>[6-7]</sup>和图像识别<sup>[8]</sup>等领域取得了一定的研究成果。除此之外,在高光谱图像半监督分类领域中,Liu 等<sup>[9]</sup>将卷积神经网络(Convolutional Neural Networks, CNN)作为阶梯网络的编码器,用来对高光谱图像进行半监督分类,在少量训练样本的条件下取得比支持向量机(Support Vector Machines, SVM)和 CNN 等经典的监督学习方法更好的分类结果。樊宏渊<sup>[10]</sup>结合图像分割中常用的卷积-反卷积结构将三维卷积神经网络(3D-CNN)作为编码器网络,解码器网络

采用 3D-CNN 的反卷积结构,取得了较高的分类精度。通过改进阶梯网络中编码器与解码器部分的网络结构,使之更适合提取高光谱图像的特征,从而提升分类精度。但文献[11]通过实验发现,仅依靠 CNN 难以充分提取光谱数据中的上下文信息,在空谱特征的提取过程中存在一定不足,而长短时记忆(Long Short Term Memory, LSTM)网络能够有效学习序列数据中的上下文信息。

为此,本文设计一种基于改进阶梯网络的高光谱半监督分类算法。首先根据高光谱图像中空间信息和光谱信息的特性,设计一种结合 3D-CNN 和 LSTM 的空谱特征提取网络 3D-CNN-LSTM,该网络可以有效降低网络的参数量以及增强特征的提取能力。然后借鉴前人对阶梯网络改进的思路,针对高光谱图像分类的问题,以 3D-CNN-LSTM 空谱特征提取网络为基础对阶梯网络进行改进,提出一种新的半监督分类网络,命名为 3D-CNN-LSTM-LN。最后为了验证基于所提网络的算法在小样本条件下的有效性,在 Pavia University 和 Indian Pines 两个数据集中进行分类实验。

## 2 理论基础

### 2.1 3D-CNN

3D-CNN 是在 2D-CNN 的基础上改进而来的,能够较好地学习第三维度的时序信息,在视频分类和动作识别等领域具有巨大的优势。二维卷积与三维卷积对于三维数据的处理过程如图 1 和图 2 所示。对比图 1 和图 2 可以看到,利用二维卷积核来提取特征会破坏三维数据的结构信息,三维卷积核能够充分提取三维数据的空间特征,并且保留时序维度上的信息。对于高光谱数据来说,3D-CNN 更适用于空谱特征的提取。三维卷积核可以同时在这三个方向上进行卷积,输出的特征图为一个三阶张量。在 3D-CNN 的第  $i$  层卷积层第  $j$  个特征图中,  $(x, y, z)$  位置处的输出可表示为

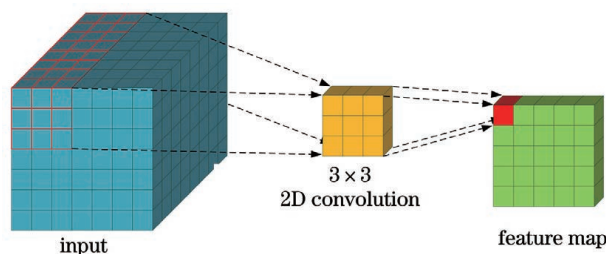


图 1 二维卷积的处理过程

Fig. 1 Process of two-dimensional convolution

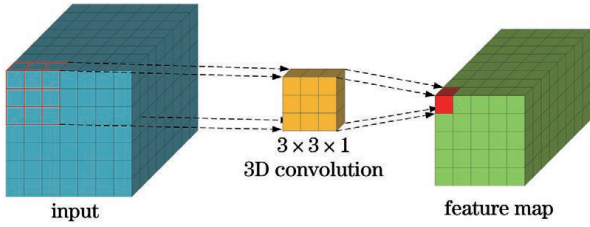


图 2 三维卷积的处理过程

Fig. 2 Process of three-dimensional convolution

$$v_{ij}^{(xyz)} = f\left(b_{ij} + \sum_{k=1}^m \sum_{w=0}^{W_i-1} \sum_{h=0}^{H_i-1} \sum_{b=0}^{B_i-1} u_{ijk}^{(hwb)} v_{(i-1)k}^{(x+w)(y+h)(z+b)}\right), \quad (1)$$

式中： $f$  为激活函数； $m$  为三维卷积核的数量； $W$  为三维卷积核的宽； $H$  为三维卷积核的高； $B$  为三维卷积核的长； $u$  为输入大小； $v$  为与输入相连接的卷积核大小。

在 CNN 中一般使用小卷积核进行特征提取，但是对于高光谱图像的空谱特征来说，使用小卷积核提取空谱特征只能捕捉到临近几个维度的光谱信息，难以获取全局空谱特征，因此本文仅使用 3D-CNN 作为局部空谱特征的提取器。

### 2.2 LSTM 网络

RNN(Recurrent Neural Network)是一种专门用于处理序列数据的神经网络，可以学习数据中的上下文信息。但是当网络中输入的序列数据长度较长时，会存在长期依赖的问题，即序列数据前后距离太大而无法学习较远距离的信息。1997 年，Hochreiter 等<sup>[12]</sup>在 RNN 的原有基础上进行改进，提出了 LSTM 网络。RNN 包括输入门和输出门，LSTM 网络在此基础上增加了遗忘门，结构如图 3 所示，其中  $C_t$  为传递到下一个单元的信息， $\oplus$  为矩阵相加， $\otimes$  为矩阵相乘， $X_t$  为第  $t$  个单元的输入信息， $\sigma$  为 Sigmoid 函数， $o_t$  为 LSTM 单元中的门控

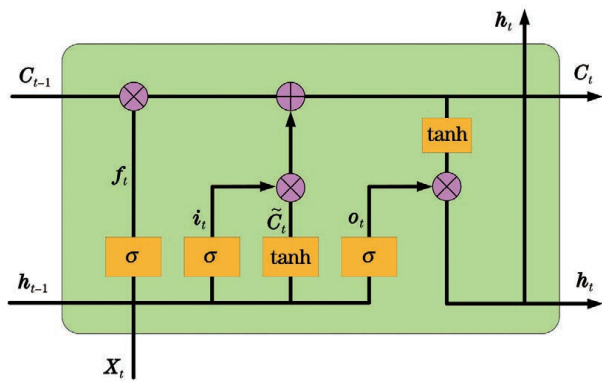


图 3 LSTM 网络的结构

Fig. 3 Structure of LSTM network

信息， $f_t$  为输出， $h_t$  为第  $t$  个单元的传递信息， $\tilde{C}_t$  为由  $\tanh$  函数得到的门控信息， $i_t$  为由 Sigmoid 函数得到的门控信息，接下来将详细介绍 LSTM 网络。

1) 遗忘门是用来决定丢弃哪些信息。首先获取上一个单元的传递信息  $h_{t-1}$  和当前单元的输入  $X_t$ ，经过遗忘门后得到一个  $[0, 1]$  的输出  $f_t$ ，其中“1”表示完全保留，“0”表示完全丢弃。 $f_t$  的计算公式为

$$f_t = \sigma[W_f(h_{t-1}, X_t) + b_f], \quad (2)$$

式中： $b_f$  为遗忘门的偏置； $W_f$  为遗忘门的权值。

2) 输入门是用来确定什么信息将被保留，这里包含两部分。一是 Sigmoid 层是决定什么值将被更新；二是 Tanh 层是用来创建一个新的向量  $\tilde{C}_t$ ， $\tilde{C}_t$  将被添加到下一步状态中。将两部分信息进行结合可以产生新的状态信息，计算公式为

$$i_t = \sigma[W_i(h_{t-1}, X_t) + b_i], \quad (3)$$

$$\tilde{C}_t = \tanh[W_c(h_{t-1}, X_t) + b_c], \quad (4)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t, \quad (5)$$

式中： $b_i$  和  $b_c$  分别为输入门和向量  $\tilde{C}_t$  的偏置； $W_i$  和  $W_c$  分别为输入门和向量  $\tilde{C}_t$  的权值； $*$  为相乘。

3) 输出门是用来得到传递到下一单元的传递信息  $h_t$ 。根据上一个传递过来的状态信息  $C_{t-1}$ ，计算当前单元的输出  $h_t$ ，计算公式为

$$o_t = \sigma[W_o(h_{t-1}, X_t) + b_o], \quad (6)$$

$$h_t = o_t * \tanh(C_t), \quad (7)$$

式中： $b_o$  为输出门的偏置； $W_o$  为输出门的权值。

### 2.3 阶梯网络

在深度学习算法中，半监督学习本质上是无监督学习与监督学习的结合<sup>[13]</sup>。无监督学习的目的是学习原始数据的特征表示，并尽可能地确保该特征表示能够保留足够多的原始数据信息；监督学习的目的是尽可能地保留与当前任务相关的信息，去除其他信息。这决定了监督学习训练与无监督学习训练之间存在着本质上的冲突，也是两类学习任务无法很好兼容在一起的原因。如果在网络中将两类任务信息分开，那么就能实现两类任务的优化。Valpola<sup>[14]</sup>借鉴深度自编码网络的思想，提出阶梯网络并用于无监督特征的提取。随后，Rasmus 等<sup>[15]</sup>对阶梯网络进行改进，提出了半监督阶梯网络(以下仍简称阶梯网络)，该网络可以同时进行监督学习任务和无监督学习任务的训练，阶梯网络的结构如图 4 所示，其中  $\tilde{h}_1$ 、 $\tilde{h}_2$  和  $\tilde{h}_3$  分别为添



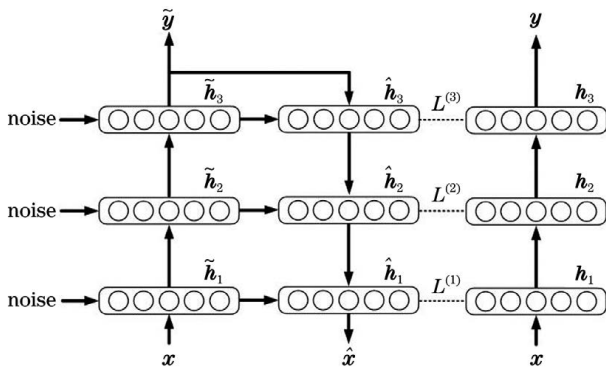


图 4 阶梯网络的结构

Fig. 4 Structure of ladder network

加噪声的第 1、2 和 3 层神经网络的编码输出,  $\hat{h}_1$ 、 $\hat{h}_2$  和  $\hat{h}_3$  分别为第 1、2 和 3 层神经网络解码后的输出,  $x$  和  $\hat{x}$  分别为原始输入向量和编码后解码的向量,  $y$  和  $\tilde{y}$  分别为未添加噪声的网络与添加噪声后的网络的预测标签,  $h_1$ 、 $h_2$  和  $h_3$  分别为未添加噪声的第 1、2 和 3 层神经网络层的编码输出,  $L^{(1)}$ 、 $L^{(2)}$  和  $L^{(3)}$  分别为解码器经过训练后的网络参数。

从网络结构来说, Rasmus 等<sup>[15]</sup> 对于阶梯网络的改进较小, 区别在于将编码部分改为两个共享参数的编码通道, 对其中一个通道加入噪声, 然后通过解码器来降噪重构, 计算重构损失  $L_{Rec}$ ; 另一个通道未加入噪声, 提取到的特征  $y$  用于分类并计算监督分类损失  $L_{Class}$ 。最后将监督分类损失、重构损失以及各层之间的损失  $L_{Layer}$  相加作为总损失, 对网络进行训练可以使其整体梯度下降。损失  $L$  的计算

公式为

$$L = L_{Rec} + L_{Class} + L_{Layers}, \quad (8)$$

其中

$$L_{Layers} = L^{(1)} + L^{(2)} + L^{(3)}. \quad (9)$$

### 3 基于改进阶梯网络的半监督分类算法

#### 3.1 3D-CNN-LSTM 空谱特征提取网络

本文联合采用 3D-CNN 与 LSTM 来提取高光谱图像的空谱特征, 设计了一种 3D-CNN-LSTM 空谱特征提取网络。CNN 中一般使用小卷积核进行特征提取, 但是对于高光谱图像的空谱特征提取来说, 使用小卷积核来提取空谱特征只能够捕捉到临近几个维度的光谱信息, 难以获取全局空谱特征, 所以仅将 3D-CNN 作为局部空谱特征的提取器。通过实验证明,  $3 \times 3 \times 3$  大小的卷积核是提取三维时空特征的有效卷积核<sup>[16]</sup>, 但是对于高光谱图像的空谱特征提取来说, 使用小卷积核来提取时空特征只能够捕捉到临近几个维度的时序信息, 难以获取全局空谱特征。因此本文仅使用  $3 \times 3 \times 1$  大小的卷积核来提取高光谱数据的空间特征, 保留光谱维度。在每一个卷积层处理后对特征进行最大池化操作, 经过三个卷积层与池化层的处理后对特征向量进行重组, 沿着光谱维将三维特征向量拉伸成一维向量; 接下来利用 LSTM 对融合空间特征的特征向量进行建模以提取全局空谱特征, 3D-CNN-LSTM 网络的结构如图 5 所示, 其中 BN 为批归一化层。

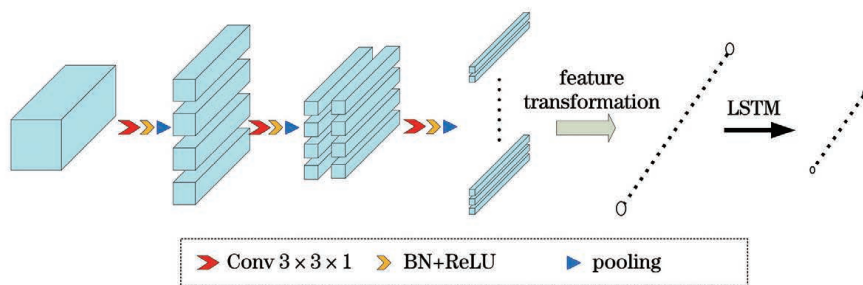


图 5 3D-CNN-LSTM 网络的结构

Fig. 5 Structure of 3D-CNN-LSTM network

#### 3.2 半监督分类网络的结构设计

为了有效提取高光谱图像的空谱特征以及提高半监督分类的效果, 使用 3D-CNN-LSTM 网络作为阶梯网络中的编码器网络, 并设计了一种新的半监督分类算法。使用 PCA (Principal Component Analysis) 算法对原始数据进行降维的预处理, 避免直接使用原始高光谱数据进行分类而造成“维数灾

难”。本文提出的基于改进阶梯网络的半监督分类算法的模型如图 6 所示。

由于阶梯网络中编码器与对应解码器各层的特征向量需要在维度上统一才可以计算重构误差损失, 而 3D-CNN-LSTM 包括 3D-CNN 和 LSTM 两种类型的网络。原始的阶梯网络使用全连接网络来构建编码器与解码器, 输入与输出都是一维

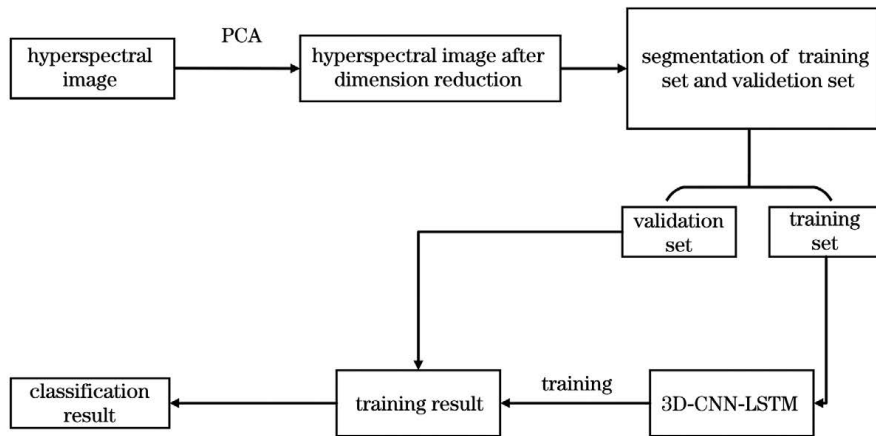


图 6 基于改进阶梯网络的半监督分类算法的模型

Fig. 6 Model of semi-supervised classification algorithm based on improved ladder network

向量,只需确保对应全连接层的神经元数量相同; CNN中包含卷积和池化等操作,而且受到步长和零填充等因素的影响,难以在解码器中按照全连接层的连接方式进行设计。图像分割的过程中, U-Net使用反卷积结构进行图像重建,因此阶梯网络中的解码器部分选择3D-CNN反卷积结构作为

解码层;LSTM可以输出任意长度的一维向量,所以同样采用LSTM作为与编码层对应的解码层,3D-CNN-LSTM阶梯网络的结构如图7所示,其中Bi-LSTM为双向长短时记忆网络, $L_{Cross-Entropy}$ 为交叉熵损失, $L_{MSE}$ 为均方误差损失,DeConv为反卷积。

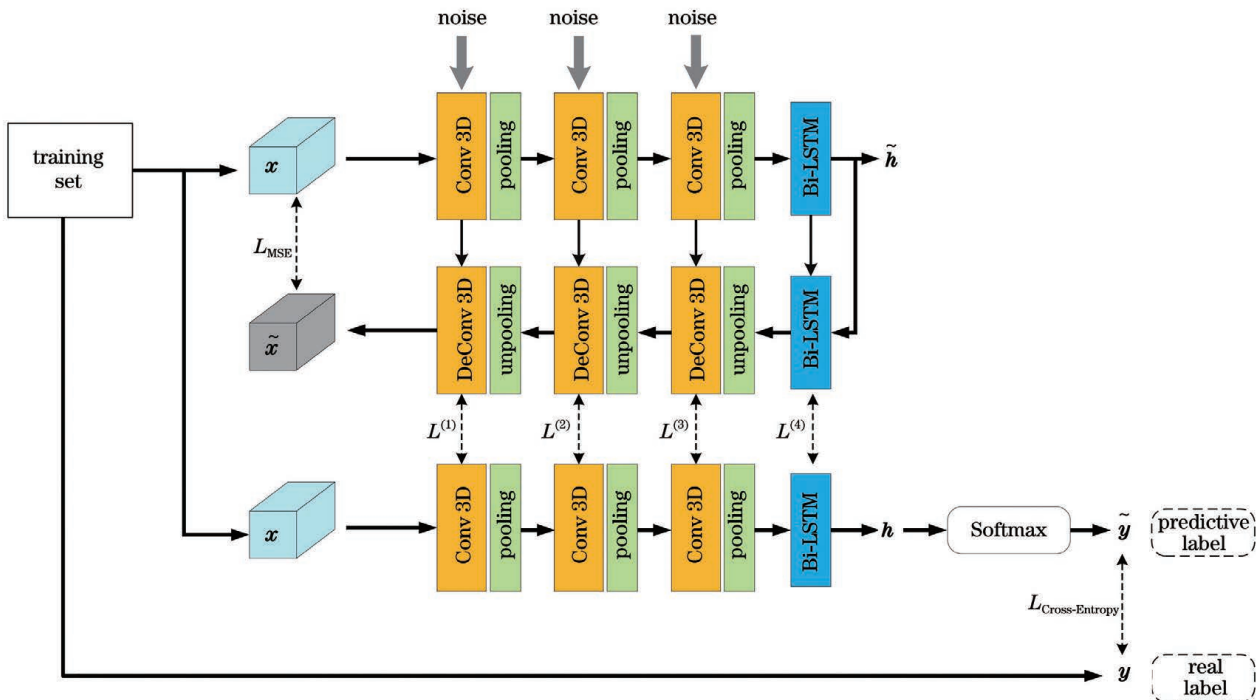


图 7 3D-CNN-LSTM 阶梯网络的结构

Fig. 7 Structure of 3D-CNN-LSTM ladder network

3D-CNN-LSTM 阶梯网络的损失由三部分组成:第一部分是经过无噪声编码通道后得到的样本预测标签和真实标签之间的交叉熵损失  $L_{Cross-Entropy}$ ;第二部分是经过解码通道后生成的重构样本与输入样本之间的均方误差损失  $L_{MSE}$ ;最后一部分是无噪

声编码通道与解码通道对应网络层得到的特征之间的误差损失  $\{L^{(1)}, L^{(2)}, L^{(3)}, L^{(4)}\}$ ,同样使用均方误差损失函数来计算。3D-CNN-LSTM 阶梯网络的损失为上述三部分损失的和,使用优化算法,即基于改进梯形网络的高光谱半监督分类算法对网络进行

训练,整体算法流程如下。

输入:高光谱图像  $\mathbf{X} \in \mathbf{R}^{W \times H \times B}$ 。

输出:高光谱图像的分类精度及分类结果。

1) 对原数据进行最小值与最大值的标准化处理并将所有像元值缩放到  $[0, 1]$  区域,然后采用 PCA 算法对数据进行降维,得到数据  $\tilde{\mathbf{X}} \in \mathbf{R}^{W \times H \times k}$ 。

2) 对高光谱图像的二维空间边缘进行零填充,以图像上所有像元为中心点依次选取  $n \times n \times k$  大小的数据块以构建样本集,其中  $n$  为数据块的边长。

3) 对中心像元有标签的样本按照训练集与验证集的比例将样本集随机划分为训练集  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{b_1}\}$  与验证集  $\{\mathbf{x}_{b_2}, \mathbf{x}_{b_2+1}, \dots, \mathbf{x}_{b_1}\}$ ,  $b$  为波段。

4) 将训练集与所有无标签样本作为训练集并输入到半监督分类模型中,使用反向传播算法对模型进行训练,当达到最大训练迭代次数时结束训练。

5) 使用验证集  $\{\mathbf{x}_{b_2}, \mathbf{x}_{b_2+1}, \dots, \mathbf{x}_{b_1}\}$  训练模型,得到验证集的分类准确率。

6) 将步骤 2) 生成的所有样本输入到训练好的

模型中,得到分类结果。

## 4 实验数据集

为了验证本文算法的性能,使用两个不同传感器所获取的高光谱图像数据进行实验,这两个数据集在空间分辨率、光谱分辨率、地物类别及标记样本数目等方面均不相同。

### 4.1 Pavia University 数据集

Pavia University 数据集是由机载成像光谱仪 (ROSIS) 于 2001 年在意大利的帕维亚大学拍摄的。未经处理的原始高光谱图像有 115 个波段,图像大小为  $610 \text{ pixel} \times 340 \text{ pixel}$ ,空间分辨率为 1.3 m,波长范围为 430~860 nm。去除水汽吸收和噪声所影响的 13 个波段,对剩下 102 个波段的图像进行实验。数据集中已标注的样本共包含 9 个不同的人造结构地物类别,训练样本数较多,不同地物间的光谱特征差异较大,可用于验证小样本条件下模型的分类能力。Pavia University 数据集的真彩色图像和地物真值图如图 8 所示。

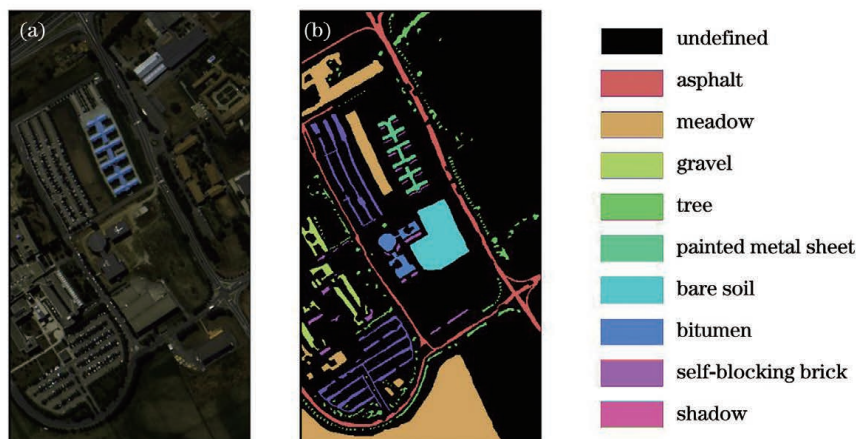


图 8 Pavia University 数据集的真彩色图像和地物真值图。(a)真彩色图像;(b)地物真值图

Fig. 8 True color image and ground truth map of Pavia University dataset. (a) True color image; (b) ground truth map

### 4.2 Indian Pines 数据集

Indian Pines 数据集是由机载成像光谱仪 (AVIRIS) 于 1992 年 6 月在美国印第安纳州西北部的农业区拍摄的。未经处理的原始高光谱图像有 224 个波段,图像大小为  $145 \text{ pixel} \times 145 \text{ pixel}$ ,空间分辨率为 20 m,波长范围为 400~2500 nm。去除水汽吸收和噪声所影响的 20 个波段 (104~108 nm, 150~163 nm, 220 nm),对剩下 200 个波段的图像进行实验。已标注的样本共包含 16 个不同种类的地物,绝大部分为农作物以及其他植物,如玉米、大豆、草和树等。由于拍摄时间为 6 月份,正

值植物生长的繁荣期和农作物发育的早期阶段,所以地物间光谱特征的相似度较高,难以区分,且各类样本数量不均,为光谱特征的学习带来较大困难。Indian Pines 数据集的真彩色图像和地物真值图如图 9 所示。

## 5 实验结果与分析

### 5.1 训练方案及模型参数

实验使用带动量的 SGD (Stochastic Gradient Descent) 优化器进行训练,网络训练周期设为 150,批归一化大小设为 100。初始学习率为 0.001,动量



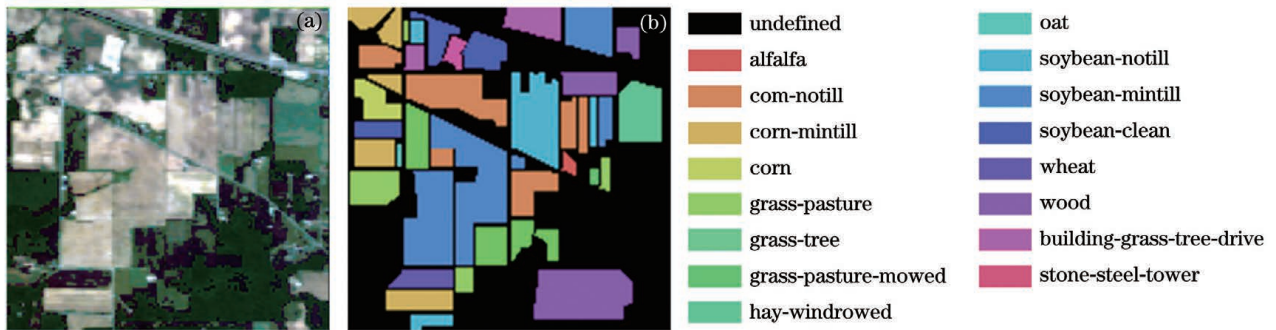


图 9 Indian Pines 数据集的真彩色图像和地物真值图。(a)真彩色图像;(b)地物真值图

Fig. 9 True color image and ground truth map of Indian Pines dataset. (a) True color image; (b) ground truth map

为 0.9, 学习率的更新采用自适应调整策略, 当测试精度在 10 个周期内不发生明显变化时, 学习率下降一个数量级。在数据准备阶段, 首先对每类样本进行随机打乱以确保数据的随机分布, 然后对训练集和验证集进行划分。在对比实验中, 分别从 Pavia University 和 Indian Pines 数据集中选择 5% 和 10% 的有标注样本作为训练集进行训练, 其余作为验证集以验证实验结果, 半监督分类算法中还需要将所有的未标注样本作为训练集以进行网络训练。

高光谱数据经过预处理后, 采用 PCA 算法将其

光谱维度降至 30。借鉴文献 [11] 中的结论, 将 Pavia University 和 Indian Pines 数据集输入到网络中, 训练样本的最优空间域尺寸分别为 11 pixel × 11 pixel 和 15 pixel × 15 pixel, 因此将两数据集的训练样本大小确定为 11 pixel × 11 pixel 和 15 pixel × 15 pixel。以 Pavia University 数据集为例, 3D-CNN-LSTM-LN 模型中编码器部分的网络结构如表 1 所示, 其网络结构与 3D-CNN-LSTM 相同, 区别在于输入样本的维度发生了变化, 而解码器部分的网络结构与编码器正好相反, 因此不再赘述。

表 1 3D-CNN-LSTM-LN 模型中编码器部分的网络结构

Table 1 Network structure of encoder in 3D-CNN-LSTM-LN model

Layer	Kernel size	Input size	Output size	Number of convolution kernels	Activation function
Input	-	-	11 × 11 × 30	-	-
Conv-1(Conv 3D)	3 × 3 × 1	11 × 11 × 30	9 × 9 × 30	4	-
BN-1	-	9 × 9 × 30	9 × 9 × 30	-	ReLU
Pooling-1(MaxPooling)	1 × 1 × 2	9 × 9 × 30	9 × 9 × 15	-	-
Conv-2(Conv 3D)	3 × 3 × 1	9 × 9 × 15	7 × 7 × 15	8	-
BN-2	-	7 × 7 × 15	7 × 7 × 15	-	ReLU
Pooling-2(MaxPooling)	1 × 1 × 2	7 × 7 × 15	7 × 7 × 7	-	-
Conv-3(Conv 3D)	3 × 3 × 1	7 × 7 × 7	5 × 5 × 7	16	-
BN-3	-	5 × 5 × 7	5 × 5 × 7	-	ReLU
Pooling-3(MaxPooling)	1 × 1 × 2	5 × 5 × 7	5 × 5 × 3	-	-
Linear-1(flatten)	-	5 × 5 × 3	1 × 75	-	-
RNN-1(Bi-LSTM)	-	1200 × 1	400 × 1	-	Tanh

## 5.2 实验结果分析

为了评价本文算法的分类性能, 将其与 4 种高光谱图像分类算法进行对比实验, 其中包括基于 3D-CNN 的监督分类 (M3D-DCNN) 算法<sup>[17]</sup>、基于

2D-CNN 的半监督分类 (SS-CNN) 算法<sup>[9]</sup>、基于 3D-CNN 的半监督分类 (S4CNN) 算法<sup>[10]</sup> 以及由本文提出的空谱特征提取网络改进的监督分类 (3D-CNN-LSTM) 算法。M3D-DCNN 算法采用三维卷积对高

光谱数据的结构特征进行空谱特征的提取,网络参数量更少,更适合进行高光谱图像的分类;SS-CNN 算法和 S4CNN 算法都是利用 CNN 对阶梯网络进行改进以实现半监督分类,区别在于前者使用 2D-CNN,后者使用 3D-CNN。以上几种算法都是近几年提出的高光谱图像分类算法,作为对比算法更具说服力。为了更好地对比分类效果,所有模型中训练集与验证集的分配比例和最大训练迭代次数均与本文实验设置相同,而训练样本大小、权重初始化、学习率以及优化器选择等条件则与原算法相同,确保对比算法达到最好的分类效果。

### 5.2.1 Pavia University 数据集的实验结果分析

不同算法在 Pavia University 数据集上的分类精度如表 2 所示。从表 2 可以看到,在 Pavia University 数据集上,本文提出的 3D-CNN-LSTM-LN 算法在总体分类精度(OA)、平均分类精度(AA)和 Kappa 系数三个分类指标上均高于其他几种算法,本文算法的总体分类精度达到了 97.04%;在小样本条件下,3D-CNN-LSTM-LN 算法在三个

分类指标上远高于用于监督分类的 M3D-DCNN 算法和 3D-CNN-LSTM 算法;尽管 3D-CNN-LSTM 算法的分类结果比 M3D-DCNN 算法差,但结合半监督思想改进之后的 3D-CNN-LSTM-LN 半监督分类算法在总体分类精度、平均分类精度和 Kappa 系数上分别比 M3D-DCNN 算法高 5.30, 7.97, 7.10 个百分点,这说明即使特征提取网络较差,但结合半监督学习的思想依然能够取得更好的分类结果;对于用于半监督分类的 SS-CNN 算法和 S4CNN 算法,从总体分类精度上看,3D-CNN-LSTM-LN 算法比 SS-CNN 算法高 4.57 个百分点,比 S4CNN 算法高 1.90 个百分点;从平均分类精度上看,3D-CNN-LSTM-LN 算法比 SS-CNN 算法高 5.63 个百分点,比 S4CNN 算法高 2.95 个百分点;从 Kappa 系数上看,3D-CNN-LSTM-LN 算法比 SS-CNN 算法高 6.00 个百分点,比 S4CNN 算法高 1.80 个百分点;3D-CNN-LSTM-LN 算法在 9 类地物的分类中,有 8 类地物的分类精度在五类算法中达到最优。

表 2 不同算法在 Pavia University 数据上的分类精度

Table 2 Classification accuracy of different algorithms on Pavia University dataset

unit: %

No.	Class	M3D-DCNN	3D-CNN-LSTM	SS-CNN	S4CNN	3D-CNN-LSTM-LN
1	Asphalt	93.25	93.03	94.98	96.94	<b>99.21</b>
2	Meadow	98.14	96.99	96.84	99.11	<b>99.63</b>
3	Gravel	76.83	75.13	89.16	86.97	<b>96.22</b>
4	Tree	97.78	91.45	93.50	98.33	<b>98.94</b>
5	Metal sheet	<b>100.00</b>	96.63	99.33	99.78	99.66
6	Bare soil	92.87	90.31	92.28	97.40	<b>99.18</b>
7	Bitumen	80.56	86.97	83.21	91.76	<b>97.79</b>
8	Brick	83.14	83.95	90.92	93.50	<b>98.13</b>
9	Shadow	94.25	88.52	96.84	97.00	<b>98.55</b>
OA	-	93.73	92.46	94.46	97.06	<b>99.03</b>
AA	-	90.62	89.13	92.96	95.64	<b>98.59</b>
Kappa	-	91.70	89.90	92.80	96.20	<b>98.80</b>

不同算法在 Pavia University 数据集上的分类结果如图 10 所示。从图 10 可以看到,由于这些算法都是基于空谱特征进行分类的,所以分类结果中的“椒盐噪声”几乎不存在,但是各种算法对分类图中地物边缘以及未标注区域的分类效果相差比较明显,相比于其他算法,本文提出的 3D-CNN-LSTM-LN 算法对道路边缘(图 10 虚线框区域)和建筑物屋顶(图 10 中实线框区域)部分的分类结果更加清

晰,未标注区域的分类结果与图 10(a)真彩色图像中的地物分布更加接近。综上可以充分说明,在 Pavia University 数据集上,本文提出的半监督算法能够取得最佳的分类结果,算法的性能最好。

### 5.2.2 Indian Pines 数据集的实验结果分析

不同算法在 Indian Pines 数据集上的分类精度如表 3 所示。从表 3 可以看到,在 Indian Pines 数据集上,3D-CNN-LSTM-LN 算法在总体分类精度、



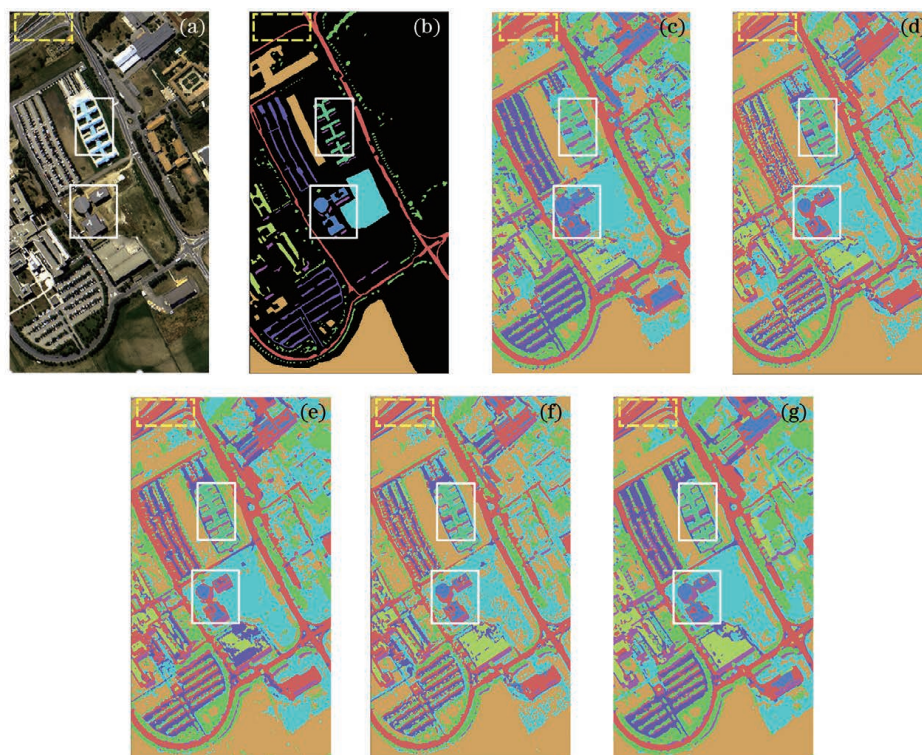


图 10 不同算法在 Pavia University 数据集上的分类结果。(a)真彩色图像;(b)地物标签图;(c)M3D-DCNN 算法;(d)3D-CNN-LSTM 算法;(e)SS-CNN 算法;(f)S4CNN 算法;(g)3D-CNN-LSTM-LN 算法

Fig. 10 Classification results of different algorithms on Pavia University dataset. (a) True color image; (b) feature label map; (c) M3D-DCNN algorithm; (d) 3D-CNN-LSTM algorithm; (e) SS-CNN algorithm; (f) S4CNN algorithm; (g) 3D-CNN-LSTM-LN algorithm

表 3 不同算法在 Indian Pines 数据上的分类精度

Table 3 Classification accuracy of different algorithms on Indian Pines dataset

unit: %

No.	Class	M3D-DCNN	3D-CNN-LSTM	SS-CNN	S4CNN	3D-CNN-LSTM-LN
1	Alfalfa	39.12	77.65	81.91	<b>97.23</b>	96.85
2	Corn-notill	77.95	82.45	83.24	91.77	<b>96.93</b>
3	Corn-mintill	88.67	82.36	74.34	89.34	<b>93.97</b>
4	Corn	87.02	66.75	79.67	96.38	<b>97.15</b>
5	Grass-pasture	95.43	93.17	<b>97.65</b>	92.93	95.49
6	Grass-tree	98.78	95.36	88.96	<b>99.36</b>	98.17
7	Grass-pasture-mowed	70.31	88.95	75.02	<b>97.72</b>	62.13
8	Hay-windrowed	95.33	97.34	89.61	98.33	<b>98.88</b>
9	Oat	53.35	56.02	<b>82.47</b>	72.03	57.15
10	Soybean-notill	81.63	90.14	82.35	92.19	<b>97.31</b>
11	Soybean-mintill	86.27	88.46	82.63	92.84	<b>98.19</b>
12	Soybean-clean	80.68	74.28	83.28	83.96	<b>92.56</b>
13	Wheat	99.45	97.67	81.36	97.95	<b>99.70</b>
14	Wood	94.59	98.32	87.64	97.71	<b>99.65</b>
15	Building-grass-tree-drive	76.95	90.24	80.27	89.72	<b>95.12</b>
16	Stone-steel-tower	95.27	46.83	<b>97.79</b>	90.98	93.35
OA	-	87.20	88.83	89.56	93.34	<b>97.04</b>
AA	-	82.51	83.22	84.26	82.40	<b>92.03</b>
Kappa	-	85.30	87.20	88.80	92.30	<b>96.60</b>

平均分类精度和 Kappa 系数三个分类指标上均高于其他几种算法,本文算法的总体分类精度达到了 99.03%;在小样本条件下,3D-CNN-LSTM-LN 算法在三个分类指标上远高于用于监督分类的 M3D-DCNN 算法和 3D-CNN-LSTM 算法;特别是对于 3D-CNN-LSTM 算法,3D-CNN-LSTM-LN 算法的总体分类精度、平均分类精度和 Kappa 系数分别比 3D-CNN-LSTM 算法高 8.21, 8.81, 9.40 个百分点,说明利用有效的特征提取网络来改进阶梯网络进行半监督分类可以有效提升分类精度,这与在 Pavia University 数据集上的实验结论相同;对比用于半监督分类的 SS-CNN 算法和 S4CNN 算法,从总体分类精度上看,3D-CNN-LSTM-LN 算法比 SS-CNN 算法高 7.08 个百分点,比 S4CNN 算法高 3.70 个百分点;从平均分类精度上看,3D-CNN-LSTM-LN 算法比 SS-CNN 算法高 7.77 个百分点,比 S4CNN 算法高 9.63 个百分点;从 Kappa 系数上看,3D-CNN-LSTM-LN 算法比 SS-CNN 算法高 7.80 个百分点,比 S4CNN 算法高 4.30 个百分点;3D-CNN-LSTM-LN 算法在 16 类地物的分类中,

有 10 类地物的分类精度在五类算法中达到最优,但是对于燕麦地物的分类精度较差,仅达到 57.15%。

采用 5 种算法对 Indian Pines 数据集进行分类,结果如图 11 所示。从图 11 可以看到,相比于其他算法,本文提出的 3D-CNN-LSTM-LN 算法对道路边缘(图 11 虚线框区域)和农田(图 11 实线框区域)部分的分类效果更加清晰,未标注区域的分类结果与图 11(a)真彩色图像中的地物分布更加接近。综合以上的实验结果来看,在 Indian Pines 数据集上的总体情况与 Pavia University 数据类似,3D-CNN-LSTM-LN 算法获得了最佳的分类结果,证明了所提算法的优越性。除此之外还可以发现,虽然 M3D-DCNN 算法和 3D-CNN-LSTM 算法的分类精度和 Kappa 系数比较接近(仅相差 1 个百分点左右),但是从图 11(c)和图 11(d)可以看到,M3D-DCNN 算法对于未标注区域的分类效果较差,忽略了大量的地物细节信息,这从侧面反映出对算法分类效果的判断不能仅靠分类精度,更需要从视觉效果来对比实际的分类结果。

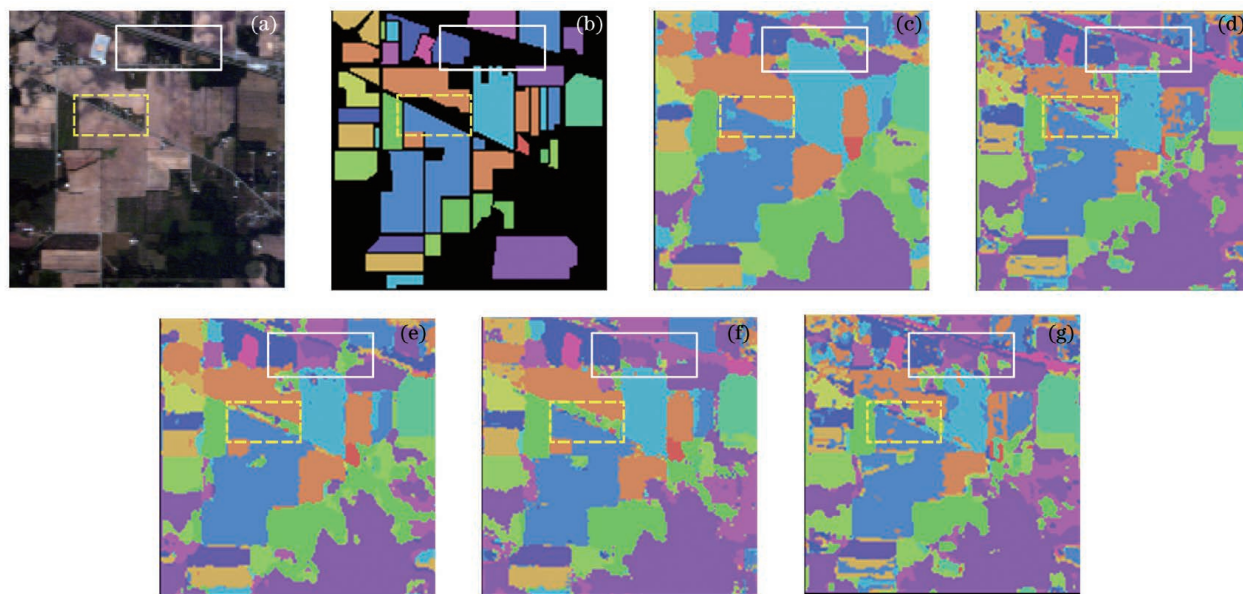


图 11 不同算法在 Indian Pines 数据集上的分类结果图。(a)真彩色图像;(b)地物标签图;(c)M3D-DCNN 算法;(d)3D-CNN-LSTM 算法;(e)SS-CNN 算法;(f)S4CNN 算法;(g)3D-CNN-LSTM-LN 算法

Fig. 11 Classification results of different algorithms on Indian Pines dataset. (a) True color image; (b) feature label map; (c) M3D-DCNN algorithm; (d) 3D-CNN-LSTM algorithm; (e) SS-CNN algorithm; (f) S4CNN algorithm; (g) 3D-CNN-LSTM-LN algorithm

## 6 结 论

本文充分挖掘高光谱图像中空间信息和光谱信息的特性将 3D-CNN 和 LSTM 结合,利用两类网络

提取不同特征的优势进行互补,设计了一种高效的空谱特征提取网络 3D-CNN-LSTM;然后针对高光光谱图像的分类问题,以 3D-CNN-LSTM 空谱特征提取网络为基础,对阶梯网络进行改进,以实现半监督

分类。在 Pavia University 和 Indian Pines 两个数据集上仅依靠 5% 和 10% 的标注样本进行分类实验,本文算法的总体分类精度分别达到了 97.04% 和 99.03%,比其他几种算法取得了更好的分类精度,本文提出的空谱特征提取网络较现有的单独依靠 CNN 提取特征更加充分。

## 参 考 文 献

- [1] Zhang H K, Li Y, Jiang Y N. Deep learning for hyperspectral imagery classification: the state of the art and prospects[J]. *Acta Automatica Sinica*, 2018, 44(6): 961-977.  
张号逵, 李映, 姜晔楠. 深度学习在高光谱图像分类领域的研究现状与展望[J]. *自动化学报*, 2018, 44(6): 961-977.
- [2] Hughes G. On the mean accuracy of statistical pattern recognizers[J]. *IEEE Transactions on Information Theory*, 1968, 14(1): 55-63.
- [3] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504-507.
- [4] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. *Neural Computation*, 2006, 18(7): 1527-1554.
- [5] Shi F Y, Wang Z Y, Liang J. Fault classification based on semi-supervised dense ladder network[J]. *CIESC Journal*, 2018, 69(7): 3083-3091.  
施方迺, 汪子扬, 梁军. 基于半监督密集阶梯网络的工业故障识别[J]. *化工学报*, 2018, 69(7): 3083-3091.
- [6] Hou L, Gui W. Research on the detection method of infrared breast cancer based on semi-supervised ladder network[J]. *Information Technology and Informatization*, 2018(6): 179-182.  
侯丽, 桂伟. 基于半监督阶梯网络的红外乳腺癌检测方法研究[J]. *信息技术与信息化*, 2018(6): 179-182.
- [7] Jin L Y, Guo S X, Ma S Z, et al. Liver segmentation in CT image based on semi-supervised ladder network[J]. *Journal of Jilin University (Information Science Edition)*, 2018, 36(2): 158-164.  
金兰依, 郭树旭, 马树志, 等. 基于半监督阶梯网络的肝脏 CT 影像分割[J]. *吉林大学学报(信息科学版)*, 2018, 36(2): 158-164.
- [8] Li X. High resolution SAR image classification based on curvelet DLN [D]. Xi'an: Xidian University, 2018: 10-35.  
李晰. 基于曲线波 DLN 的高分辨 SAR 图像地物分类[D]. 西安: 西安电子科技大学, 2018: 10-35.
- [9] Liu B, Yu X C, Zhang P Q, et al. A semi-supervised convolutional neural network for hyperspectral image classification[J]. *Remote Sensing Letters*, 2017, 8(9): 839-848.
- [10] Fan H Y. Semi-supervised hyperspectral image classification based on deep learning[D]. Xi'an: Xidian University, 2019: 7-20.  
樊宏渊. 基于深度学习的半监督高光谱图像分类[D]. 西安: 西安电子科技大学, 2019: 7-20.
- [11] Guan S H, Yang G, Li H, et al. Hyperspectral image classification based on 3-D convolutional recurrent neural network[J]. *Laser Technology*, 2020, 44(4): 485-491.  
关世豪, 杨枕, 李豪, 等. 3 维卷积递归神经网络的高光谱图像分类方法[J]. *激光技术*, 2020, 44(4): 485-491.
- [12] Hochreiter S, Schmidhuber J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [13] Fu C X. Research on semi-supervised learning with ladder network method and application[D]. Changsha: National University of Defense Technology, 2016: 6-15.  
符春晓. 基于阶梯网络的半监督深度学习方法及其应用研究[D]. 长沙: 国防科学技术大学, 2016: 6-15.
- [14] Valpola H. From neural PCA to deep unsupervised learning [EB/OL]. (2014-11-28) [2020-10-25]. <http://arxiv.org/abs/1411.7783>.
- [15] Rasmus A, Valpola H, Honkala M, et al. Semi-supervised learning with ladder networks [EB/OL]. (2015-07-09) [2020-10-25]. <https://arxiv.org/abs/1507.02672v2>.
- [16] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C] // *Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012, December 3-6, 2012, Lake Tahoe, Nevada, United States*. [S.l.: s.n.], 2012: 1106-1114.
- [17] He M Y, Li B, Chen H H. Multi-scale 3D deep convolutional neural network for hyperspectral image classification[C]//2017 IEEE International Conference on Image Processing (ICIP), September 17-20, 2017, Beijing, China. New York: IEEE Press, 2017: 3904-3908.