

## 基于多尺度自适应胶囊网络的高光谱遥感分类

张根<sup>1,2,3</sup>, 丁小辉<sup>1,3\*</sup>, 杨骥<sup>1,3</sup>, 王华<sup>2</sup><sup>1</sup>广东省科学院广州地理研究所, 广东 广州 510070;<sup>2</sup>广东工业大学土木与交通工程学院, 广东 广州 510006;<sup>3</sup>南方海洋科学与工程广东省实验室(广州), 广东 广州 511458

**摘要** 胶囊网络(CapsNet)是一种新型的神经网络,已被广泛应用于高光谱遥感分类。针对胶囊网络所面临的过拟合和梯度消失等问题,提出一种基于多尺度自适应胶囊网络(MSCaps)的高光谱遥感分类方法。利用多尺度(不同尺寸的输入图像)的卷积层来提取地物的空间特征与光谱特征,并采用一种非迭代自适应路由算法进一步改进CapsNet的结构,避免了耦合系数 $c_{ij}$ 的稀疏化引起的过拟合问题。利用PU和SA两个公共高光谱数据集从总体分类精度(OA)和模型训练效率两方面评价了MSCaps的分类性能。在分类精度上,利用MSCaps与支持向量机(SVM)、随机森林(RF)、深度卷积神经网络(CNN)、CapsNet、多尺度的胶囊网络(MCaps)、基于非迭代自适应路由算法的胶囊网络(ARWI-Caps)、多尺度的卷积神经网络(MSCNN)对原始影像的OA进行对比;还与SVM、RF对经主成分分析(PCA)特征提取后的影像的OA进行对比,其中两种分类方法分别简称为PCA-SVM与PCA-RF。在训练效率上,对MSCaps的训练时间和CNN、CapsNet、MSCNN的训练时间进行对比。实验结果表明:MSCaps对PU和SA数据集的OA分别为99.14%和95.38%,分类精度均高于SVM、RF、PCA-SVM、PCA-RF、CNN、CapsNet、MCaps、ARWI-Caps及MSCNN;在模型训练效率方面,MSCaps在PU和SA两种数据集上的耗时分别约为CapsNet的1/3与1/4,训练效率较CapsNet有大幅提高。因此,所提MSCaps可以有效解决基于CapsNet的高光谱遥感分类问题。

**关键词** 高光谱遥感; 胶囊网络; 非迭代自适应路由算法; 深度学习; 遥感分类

中图分类号 P237

文献标志码 A

DOI: 10.3788/LOP202259.2428004

## Hyperspectral Remote Sensing Classification Using Multi-Scale Adaptive Capsule Network

Zhang Gen<sup>1,2,3</sup>, Ding Xiaohui<sup>1,3\*</sup>, Yang Ji<sup>1,3</sup>, Wang Hua<sup>2</sup><sup>1</sup>Guangzhou Institute of Geography, Guangdong Academy of Sciences, Guangzhou 510070, Guangdong, China;<sup>2</sup>School of Civil and Transportation Engineering, Guangdong University of Technology, Guangzhou 510006, Guangdong, China;<sup>3</sup>Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 511458, Guangdong, China

**Abstract** A capsule network (CapsNet) is a novel neural network that has been widely used in the classification of hyperspectral remote sensing. However, it is faced with overfitting and gradient vanishing. To solve this problem, this paper proposes a hyperspectral remote sensing classification method using a multi-scale adaptive capsule network (MSCaps). The multi-scale convolution layer was used to extract the spatial and spectral features of the ground object from the input images with different sizes (namely, multi-scale). To solve the overfitting problem caused by the sparseness of the coupling coefficient  $c_{ij}$ , we applied an adaptive routing algorithm without iteration to further improve the CapsNet structure. To validate the proposed model, we evaluated the classification performances of MSCaps using overall classification accuracy (OA) and model training efficiency on two public hyperspectral remote sensing datasets, namely, the Pavia University (PU) and Salinas-A (SA) datasets. The OA of MSCaps was measured and compared with that of the benchmarks, including support vector machine (SVM), random forest (RF), deep convolutional neural network (CNN),

收稿日期: 2021-09-07; 修回日期: 2021-10-20; 录用日期: 2021-11-02

基金项目: 国家自然科学基金(41976189)、广东省科学院建设国内一流研究机构专项资金(2019GDASYL-0301001, 2021GDASYL-20210103001)、南方海洋科学与工程广东省实验室(广州)重大专项团队项目(GML2019ZD0301)

通信作者: \*dxh2017@sina.com

CapsNet, multi-scale capsule network (MCaps), capsule network using adaptive routing algorithm without iteration (ARWI-Caps), and the multi-scale CNN (MSCNN) on original images. Additionally, the OA of MSCaps was compared with that of SVM and RF on images extracted using principal component analysis (PCA). They are named PCA-SVM and PCA-RF, respectively. The training efficiency of MSCaps was compared with that of CNN, CapsNet, and MSCNN. The experimental results show that the OA of CapsNet on the PU and SA datasets is 99.14% and 95.38%, respectively, which is higher than that of the benchmarks. Additionally, the training time of MSCaps is  $\sim 1/3$  and  $\sim 1/4$  of that of CapsNet on the PU and SA datasets, respectively. Thus, the training efficiency of MSCaps is significantly higher than that of CapsNet. Therefore, the proposed hyperspectral remote sensing classification method using MSCaps has a good application potential and is an effective alternative for hyperspectral remote sensing classification.

**Key words** hyperspectral remote sensing; capsule network; adaptive routing algorithm without iteration; deep learning; remote sensing classification

## 1 引言

土地利用类型信息在土地资源监测、城市规划与管理、气候变化研究、生态环境保护等方面具有重要作用,尤其在生态环境演替与气候变化关系、区域可持续发展研究等方面,土地利用类型是不可或缺的重要信息之一<sup>[1-2]</sup>。其中,精准的土地利用类型信息是进一步促进与深化相关研究的重要基础,然而随着经济的快速发展,土地利用斑块正变得越来越破碎,这对精准获取土地利用类型信息提出了巨大挑战。而随着遥感技术的不断发展,高空间分辨率、高光谱分辨率、高时间分辨率的遥感影像为实时精准地获取土地利用类型信息奠定了坚实的数据基础。

其中,高光谱遥感影像(HSI)具有光谱分辨率高、波谱连续及“图谱合一”等优点,可为土地利用类型精准分类提供有效数据。HSI往往包含数十至数百个窄波段的地物光谱信号,可为实现土地利用类型精细分类提供丰富的光谱信息,如提高对相似地物类型(不同材质的道路、不同生长期的植被等)的识别能力<sup>[3]</sup>。同时,随着成像技术的发展,HSI的空间分辨率得到了显著提升,提供的空间纹理信息更加详细,能够更清晰地描述地物的空间特征,从而实现对土地利用类型的精细分类<sup>[4]</sup>。然而,在样本数目有限的情况下,HSI容易引发“维数灾难”<sup>[5]</sup>。为解决这个问题,目前常利用特征提取和特征选择的方法来降低高光谱数据的维度<sup>[6-7]</sup>,并在此基础上结合支持向量机(SVM)和随机森林(RF)等方法进行高光谱遥感分类。但上述分类器都是基于人工特征构造的且大部分都属于浅层分类器,因而难以对具有复杂结构的土地利用类型进行精准分类。

深度学习是近年来人工智能领域的重大研究突破,其深度网络结构可有效模拟人类大脑进行自主特征提取,以充分利用低层与高层语义特征,满足分类、目标识别等应用需求<sup>[8]</sup>。深度卷积神经网络(CNN)是深度学习领域中具有代表性的“端到端”特征学习方法,由于具有较好的鲁棒性和泛化能力,已被广泛应用于图像分类、目标识别等领域,如在遥感领域,CNN被广泛应用于土地利用、农作物和湿地等的分类<sup>[9-17]</sup>。但

是,CNN在训练过程中会丢失地物特征之间的空间关系,限制了其分类精度的进一步提高<sup>[18]</sup>。Sabour等<sup>[19]</sup>提出了一种新型的神经网络,称为胶囊网络(CapsNet)。CapsNet用向量神经元,即胶囊(capsule),代替CNN的标量神经元,并采用动态路由算法更新低层胶囊与高层胶囊之间的耦合系数,得到局部与整体的关系<sup>[20]</sup>。CapsNet在图像分割、目标识别等领域均表现了优异的性能,在高光谱遥感分类领域中精度优于CNN<sup>[21-27]</sup>。但在复杂数据集分类中,由于其鲁棒性受到限制,分类性能较差,原因主要有两个方面: CapsNet只使用浅层CNN来提取底层结构特征,忽略了高级语义特征;浅卷积层通过使用大量的卷积核来获得较大的感受野,大幅增加了可训练参数的数量,不但容易出现过拟合现象,而且极大地降低了训练效率。

为解决CapsNet面临的问题,进一步提高CapsNet的遥感分类精度,本文提出一种基于多尺度自适应胶囊网络(MSCaps)的高光谱遥感分类方法,利用多尺度(不同尺寸的输入图像)的卷积层来提取地物的空间信息与光谱信息。同时,为了进一步加深胶囊网络架构以提取更高层次的语义特征,避免耦合系数 $c_j$ 在训练过程中逐步稀疏化,继而引起过拟合现象,本文利用一种非迭代自适应路由算法进一步改进CapsNet的网络结构,以解决模型的过拟合问题,从而增加模型的适用性,提高分类的效果。为验证MSCaps的有效性,利用Pavia University(PU)和Salinas-A(SA)两个公共高光谱数据集分别对由不同材质构成的建筑物屋顶、处于不同生长周期的植被等复杂地类进行区分。

## 2 方法内容

### 2.1 胶囊网络

与传统CNN模型相比较,CapsNet利用向量化的神经元(胶囊)代替传统标量神经元,对分类目标的仿射变换具有更强的鲁棒性,所需的训练样本数量也更少。一个胶囊实际上是由一组标量神经元组成的,其方向和长度分别表示实体的属性和实体存在的概率。CapsNet由输入层、输出层、卷积层、主胶囊(PrimaryCaps)层和数字胶囊(DigitCaps)层组成。其

中卷积层用于提取地物目标的低层次特征, PrimaryCaps层用于表达特征之间的空间关系,并将提取的特征传递给 DigitCaps层,并利用动态路由算法进行预测。CapsNet的目标函数定义为

$$L_k = T_k \max(0, m^+ - \|\mathbf{v}_k\|)^2 + \lambda(1 - T_k) \max(0, \|\mathbf{v}_k\| - m^-)^2, \quad (1)$$

式中:  $T_k$  表示第  $k$  个目标标签;  $\|\mathbf{v}_k\|$  表示胶囊向量的模长;  $m^+ = 0.9$  和  $m^- = 0.1$  分别表示最大与最小边缘阈值;  $\lambda$  是用于阻止网络陷入局部最优的权重系数, 通常为 0.5。

动态路由算法(dynamic routing algorithm)根据低层( $l$ 层)胶囊层与高层( $l+1$ 层)胶囊层之间的相似性来调整耦合系数  $c$ , 并以此更新网络之间的权重  $W$ , 算法流程如图 1 所示。若  $l$  层的第  $i$  个胶囊与  $l+1$  层的第  $j$  个胶囊之间的相似性越大, 则耦合系数  $c_{ij}$  越大, 具体计算公式为

$$c_{ij} = \frac{\exp b_{ij}}{\sum_k \exp b_{ij}}, \quad (2)$$

式中:  $b_{ij}$  表示胶囊  $i$  和胶囊  $j$  耦合的先验概率, 且初始值为 0。其更新方式为

$$b_{ij} \leftarrow b_{ij} + \hat{\mathbf{u}}_{ji} \cdot \mathbf{v}_j, \quad (3)$$

式中:  $\hat{\mathbf{u}}_{ji}$  由胶囊  $\mathbf{u}_i$  经过仿射变换而获得;  $\mathbf{v}_j$  为胶囊  $j$  的输出值, 为保证其长度介于 0 至 1 之间, 使用非线性压缩函数压缩长度。

$$\hat{\mathbf{u}}_{ji} = W_{ij} \mathbf{u}_i, \quad (4)$$

$$\mathbf{v}_j = \frac{\|\mathbf{s}_j\|^2}{1 + \|\mathbf{s}_j\|^2} \times \frac{\mathbf{s}_j}{\|\mathbf{s}_j\|}, \quad (5)$$

式中:  $\mathbf{s}_j$  为  $l+1$  层第  $j$  个胶囊的输入向量。其计算公式为

$$\mathbf{s}_j = \sum_i c_{ij} \hat{\mathbf{u}}_{ji} \quad (6)$$

Algorithm 1: dynamic routing algorithm

```

1: procedure routing ( $\hat{\mathbf{u}}_{ji}, r, l$ )
2:   for all capsule  $i$  in layer  $l$  and capsule  $j$  in layer  $l+1$ :  $b_{ij} \leftarrow 0$ 
3:   for  $r$  iterations do
4:     for all capsule  $i$  in layer  $l$ :  $c \leftarrow \text{Softmax}(b_l)$ 
5:     for all capsule  $j$  in layer  $l+1$ :  $\mathbf{s}_j \leftarrow \sum_i c_{ij} \hat{\mathbf{u}}_{ji}$ 
6:     for all capsule  $j$  in layer  $l+1$ :  $\mathbf{v}_j \leftarrow \text{squash}(\mathbf{s}_j)$ 
7:     for all capsule  $i$  in layer  $l$  and capsule  $j$  in layer  $l+1$ :  $b_{ij} \leftarrow b_{ij} + \hat{\mathbf{u}}_{ji} \cdot \mathbf{v}_j$ 
8:   return  $\mathbf{v}_j$ 
    
```

图 1 动态路由算法的流程

Fig. 1 Flowchart of dynamic routing algorithm

## 2.2 多尺度胶囊网络

### 2.2.1 MSCaps 网络结构

所提 MSCaps 具有多分支结构, 每个分支对应一个特定大小(size)的输入图像, 并根据不同大小的感受野从输入图像中提取地物特征。MSCaps 的地物特征编码过程主要包含特征提取和编码两个阶段。第一阶段, 分别从不同尺寸的图像(对应 3 个分支)中提取地物的结构与语义特征, 其中第一、第二分支的前两个卷积层用于提取较高级别的地物特征, 第三分支的卷积层用于提取相对低级别的地物特征, 分别对提取的地物特征进行编码, 表达特征之间的空间关系, 形成不同大小的胶囊层, 并利用 Conca 层对胶囊层进行合并, 通过非迭代自适应路由算法将地物特征传递给更高层的胶囊层。在第二地物特征解码阶段, 采用 2 个全连接层进行图像重建。MSCaps 的具体结构(卷积层大小、步长、激活函数)如图 2 所示, 其中  $c$  表示类别的数量, FC 表示全连接层, ReLU 和 Sigmoid 表示激活函数, Conv 表示卷积层。

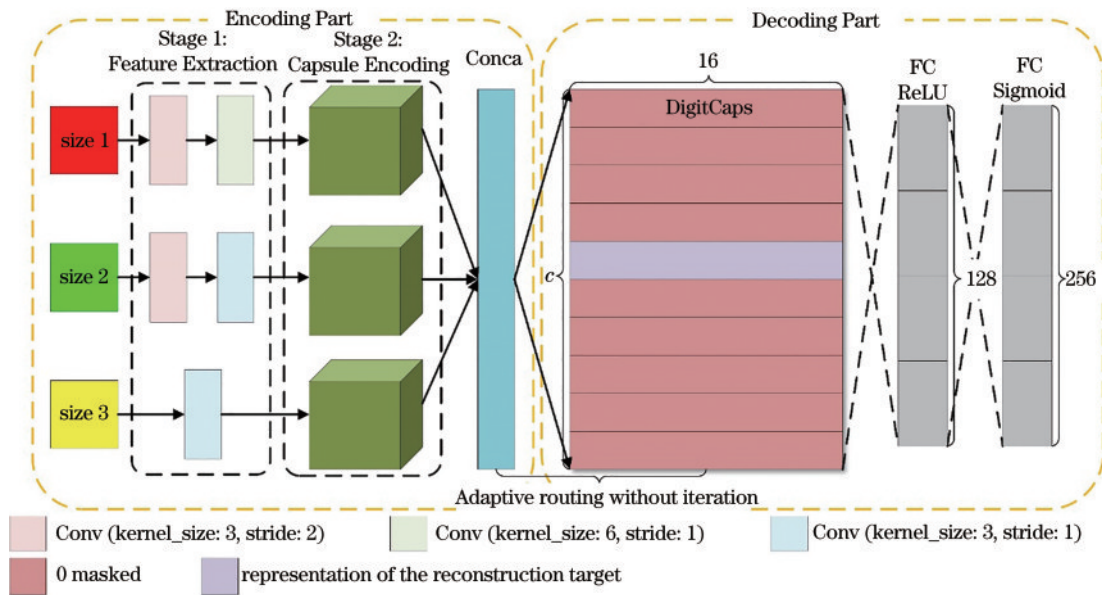


图 2 MSCaps 网络结构

Fig. 2 MSCaps architecture

2.2.2 非迭代自适应路由算法

在标准的 CapsNet 中,在算法迭代过程胶囊间的耦合系数  $c_{ij}$  会逐渐稀疏化,并引起过拟合现象。为避免这个问题,利用一种非迭代自适应路由算法 (adaptive routing algorithm without iteration) 来进一步改进 MSCaps 的结构,算法流程如图 3 所示。该算法引入超参数  $\gamma$  代替胶囊间的耦合系数,以增强不同网络层之间的梯度,避免后向传播过程中出现梯度消失现象。

Algorithm 2: adaptive routing algorithm without iteration

```

1: procedure routing (  $\hat{u}_{j|l}$ ,  $r$ ,  $l$  )
2:   capsule  $i$  in layer  $l$  and capsule  $j$  in layer  $l+1$ 
3:    $s_j \leftarrow \sum_i \hat{u}_{j|l}$ 
4:    $v_j \leftarrow \text{squash}(\gamma s_j)$ 
5:    $\hat{u}_{j|l} \leftarrow v_j + \hat{u}_{j|l}$ 
6:   return  $v_j$ 
7: end procedure
    
```

图 3 非迭代自适应路由算法流程

Fig. 3 Flowchart of adaptive routing algorithm without iteration

3 数据来源

实验采用的 PU 和 SA 两个公共高光谱数据集来源于高光谱数据集网站: [http://www.ehu.eu/ccwintco/index.php?title=Hyperspectral\\_Remote\\_Sensing\\_Scenes](http://www.ehu.eu/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes)。其中,PU 数据集是 2001 年德国宇航中心研制的 ROSIS 传感器获取的,观测区域为意大利北部帕维亚大学,影像大小为 610340 像素,光谱覆盖范围为 430~860 nm,分辨率为 1.3 m,去除受水汽吸收作用影响较大的波段后,剩余 103 个波段,如图 4(a)所示,对应的真实地物包含 9 种类型,如图 4(b)所示。SA 数据集是 1992 年美国 NASA 喷气推进实验室研制的 AVIRIS 传感器获取的,观测区域为美国加利福尼亚州某山谷,影像大小为 8683 像素,波长范围为 400~2500 nm,分辨率为 3.7 m,去除受水汽吸收作用影响较大的波段(第 108 个波段至第 112 个波段,第 154 个波段至第 167 个波段,第 224 个波段)后,剩余 204 个波段,如图 5(a)所示,对应的真实地物样本包含 6 种类型,如图 5(b)所示。用于 PU 和 SA 数据集分类实验的训练样本、验证样本及测试样本数量都是根据图 4(b)和图 5(b)所示的真实地物决定的,数量分别如表 1 和表 2 所示。

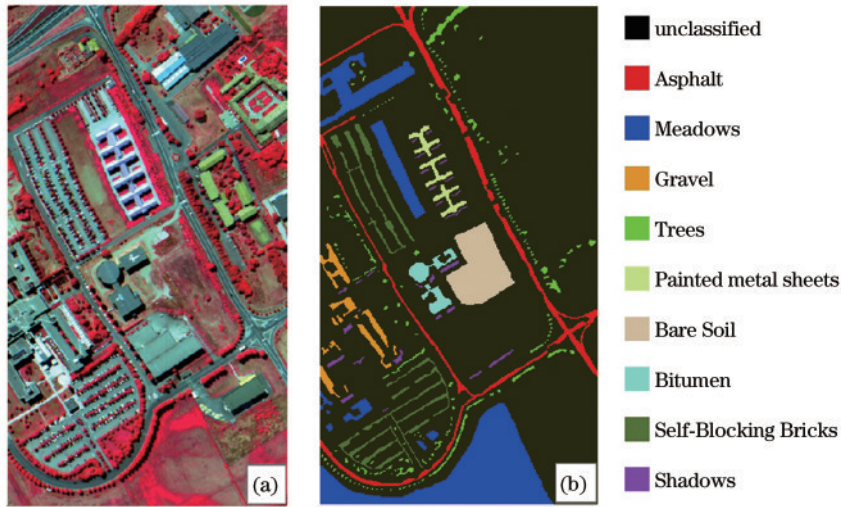


图 4 实验数据集 1。(a) PU 数据集;(b) PU 数据集真值

Fig. 4 Experimental dataset 1. (a) PU dataset; (b) ground truth of PU dataset

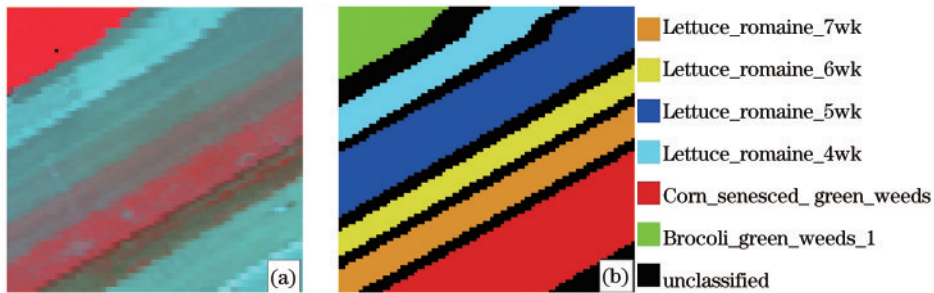


图 5 实验数据集 2。(a) SA 数据集;(b) SA 数据集真值

Fig. 5 Experimental dataset 2. (a) SA dataset; (b) ground truth of SA dataset

表 1 PU 数据集中各地物类型的训练样本、验证样本及测试样本数量

Table 1 Number of training, verification, and test samples of various types of objects on PU dataset

Class No.	Land cover	Training	Validation	Test
1	Asphalt	1000	1000	4631
2	Meadows	1000	1001	16650
3	Gravel	460	461	1180
4	Trees	890	891	1285
5	Painted metal sheets	400	401	546
6	Bare Soil	1000	1001	3030
7	Bitumen	400	401	531
8	Self-Blocking Bricks	1000	1001	1683
9	Shadows	260	261	428

表 2 SA 数据集中各地物类型的训练样本、验证样本及测试样本数量

Table 2 Number of training, verification, and test samples of various types of objects on SA dataset

Class No.	Land cover	Training	Validation	Test
1	Brocoli_green_weeds_1	100	100	191
2	Corn_senesced_green_weeds	390	390	563
3	Lettuce_romaine_4wk	150	150	316
4	Lettuce_romaine_5wk	470	470	585
5	Lettuce_romaine_6wk	210	210	254
6	Lettuce_romaine_7wk	250	250	299

## 4 实验结果和分析

### 4.1 深度学习网络架构

为了验证所提算法的有效性,利用PU和SA数据集对算法进行验证,并与具有相似网络架构的多尺度的卷积神经网络(MSCNN)进行对比。为了避免模型结构过深而出现梯度消失,影响分类性能,将MSCNN和MSCaps模型设计成相似的网络架构,即MSCNN的每个分支均设置两个卷积层用于提取特征,卷积核大小及卷积步长和MSCaps相同。而在MSCaps中,在前两个分支设置两个卷积层,第三个分支设置一个卷积层,卷积核大小及步长如图2所示,MSCNN和MSCaps的具体架构如图6所示。为进行有效对比,将CNN设置为依次包含一个采用ReLU激活函数的卷积层、批量标准化层(BN)、最大池化层(Pooling)、一个采用ReLU激活函数的卷积层的架构来进行特征提取,并在特征提取之后,采用展平操作(flatten)、丢弃操作(dropout)对提取的神经元进行处理,最后采用全连接层(dense layer)进行类别预测。CapsNet用于特征提取的网络结构与CNN类似。

输入图像的尺寸对模型的性能至关重要,为了得到MSCNN和MSCaps模型的最佳输入图像尺寸,采用不同尺寸的输入图像在PU和SA数据集上对CNN和CapsNet模型进行训练。将输入尺寸设置为 $5 \times 5$ 、 $7 \times 7$ 、 $11 \times 11$ 、 $15 \times 15$ 、 $19 \times 19$ 、 $23 \times 23$ 、 $27 \times 27$ 、 $31 \times 31$ 及 $35 \times 35$ ,算法总体分类精度随输入图像尺寸的变化情况如图7所示。如图7中虚线框所示,当输入

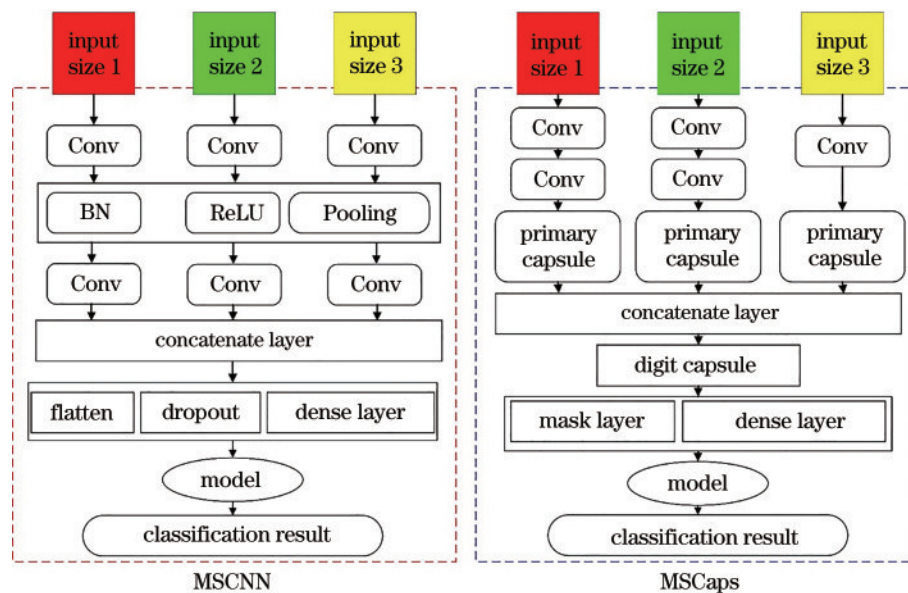


图 6 MSCNN 和 MSCaps 网络架构

Fig. 6 Network architectures of MSCNN and MSCaps

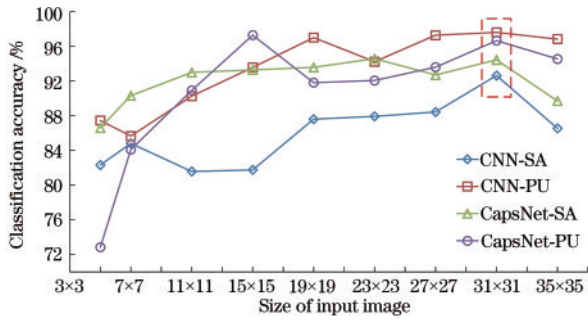


图 7 不同输入尺寸下的分类精度

Fig. 7 Classification accuracy under different input sizes

图像大小为  $31 \times 31$  时, CNN 与 CapsNet 在 SA 和 PU 数据集上的分类精度几乎同时达到最大。因此, 设置 MSCNN 和 MSCaps 第一分支的输入图像大小为  $31 \times 31$ 。从图 7 可以看出, 当输入图像大小介于  $11 \times 11$  至  $19 \times 19$ 、 $23 \times 23$  至  $27 \times 27$  时, CNN 和 CapsNet 的分类精度几乎都呈上升趋势, 同时为了保证输入图像在卷积后能与第一分支的输出进行连接, 设置 MSCNN 和 MSCaps 的第二与第三分支的输入图像大小分别为  $25 \times 25$  及  $13 \times 13$ 。

#### 4.2 实验设置

由于基于深度学习的遥感分类方法具有自动提取地物特征的优点, 因此采用基于 CNN、MSCNN、CapsNet 及 MSCaps 的遥感分类方法对具有 103 与 204 个波段的 PU 与 SA 数据集进行地物类型信息提

取。为进一步验证基于 MSCaps 的高光谱遥感方法较传统分类器 (SVM 与 RF) 的优势, 获取了 SVM 与 RF 对原始影像及经 PCA 特征提取后的影像的分类精度, 并与 CNN、MSCNN、CapsNet 与 MSCaps 的分类精度进行对比。为表述方便, SVM 与 RF 对经 PCA 特征提取后的影像进行地物类型信息提取的方法分别简称为 PCA-SVM 与 PCA-RF。同时, 为验证多尺度和非迭代自适应路由算法对 MSCaps 分类性能的贡献度, 将 MSCaps 与多尺度的胶囊网络 (MCaps) 以及基于非迭代自适应路由算法的胶囊网络 (ARWI-Caps) 作对比。

为了进一步验证 MSCaps 分类性能的优越性, 采用 McNemar's test 来评价显著性, 并设置置信水平为 0.05。本文所有深度学习算法均采用 Python 编程语言在 Keras 架构下实现, 算法运行 CPU 环境为 Intel Core Xeon(R) 2.1 GHz, 具有 128 G 内存空间。CNN、CapsNet、MCaps、ARWI-Caps、MSCNN 及 MSCaps 模型在训练完成后, 通过微调获得最佳参数配置。

#### 4.3 实验分类结果

##### 4.3.1 PU 数据集分类结果

在 PU 数据集上, SVM、RF、PCA-SVM、PCA-RF、CNN、CapsNet、MCaps、ARWI-Caps、MSCNN 以及 MSCaps 的总体分类精度分别为 91.60%、87.38%、91.86%、88.44%、95.84%、96.73%、98.22%、98.96%、98.71%、99.14%, 其中, MSCaps 的总体分类精度最高, ARWI-Caps 次之, RF 的总体分类精度最低。从图 8 可

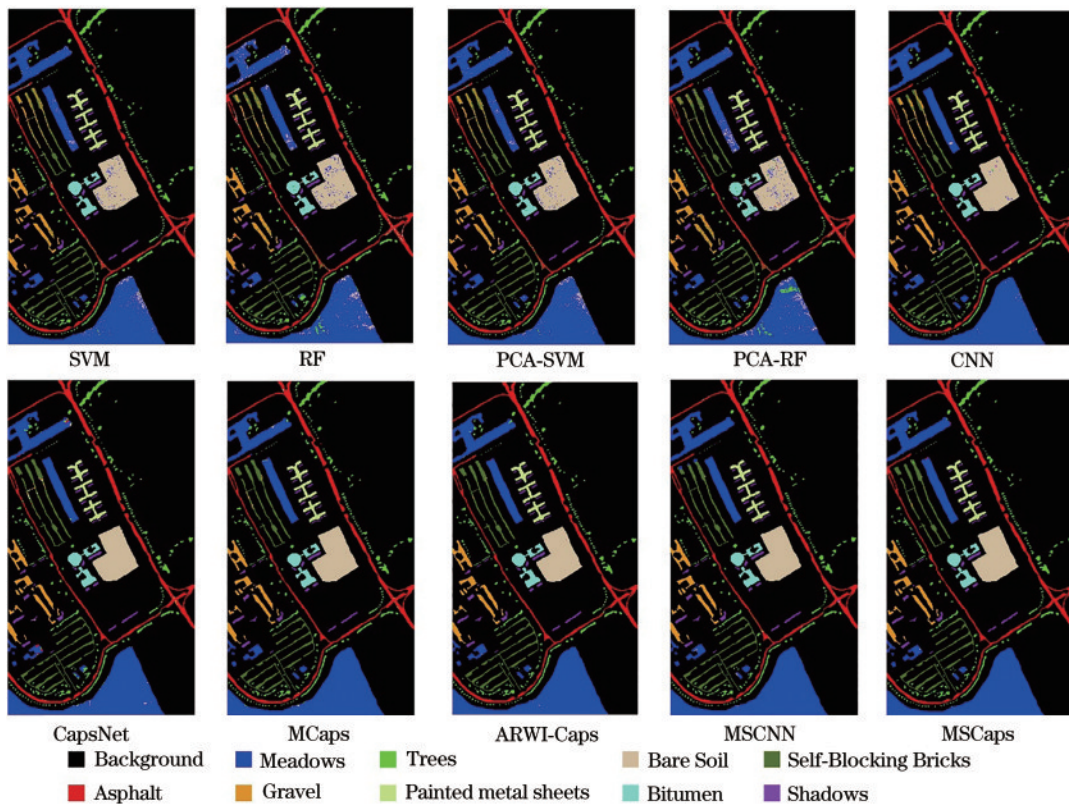


图 8 不同算法对 PU 数据集的分类结果

Fig. 8 Classification results of different algorithms on PU dataset

可以看出, SVM、RF、PCA-SVM、PCA-RF、CNN、CapsNet、MCaps、ARWI-Caps、MSCNN 及 MSCaps 算法都存在不同程度的错分现象, 其中 MSCaps 错分现象最少, 而传统的 SVM、RF、PCA-SVM、PCA-RF、CNN 错分现象较为严重, 存在大量的“椒盐”现象。大多数错分现象发生在土地利用类型较复杂的区域, 如 Meadows 误分为 Bare Soil, Self-Blocking Bricks 被误分为 Gravel 等。为进一步对比上述算法的分类性能, 表 3 列出了各算法对各地物类别的分类精度。相较于其他 5 种对比算法, MSCaps 对 Asphalt、Meadows、Bare Soil、Bitumen 和 Self-Blocking Bricks 5 类地物的分类精度最高, 尤其是对 Meadows、Bare Soil、Self-Blocking Bricks 及 Gravel 较难区分的地物, 精度分别达 99.97%、99.84%、99.62% 及 98.31%。对于 Bare Soil 地类, RF 的分类精度最低, 为 81.42%, 较 MSCaps 低 18.42 个百

分点。对于 Gravel 地类, SVM、RF、PCA-SVM、CNN 的分类精度分别为 76.32%、79.94%、76.83%、76.39%, 远低于 CapsNet、MCaps、ARWI-Caps、MSCNN 及 MSCaps; 而 MSCaps 的分类精度较 CapsNet 提高了 2.96 个百分点, 较 MCaps 提高了 2.12 个百分点, 与 ARWI-Caps、MSCNN 的分类精度接近。上述的实验结果表明: 基于 MSCaps 的高光谱遥感分类方法可以充分发挥高光谱数据的优势, 区分由不同材质组成的地物(如分别由 Gravel、Painted metal sheets 及 Bitumen 构成的建筑物屋顶)。为进一步验证所提分类算法的有效性, 在显著性水平为 0.05 的情况下, 利用 McNemar's test 检验所提算法分类性能优越性的置信水平。如表 3 所示, SVM、PCA-SVM、PCA-RF、CNN、CapsNet、MCaps、MSCNN 的  $p$ -value 值远小于 0.05, ARWI-Caps 的  $p$ -value 值为 0.011, 亦小于 0.05。

表 3 不同算法对 PU 数据集各地物类别的分类精度、总体分类精度、Kappa 系数( $K$ )及  $p$ -value

Table 3 Accuracy, overall accuracy, Kappa coefficient ( $K$ ), and  $p$ -value of different algorithms on PU dataset

No.	Accuracy / %									
	SVM	RF	PCA-SVM	PCA-RF	CNN	CapsNet	MCaps	ARWI-Caps	MSCNN	MSCaps
1	99.21	98.89	99.38	97.56	99.37	98.05	99.09	99.42	99.44	99.49
2	98.99	98.32	98.87	98.24	99.78	99.96	99.93	99.97	99.51	99.97
3	76.32	79.94	76.83	89.77	76.39	95.35	96.19	98.49	98.45	98.31
4	95.91	89.58	95.86	87.10	98.96	92.76	95.46	95.77	99.74	97.15
5	99.85	99.48	99.85	99.85	99.93	96.13	98.82	99.70	99.56	99.85
6	90.12	81.42	89.86	82.98	98.13	98.45	99.41	99.83	99.23	99.84
7	92.38	88.35	92.52	96.38	98.30	99.33	99.25	99.89	98.81	99.92
8	93.39	91.31	93.45	89.68	97.64	97.88	99.48	99.51	98.17	99.62
9	100	99.79	100	99.89	99.58	97.15	97.95	99.68	98.03	99.89
OA / %	91.60	87.38	91.86	88.44	95.84	96.73	98.22	98.96	98.71	99.14
$K$	0.82	0.74	0.84	0.76	0.91	0.93	0.96	0.97	0.97	0.99
$p$	$\ll 0.05$	0.0	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	0.011	$1.501 \times 10^{-5}$	

#### 4.3.2 SA 数据集分类结果

对 SA 数据集, SVM、RF、PCA-SVM、PCA-RF、CNN、CapsNet、MCaps、ARWI-Caps、MSCNN 及 MSCaps 的总体分类精度分别为 81.57%、73.41%、82.13%、82.21%、85.01%、87.27%、92.16%、95.10%、94.07% 及 95.38%, 其中 RF 的总体分类精度最低, MSCaps 的总体分类精度最高。从图 9 可以看出: 各算法均出现了不同程度的错分现象; 其中 SVM、RF、PCA-SVM、PCA-RF、CNN、CapsNet 的错分现象最为严重, 且存在“椒盐”现象。在所有的地物类型中, Corn\_senesced\_green\_weeds 的误分现象最为严重, 被 SVM、PCA-SVM、CNN、ARWI-Caps 及 MSCNN 大量地误分为 Lettuce\_romaine\_7wk, 分别被 RF、PCA-RF 大量地误分为 Lettuce\_romaine\_6wk 与 Lettuce\_romaine\_5wk, 被 CapsNet 大量地误分为 Lettuce\_romaine\_5wk 与 Lettuce\_romaine\_7wk, 被 MCaps 误分为 Lettuce\_romaine\_7wk 与 Lettuce\_romaine\_6wk, 所提算

法误分现象最少, 只有少量的 Corn\_senesced\_green\_weeds 被误分为 Lettuce\_romaine\_7wk。

为进一步对比各分类算法, 表 4 给出了 SVM、RF、PCA-SVM、PCA-RF、CNN、CapsNet、MCaps、ARWI-Caps、MSCNN 及 MSCaps 对各类地物的分类精度。由表 4 可知: 所提 MSCaps 获取的各类地物, 除了 Lettuce\_romaine\_7wk, 分类精度都高于其他对比算法; 尤其是对于误分现象最为严重的 Lettuce\_romaine\_6wk, MSCaps 获取的分类精度为 100%, 较其他 9 种算法分别提高了 12.47 个百分点、34.69 个百分点、13.14 个百分点、28.83 个百分点、6.52 个百分点、1.61 个百分点、10.01 个百分点、3.3 个百分点及 3.3 个百分点, 远高于 SVM、RF、PCA-SVM、PCA-RF 的分类精度。而对于 Lettuce\_romaine\_7wk, MSCaps 的分类精度接近 ARWI-Caps 获得的最高分类精度 (92.15%), 较 CapsNet 提高了 1.17 个百分点, 比 SVM 获得的最低分类精度高 16.18 个百分点。由此可见, 基于 MSCaps

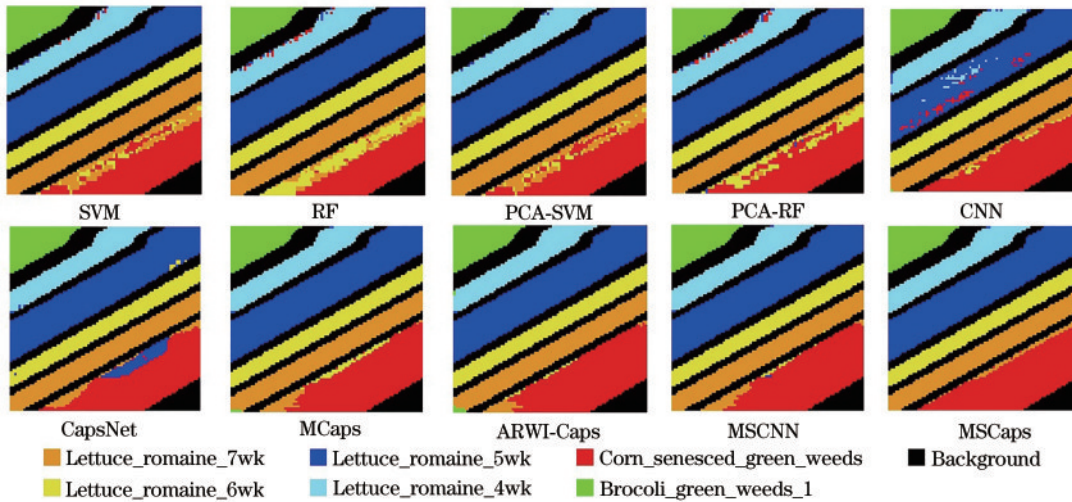


图 9 不同算法对 SA 数据集的分类结果

Fig. 9 Classification results of different algorithms on SA dataset

表 4 不同算法对 PU 数据集各地物类别的分类精度、总体分类精度、Kappa 系数(K)及 p-value

Table 4 Accuracy, overall accuracy, Kappa coefficient (K), and p-value of different algorithms on SA dataset

No.	Accuracy / %									
	SVM	RF	PCA-SVM	PCA-RF	CNN	CapsNet	MCaps	ARWI-Caps	MSCNN	MSCaps
1	100	100	100	100	99.74	100	98.73	96.54	100	100
2	100	97.08	99.26	96.99	93.26	100	100	100	100	100
3	99.84	99.83	99.88	100	93.92	100	100	100	100	100
4	99.35	99.09	99.35	98.58	99.08	90.55	99.74	99.87	99.54	99.87
5	87.53	65.31	86.86	71.17	93.48	98.39	89.99	96.70	96.70	100
6	72.70	80.75	73.64	92.26	84.36	87.71	89.92	92.15	88.96	88.88
OA/%	81.57	73.41	82.13	82.21	85.01	87.27	92.16	95.10	94.07	95.38
K	0.66	0.52	0.69	0.67	0.72	0.76	0.85	0.90	0.89	0.91
p	≪ 0.05	≪ 0.05	≪ 0.05	≪ 0.05	≪ 0.05	≪ 0.05	≪ 0.05	0.033	0.025	

的高光谱遥感分类方法可有效地区分处于不同生长期的植被(如 Lettuce\_roumaine)。同样地,在显著性水平为 0.05 的情况下,利用 McNemar's test 检验所提算法分类性能优越性的置信水平。如表 4 所示,SVM、RF、PCA-SVM、PCA-RF、CNN、CapsNet、MCaps 的 p-value 值远小于 0.05,而 ARWI-Caps 的 p-value 值为 0.033, MSCNN 的 p-value 值为 0.025,亦小于 0.05。

#### 4.4 训练时间比较

为了进一步验证所提 MSCaps 的有效性,利用模型训练时间比较了 CNN、CapsNet、MSCNN 及 MSCaps 的计算效率。为尽可能公平地进行对比,在保证网络结构尽可能相似的前提下,设置 CNN 及 CapsNet 的输入图像尺寸为 31×31,而 MSCNN 及 MSCaps 各分支的输入图像尺寸分别设置为 31×31、25×25 及 13×13。各模型的训练时间如图 10 所示。从图 10 可以看出:CNN 训练所耗费的时间最少,在 PU 和 SA 数据集上的训练时间分别约为 8.8×10<sup>3</sup> s 和

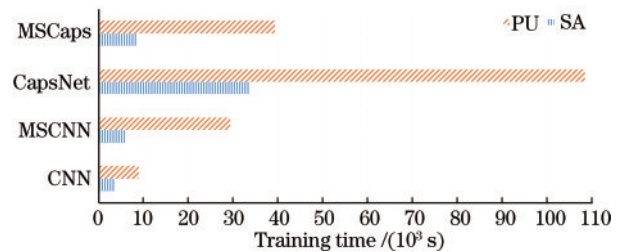


图 10 各模型在 PU 数据集和 SA 数据集上的训练时间

Fig. 10 Training time of each model on PU and SA datasets

3.5×10<sup>3</sup> s; CapsNet 训练所耗费的时间最多,在 PU 和 SA 数据集上的训练时间分别约为 108×10<sup>3</sup> s 和 33×10<sup>3</sup> s; 而所提 MSCaps 在 PU 和 SA 数据集上的训练耗时接近 MSCNN, 分别约为 39×10<sup>3</sup> s 与 8.4×10<sup>3</sup> s。MSCaps 在 PU 和 SA 数据集上的训练耗时分别约为 CNN 的 4 倍与 2 倍; MSCaps 训练效率较之 CapsNet 有大幅提高,在两种数据集上的耗时约为 CapsNet 的 1/3 与 1/4。



## 5 讨 论

针对 CapsNet 面临的问题,利用非迭代自适应路由算法,提出了 MSCaps 用于高光谱遥感分类,并利用 PU 与 SA 两个公共数据集对 MSCaps 的分类性能进行验证,并与 SVM、RF、PCA-SVM、PCA-RF、CNN、CapsNet、MCaps、ARWI-Caps 及 MSCNN 进行对比。实验结果表明,所提 MSCaps 对 PU 和 SA 的分类精度高于其他对比方法,且训练效率较 CapsNet 有了大幅提升。

首先,所提 MSCaps 可自动提取和学习地物的高层次抽象特征,因而分类性能要优于 SVM、RF、PCA-SVM 及 PCA-RF 等浅层分类器。其次,传统的 CNN、CapsNet 等深度学习模型因采用大量的卷积核提取地物特征,因而容易面临过拟合的问题。为解决上述问题,一方面,设计了具有多分支结构的多尺度卷积网络,每一个分支对应一个一定尺寸的输入图像,以尽可能贴合地物的实际大小,减少卷积过程中其他地物特征的干扰;另一方面,利用非迭代自适应路由算法对胶囊网络架构进行进一步改进,采用系数  $\gamma$  取代耦合系数  $c$ ,以避免耦合系数稀疏化引起的梯度消失,通过增强梯度信号来避免过拟合现象。因而基于所提 MSCaps 获得的高光谱遥感分类精度较 CNN、CapsNet、MCaps 及 ARWI-Caps 等更高。最后,利用的非迭代自适应路由算法避免了路由协议中的迭代过程,因而训练效率较 CapsNet、MCaps 等基于动态路由算法的胶囊网络模型有较大幅度提升。CNN 及 MSCNN 中的池化层会在特征提取阶段丢失大量的地物特征信息,且 CNN 与 MSCNN 无法表达地物特征之间的空间关系,因而 CNN 与 MSCNN 的分类精度较 MSCaps 更低。而在模型训练效率方面,CNN 与 MSCNN 结构因不具备路由协议,因此模型(尤其是 CNN 模型)的训练效率高于 CapsNet 和 MSCaps。

## 6 结 论

实验结果表明:所提 MSCaps 在高光谱遥感分类方面具有较好的应用潜力,可充分利用高光谱数据的优势区分复杂地物(如 PU 数据集中由 Painted metal sheets 与 Bitumen 构成的屋顶,SA 数据集中处于不同生长期的 Lettuce\_romaine)。然而所提方法仍存在不足:在卷积过程当中,在卷积窗口的边界区域仍然容易混入其他地物的特征。面向对象的遥感分类可有效解决上述问题,因此,下一步将进一步研究面向对象的多尺度胶囊网络以用于多光谱或高光谱遥感分类,以提高地物分类精度。

### 参 考 文 献

[1] Feddema J J, Oleson K W, Bonan G B, et al. The importance of land-cover change in simulating future climates[J]. *Science*, 2005, 310(5754): 1674-1678.

[2] 唐华俊, 吴文斌, 杨鹏, 等. 土地利用/土地覆被变化(LUCC)模型研究进展[J]. *地理学报*, 2009, 64(4): 456-468.  
Tang H J, Wu W B, Yang P, et al. Recent progresses of land use and land cover change(LUCC) models[J]. *Acta Geographica Sinica*, 2009, 64(4): 456-468.

[3] Ding X H, Zhang S Q, Li H P, et al. A restrictive polymorphic ant colony algorithm for the optimal band selection of hyperspectral remote sensing images[J]. *International Journal of Remote Sensing*, 2020, 41(3): 1093-1117.

[4] Hu Y B, Zhang J, Ma Y, et al. Hyperspectral coastal wetland classification based on a multiobject convolutional neural network model and decision fusion[J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 16(7): 1110-1114.

[5] 丁小辉, 李华朋, 张树清. 基于多态蚁群算法的高光谱遥感影像最优波段选择[J]. *遥感技术与应用*, 2016, 31(2): 275-284.  
Ding X H, Li H P, Zhang S Q. Optimized band selection of hyperspectral remote sensing image based on polymorphic ant colony algorithm[J]. *Remote Sensing Technology and Application*, 2016, 31(2): 275-284.

[6] 赵伍迪, 李山山, 李安, 等. 结合深度学习的高光谱与多源遥感数据融合分类[J]. *遥感学报*, 2021, 25(7): 1489-1502.  
Zhao W D, Li S S, Li A, et al. Deep fusion of hyperspectral images and multi-source remote sensing data for classification with convolutional neural network[J]. *National Remote Sensing Bulletin*, 2021, 25(7): 1489-1502.

[7] 唐璜, 刘正军, 杨懿, 等. 基于特征增强和 ELU 的神经网络建筑物提取研究[J]. *地球信息科学学报*, 2021, 23(4): 692-709.  
Tang Y, Liu Z J, Yang Y, et al. Research on building extraction based on neural network with feature enhancement and ELU activation function[J]. *Journal of Geo-Information Science*, 2021, 23(4): 692-709.

[8] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504-507.

[9] Li H P, Zhang C, Zhang S Q, et al. A hybrid OSVM-OCNN method for crop classification from fine spatial resolution remotely sensed imagery[J]. *Remote Sensing*, 2019, 11(20): 2370.

[10] Zhang C, Pan X, Li H P, et al. A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 140: 133-144.

[11] Othman E, Bazi Y, Alajlan N, et al. Using convolutional features and a sparse autoencoder for land-use scene classification[J]. *International Journal of Remote Sensing*, 2016, 37(10): 2149-2167.

[12] Sharma A, Liu X W, Yang X J, et al. A patch-based convolutional neural network for remote sensing image classification[J]. *Neural Networks*, 2017, 95: 19-28.

[13] 付松琳, 谢晨波, 李路, 等. 基于激光雷达探测技术的

- PM<sub>2.5</sub>浓度辨识研究[J]. 光学学报, 2021, 41(9): 0928001.
- Fu S L, Xie C B, Li L, et al. PM<sub>2.5</sub> concentration identification based on lidar detection[J]. Acta Optica Sinica, 2021, 41(9): 0928001.
- [14] 刘扬, 付征叶, 郑逢斌. 高分辨率遥感影像目标分类与识别研究进展[J]. 地球信息科学学报, 2015, 17(9): 1080-1091.
- Liu Y, Fu Z Y, Zheng F B. Review on high resolution remote sensing image classification and recognition[J]. Journal of Geo-Information Science, 2015, 17(9): 1080-1091.
- [15] 农元君, 王俊杰, 赵雪冰, 等. 遥感目标空间关系检测方法[J]. 光学学报, 2021, 41(16): 1628001.
- Nong Y J, Wang J J, Zhao X B, et al. Spatial relationship detection method of remote sensing objects [J]. Acta Optica Sinica, 2021, 41(16): 1628001.
- [16] 彭晓钰, 杨锦清, 吴春红, 等. 基于后向传播神经网络的激光定位系统动态范围的提升[J]. 光学学报, 2021, 41(6): 0620001.
- Peng X Y, Yang J Q, Wu C H, et al. Improvement of dynamic range of laser positioning system based on back propagation neural network[J]. Acta Optica Sinica, 2021, 41(6): 0620001.
- [17] 刘金香, 班伟, 陈宇, 等. 融合多维度CNN的高光谱遥感图像分类算法[J]. 中国激光, 2021, 48(16): 1610003.
- Liu J X, Ban W, Chen Y, et al. Multi-dimensional CNN fused algorithm for hyperspectral remote sensing image classification[J]. Chinese Journal of Lasers, 2021, 48(16): 1610003.
- [18] Xiang C Q, Zhang L, Tang Y, et al. MS-CapsNet: a novel multi-scale capsule network[J]. IEEE Signal Processing Letters, 2018, 25(12): 1850-1854.
- [19] Sabour S, Frosst N, Hinton G E. Dynamic routing between capsules[C]//Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA. New York: Curran Associates, 2017: 3856-3866.
- [20] Ren Q, Shang S H, He L H. Adaptive routing between capsules[EB/OL]. (2019-11-19) [2021-02-04]. <https://arxiv.org/abs/1911.08119>.
- [21] Nguyen C D T, Dao H H, Huynh M T, et al. ResCap: residual capsules network for medical image segmentation [C]//Submissions to the 2019 Kidney Tumor Segmentation Challenge, January 1, 2019, Shenzhen, China. Minneapolis: University of Minnesota Libraries Publishing, 2019.
- [22] Duarte K, Rawat Y, Shah M. VideoCapsuleNet: a simplified network for action detection[EB/OL]. (2018-05-21)[2021-02-04]. <https://arxiv.org/abs/1805.08162>.
- [23] Chen R L, Jalal M A, Mihaylova L, et al. Learning capsules for vehicle logo recognition[C]//2018 21st International Conference on Information Fusion (FUSION), July 10-13, 2018, Cambridge, UK. New York: IEEE Press, 2018: 565-572.
- [24] Beşer F, Kizrak M A, Bolat B, et al. Recognition of sign language using capsule networks[C]//2018 26th Signal Processing and Communications Applications Conference (SIU), May 2-5, 2018, Izmir, Turkey. New York: IEEE Press, 2018: 17914316.
- [25] Paoletti M E, Haut J M, Fernandez-Beltran R, et al. Capsule networks for hyperspectral image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(4): 2145-2160.
- [26] Wang W Y, Li H C, Pan L, et al. Hyperspectral image classification based on capsule network[C]//2018 IEEE International Geoscience and Remote Sensing Symposium, July 22-27, 2018, Valencia, Spain. New York: IEEE Press, 2018: 3571-3574.
- [27] Deng F, Pu S L, Chen X H, et al. Hyperspectral image classification with capsule network using limited training samples[J]. Sensors, 2018, 18(9): 3153.