

利用胶囊网络实现高光谱影像空谱联合分类

高奎亮, 余旭初, 张鹏强, 谭熊, 刘冰

引用本文:

高奎亮, 余旭初, 张鹏强, 谭熊, 刘冰. 利用胶囊网络实现高光谱影像空谱联合分类[J]. 武汉大学学报·信息科学版, 2022, 47(3): 428–437.

GAO Kuiliang, YU Xuchu, ZHANG Pengqiang, TAN Xiong, LIU Bing. Hyperspectral Image Spatial–Spectral Classification Using Capsule Network Based Method[J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(3): 428–437.

相似文章推荐 (请使用火狐或IE浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

利用动态上采样滤波深度网络进行多角度遥感影像超分辨率重建

Multi-angle Remote Sensing Images Super-Resolution Reconstruction Using Dynamic Upsampling Filter Deep Network

武汉大学学报·信息科学版. 2021, 46(11): 1716–1726 <https://doi.org/10.13203/j.whugis20200651>

基于深度学习的立体影像密集匹配方法综述

A Review of Dense Stereo Image Matching Methods Based on Deep Learning

武汉大学学报·信息科学版. 2021, 46(2): 193–202 <https://doi.org/10.13203/j.whugis20200620>

一种综合利用图像和光谱信息的物体真假模式识别方法

A Method of True and Fake Objects Pattern Recognition Integrating Image Information and Spectral Information

武汉大学学报·信息科学版. 2019, 44(8): 1174–1181 <https://doi.org/10.13203/j.whugis20190139>

一种利用两阶段学习模型的水下阵列定位方法

An Underwater Array Localization Method Using Two-Stage Learning Model

武汉大学学报·信息科学版. 2021, 46(12): 1889–1899 <https://doi.org/10.13203/j.whugis20210466>

一种基于改进双边滤波的鲁棒高光谱遥感图像特征提取方法

Robust Hyperspectral Image Feature Extraction Based on Improved Bilateral Filtering

武汉大学学报·信息科学版. 2020, 45(4): 504–510 <https://doi.org/10.13203/j.whugis20180267>



利用胶囊网络实现高光谱影像空谱联合分类

高奎亮¹ 余旭初¹ 张鹏强¹ 谭熊¹ 刘冰¹

¹ 信息工程大学,河南 郑州,450001

摘要:卷积神经网络等深度学习模型已经在高光谱影像分类任务中取得了理想的结果。然而,由于传统神经元只能进行标量计算,现有的深度学习模型无法对高光谱影像特征的实例化参数进行建模,因此无法在邻域范围受限的条件下获得令人满意的分类效果。通过引入胶囊网络结构设计了一种新型网络模型,该模型利用胶囊神经元进行向量计算,并利用权重矩阵编码特征间的空间关系,能够进一步提高高光谱影像的分类精度。在帕维亚大学、印第安纳松树林和萨利纳斯山谷数据集上进行验证,实验结果表明,所提出的网络模型较传统算法和卷积神经网络分类模型而言具有更加优异的分类性能,且对训练样本数量和像素邻域范围具有更好的适应性。

关键词:高光谱影像分类;胶囊网络;胶囊神经元;卷积层;深度学习

中图分类号:P237

文献标志码:A

高光谱影像同时包含了丰富的空间和光谱信息,在矿岩探测、精细农业、军事侦察等领域发挥着重要作用。高光谱影像分类是高光谱影像处理与分析中的关键技术之一,其目的是为每个影像像素确定唯一的地物类别。逻辑回归、随机森林和最近邻分类器等传统算法最先被应用于高光谱影像分类。然而,受高光谱影像光谱冗余度高、标记样本少和空间变异性大等问题的制约^[1],上述方法难以取得令人满意的分类效果。

近年来,随着深度学习的兴起和不断发展,栈式自编码器^[2-3]、深度置信网络^[4-5]和循环神经网络^[6-7]等深度模型被应用于高光谱影像分类,并取得了一定效果。然而,上述方法均无法直接处理呈二维的图像结构,均需要通过一定的预处理步骤将高光谱数据转换成一维向量作为模型输入。高光谱影像具有典型的三维立方体结构,这种做法显然丢失了影像中重要的空谱联合信息。

卷积神经网络(convolutional neural network, CNN)利用其独特的卷积运算能够直接处理高维的二维图像数据,具有良好的特征提取能力。因此,利用CNN进行高光谱影像分类已经成为近年来的研究热点。目前,基于CNN的高光谱影像分类方法可分为3种:基于一维(1 dimensional, 1D)-CNN的分类方法^[8-10]、基于二维

(2 dimensional, 2D)-CNN的分类方法^[11-13]和基于三维(3 dimensional, 3D)-CNN的分类方法^[1, 14-16]。文献[8]率先使用1D-CNN进行高光谱影像分类,在训练样本充足的条件下,分类精度高于传统支持向量机(support vector machine, SVM)方法。文献[14]将像素邻域数据作为2D-CNN的输入,影像分类精度进一步提高。文献[1]分别构造了1D-CNN、2D-CNN和3D-CNN 3种分类模型,并对最佳网络结构、卷积核数量和分类效果进行了详细的研究。除此之外,CNN还与迁移学习^[17]、主动学习^[18]、半监督分类^[19]等其他方法相结合,在一定条件下均能获得良好的分类效果。

众所周知,CNN在诸多计算机视觉任务中已经取得了优异表现,但其仍然存在固有缺陷。CNN中,神经元的输入输出均为标量,特征表达能力受到限制。同时,CNN依靠池化操作实现一定程度的平移不变性。然而,这种机制使得CNN模型无法对特征对象的实例化参数进行建模,只能以标量输出的大小表示特征存在的概率而无法对特征之间的空间关系进行建模^[20]。利用CNN进行高光谱影像分类时,提高精度的常用方法可概括为3点:(1)扩充训练数据规模。例如,文献[1]通过对原数据添加噪声干扰的方式生成

虚拟训练样本,文献[10]和文献[21]均通过构造训练样本对的方式增加训练数据量。这种方法能够使 CNN 模型得到更为充分的训练,但不能解决其固有缺陷。(2)增大像素邻域范围。大量实验表明^[1, 12],适当扩大中心像素的邻域范围能够有效利用影像中的空间信息,提高分类精度。然而,CNN 在小邻域条件下的分类精度并不理想。(3)调整网络结构,不断提高网络的广度和深度,从而提高模型的抽象拟合能力^[16, 22-23]。然而,这种方法会产生不必要的参数冗余和训练开销,同时使网络难以训练,容易产生过拟合现象。

仅从扩充训练数据规模和增大像素邻域的角度改善分类效果存在局限性,不断加深 CNN 结构也不能改变其固有缺陷。高光谱影像具有“图谱合一”的优势特点,如何在训练样本受限和小邻域条件下充分利用高光谱影像中的空间和光谱信息是进一步提高分类精度的关键。不同于 CNN 的标量输入输出,胶囊网络^[24]中的每一个胶囊神经元均为向量输入输出,这极大地提高了神经元的特征表达能力,使胶囊网络能够对特征对象的实例化参数进行充分建模。因此,胶囊网络能够在更高层次上对高光谱影像中的空谱联合信息进行更为抽象的表达。

本文以胶囊网络为基础,设计了一种适用于高光谱影像分类的新型网络模型。该模型由传统卷积层、胶囊层和全连接层组成,具有更强的特征表达能力。实验结果表明,本文网络模型的性能优于传统算法和 CNN 分类模型,且在进一步减小训练样本数量和像素邻域的条件下体现出更好的适应性。

1 胶囊网络与本文模型

1.1 卷积层

卷积层是 CNN 中的核心部分,它能够直接对输入图像进行卷积运算,并逐层提取深度特征。为提取输入图像中的多种特征,一个卷积层内通常包括多个卷积核,从而得到多个输出特征图。二维卷积运算的表达式为:

$$v_{l,j}^{x,y} = f(b_{l,j} + \sum_m \sum_{p=0}^{P_l-1} \sum_{q=0}^{Q_l-1} w_{l,j,m}^{p,q} v_{(l-1),m}^{(x+p)(y+q)}) \quad (1)$$

式中, $v_{l,j}^{x,y}$ 表示第 l 层第 j 个特征图 (x,y) 位置处神经元的值; $b_{l,j}$ 表示第 l 层第 j 个特征图的偏置; $w_{l,j,m}^{p,q}$ 表示第 l 层第 j 个卷积核在位置 (p,q) 处的权值; m 为 $l-1$ 层中特征图的个数; P_l 和 Q_l 分别表示卷积核的高度和宽度; $f(\cdot)$ 为非线性激活函数。

1.2 胶囊网络

1.2.1 信息传递与动态路由

胶囊是一个包含丰富信息的向量神经元,由多个传统神经元组合构成,其活动向量的模长代表某特定实体的存在概率,方向则负责编码该实体的实例化参数。相邻两个胶囊层之间的信息传递过程包括线性组合和动态路由两个阶段。如图 1 所示,位于高层级的胶囊 j 与 3 个低层级胶囊相连,每个低层级胶囊的输出向量分别记为 u_1 、 u_2 和 u_3 。

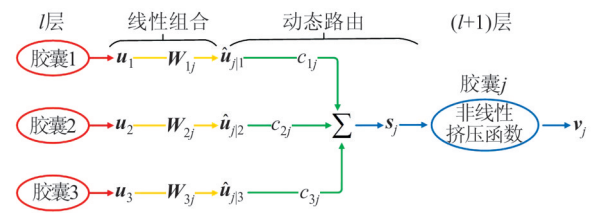


图 1 相邻胶囊层间的信息传递过程

Fig.1 Information Transfer Between Adjacent Capsule Layers

为了编码低层特征与高层特征之间的抽象空间关系,低层胶囊将自身输出向量乘以对应的权重矩阵 W_{ij} (i 和 j 分别索引高层次和低层次胶囊),从而得到预测向量,计算公式为:

$$\hat{u}_{ji} = W_{ij} u_i, \quad i = 1, 2, 3 \quad (2)$$

将每个低层胶囊的预测向量乘以相应的耦合系数 c_{ij} 并进行求和计算,得到高层胶囊的输入向量 s_j ,计算公式为:

$$s_j = \sum_i c_{ij} \hat{u}_{ji}, \quad i = 1, 2, 3 \quad (3)$$

非线性挤压函数能够将输入向量 s_j 的模长压缩到 0~1 之间而保持向量方向不变,使输出向量 v_j 的模长代表特定对象的存在概率。其计算公式为:

$$v_j = \frac{\|s_j\|^2 s_j}{1 + \|s_j\|^2} \quad (4)$$

耦合系数 c_{ij} 决定了低层胶囊与高层胶囊之间信息传递的权重,其取值大小由动态路由算法计算得到。动态路由算法包括了 softmax 归一化、向量加权求和、向量压缩和更新耦合系数 4 个部分。相邻两胶囊层之间的先验概率系数 b_{ij} 最初被初始化为 0,因此在第一次迭代时经 softmax 计算得到的耦合系数均保持相同。接下来的迭代过程中,系数 b_{ij} 利用低层胶囊的预测向量 \hat{u}_{ji} 和高层胶囊的输出向量 v_j 的内积进行更新,计算公式为:

$$b_{ij} \leftarrow b_{ij} + \hat{u}_{ji} \cdot v_j \quad (5)$$

可以看出,当 \hat{u}_{ji} 与 v_j 的内积越大时,两胶囊之间的耦合系数越大,即低层胶囊更倾向于将特征信息发送给该高层胶囊,从而实现信息路由的目的。

1.2.2 损失函数

胶囊网络的设计原理赋予了其能够同时识别多个实体对象的能力。因此,胶囊网络不再使用传统的交叉熵损失函数,而是使用边缘损失函数 L_c 作为网络优化的目标函数:

$$L_c = T_c \max(0, m^+ - \|v_c\|)^2 + \lambda(1 - T_c) \max(0, \|v_c\| - m^-)^2 \quad (6)$$

式中, c 表示类别; T_c 为关于分类的函数,当第 c 类对象出现时, T_c 为1,否则 T_c 为0; m^+ 为上边缘阈值,通常取0.9,用于惩罚预测存在而实际不存在的错误情况; m^- 为下边缘阈值,通常取0.1,用于

惩罚实际存在但预测不存在的错误情况; λ 为稀疏系数,负责调整二者比重,取值通常为0.5。

需要说明的是,式(6)仅是一个胶囊神经元的损失函数,胶囊层的整体损失应为每个胶囊的损失之和。为了起到正则化效果,胶囊网络增加了重构损失,即在胶囊层之后添加一个全连接网络,并计算全连接网络的输出与原始输入之间差的平方和。胶囊网络总损失 L_{total} 的计算公式为:

$$L_{total} = \sum_c L_c + \alpha \|x - x'\|^2$$

式中, x 为网络的输入数据; x' 为网络的重构表达; α 为权重因子。

1.3 本文网络模型

图2展示了本文设计的胶囊网络模型(P-CapsNet)。整个网络模型可分为分类和重构两个部分,共包括1个传统卷积层、1个主胶囊层、1个数字胶囊层和3个全连接层。

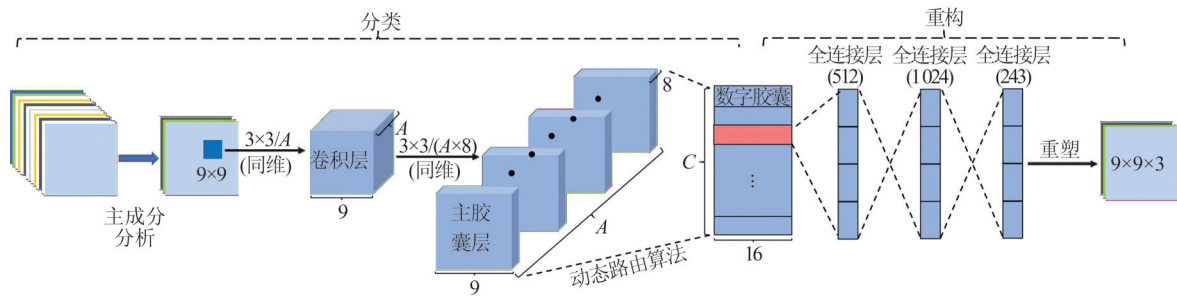


图2 本文设计的网络模型

Fig.2 Network Model of This Paper

高光谱影像通常包括上百个连续的波段,这就存在光谱信息冗余的问题。参照相关的实验经验,首先利用主成分分析算法将高光谱影像压缩至3维,以保留其主要的光谱特征。参照文献[16],选取中心像素周围 9×9 大小的邻域数据作为网络输入。在传统卷积层中,使用 A 个 3×3 像素大小的二维卷积核对输入数据进行卷积运算,以提取影像中的局部特征,分别将主胶囊层和数字胶囊层中的胶囊神经元数量设置为8个和16个(参数设置见§2.2)。主胶囊层的实现实际上为一组卷积运算,卷积核数量设置为 $A \times 8$,即对卷积层的输出特征图进行了 $A \times 8$ 次卷积运算,并将卷积结果封装成 $A \times 9 \times 9$ 个长度为8像素的胶囊向量。数字胶囊层中共包括 C 个长度为16像素的胶囊神经元,其中 C 为类别数。两个胶囊层之间采用动态路由算法进行信息传递。需要说明的是,数字胶囊层中的胶囊向量的模长代表了输入数据属于某一类别的概率,即模长最大的胶囊向量对应网络的预测类别。除此之外,选取数

字胶囊层中模长最大的胶囊向量进行影像重构,即将其依次连接3个全连接层,并将最后一个全连接层的输出结果重塑为大小为 $9 \times 9 \times 3$ 像素的影像数据(与输入影像一致)。按照式(6)对重构影像和输入影像之间的偏差进行计算,以进一步增强胶囊网络的抽象建模能力。卷积层、胶囊层和前两个全连接层均采用纠正线性单元(ReLU)函数进行非线性激活,以提高网络的收敛速度,最后一个全连接层采用Sigmoid函数进行激活。

2 高光谱影像分类实验结果与分析

实验的运行环境为16 GB内存,i7-9750H处理器,RTX2070显卡。实验所有程序均基于Python语言和深度学习库Tensorflow开发实现。

2.1 实验数据

选择帕维亚大学、印第安纳松树林、萨利纳斯山谷3个常用高光谱数据集作为实验数据,并分别将其划分为训练集、确认集和测试集。

1)帕维亚大学。该影像由反射光学系统成像光谱仪(reflective optics system imaging spectrometer, ROSIS)传感器拍摄获得,影像大小为 610×340 像素,光谱范围为 $0.43 \sim 0.86 \mu\text{m}$,像素分辨率为 1.3 m ,共包括9个地物类别和103个波段,具体情况如表1所示。

表1 帕维亚大学数据集划分情况/块
Tab.1 Division Situation of Pavia University Dataset/Patch

类别	数据量		
	训练集	确认集	测试集
柏油路	180	20	6 631
草地	180	20	18 649
沙砾	180	20	2 099
树木	180	20	3 064
金属板	180	20	1 345
裸土	180	20	5 029
沥青	180	20	1 330
砖体	180	20	3 682
阴影	180	20	947
总计	1 620	180	42 776

2)印第安纳松树林。该影像由航空可见红外成像光谱仪(airborne visible infrared imaging spectrometer, AVIRIS)传感器获得,大小为 145×145 像素,光谱范围为 $0.40 \sim 2.50 \mu\text{m}$,像素分辨率为 20 m ,共包括200个波段和16个地物类别。本文中仅选取了9个标注样本数量较多的地物类别进行实验,具体情况如表2所示。

表2 印第安纳松树林数据集划分情况/块
Tab.2 Division Situation of Indian Pines Dataset/Patch

类别	数据量		
	训练集	确认集	测试集
免耕玉米	180	20	1 428
少耕玉米	180	20	830
草-牧场	180	20	483
草-树	180	20	730
堆积干草	180	20	478
免耕大豆	180	20	972
少耕大豆	180	20	2 455
大豆	180	20	593
树木	180	20	1 265
总计	1 620	180	9 234

3)萨利纳斯山谷。该影像由AVIRIS拍摄获得,大小为 512×217 像素,光谱范围为 $0.40 \sim 2.50 \mu\text{m}$,像素分辨率为 3.7 m ,共包括16个地物类别和204个波段,具体情况如表3所示。

表3 萨利纳斯山谷数据集划分情况/块
Tab.3 Division Situation of Salinas Dataset/Patch

类别	数据量		
	训练集	确认集	测试集
椰菜_绿_野草_1	180	20	2 009
椰菜_绿_野草_2	180	20	3 726
休耕地	180	20	1 976
粗糙的休耕地	180	20	1 394
平滑的休耕地	180	20	2 678
残株	180	20	3 959
芹菜	180	20	3 579
未结果实的葡萄	180	20	11 271
正在开发的葡萄园土壤	180	20	6 203
开始衰老的玉米	180	20	3 278
长叶莴苣(4周)	180	20	1 068
长叶莴苣(5周)	180	20	1 927
长叶莴苣(6周)	180	20	916
长叶莴苣(7周)	180	20	1 070
未结果实的葡萄园	180	20	7 268
葡萄园小路	180	20	1 807
总计	2 880	320	54 129

2.2 实验结果与分析

卷积核数量对神经网络的性能有着重要影响。如果卷积核数量过少,模型无法提取出足够的特征;如果卷积核数量过多,会导致网络参数冗余。因此,针对3组高光谱数据集,分别采用试错法确定P-CapsNet的最佳卷积核数量,结果如表4所示。帕维亚大学和印第安纳松树林数据集的最佳卷积核数量分别为48和40;对于萨利纳斯山谷数据集而言,当卷积核数量设置为32和40时,模型均具有较好的分类性能,但考虑到训练成本,最终将卷积核数量设置为32。

表4 不同卷积核数量对应的总体分类精度/%
Tab.4 Overall Classification Accuracy Corresponding to Different Numbers of Convolution Kernels/%

数据集	卷积层中卷积核数量							
	16	24	32	40	48	56	64	
帕维亚大学	93.76	93.47	93.54	93.74	95.79	94.81	93.43	
印第安纳松树林	94.03	94.19	94.40	95.51	94.12	94.37	93.84	
萨利纳斯山谷	94.81	95.02	96.03	96.04	95.58	94.78	94.33	

胶囊神经元的大小直接影响着模型的实例化参数建模能力,表5给出了3组数据集的总体分类精度与胶囊神经元大小之间的关系。参照文献[20]和文献[24]中的实验经验,数字胶囊层神经元大小设置为主胶囊层神经元大小的2倍。表5中,2/4表示主胶囊层神经元大小为2,数字胶囊层神经元大小为4。可以发现,随着胶囊神经元增大,分类精度先上升后下降。这表明胶囊

神经元过大或者过小都会影响模型的性能。

表5 不同胶囊神经元大小对应的总体分类精度/%
Tab.5 Overall Classification Accuracy Corresponding to Different Sizes of Capsule Neurons/%

数据集	胶囊神经元大小			
	2/4	4/8	8/16	16/32
帕维亚大学	93.73	94.67	95.79	95.02
印第安纳松树林	94.34	95.33	95.51	95.18
萨利纳斯山谷	95.18	95.67	96.03	95.95

相邻胶囊层之间利用动态路由算法进行信息传递,因此需要确定路由算法的最佳迭代次数。图3和图4分别展示了在不同数据集上分类精度和训练时间随路由迭代次数的变化。可以发现,训练时间随路由迭代次数的增加有明显增长,这说明动态路由算法在训练过程中产生了较大的资源开销。综合图3、图4结果,最终将帕维亚大学、印第安纳松树林和萨利纳斯山谷数据集上的路由迭代次数分别设置为2、1、1。

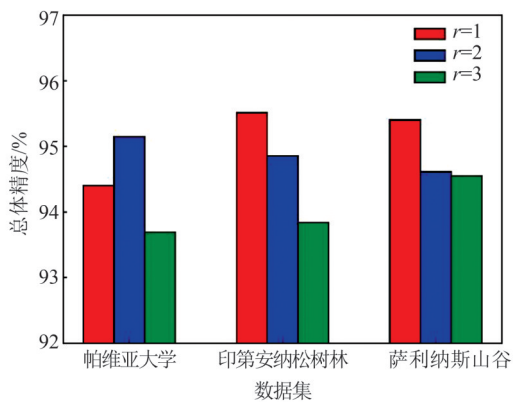


图3 总体分类精度与r的关系

Fig.3 Relationship Between Overall Accuracy and r

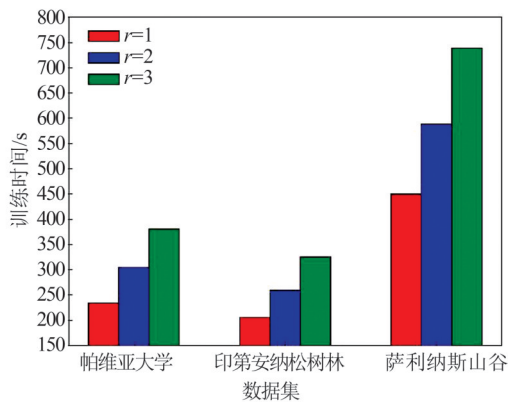


图4 各数据集训练时间与r的关系

Fig.4 Relationship Between Training Time and r

实验中,采用Xavier方法^[25]对卷积模板进行初始化,并采用Adam算法作为优化算法。参照

现有CNN模型设计的一般经验,设置学习率为0.001,训练迭代次数为100次。P-CapsNet在训练集上进行参数优化,选择在确认集上分类精度最高的网络模型作为最终训练结果,并利用测试集评估模型的性能。为了对模型进行定量评价,选择总体精度(overall accuracy, OA)、平均精度(average accuracy, AA)和Kappa系数作为评价指标。

为了验证P-CapsNet的有效性,将分类结果分别与SVM、扩展属性学轮廓特征(extended morphological profiles, EMP)^[26]、2D-CNN^[11]、3D-CNN^[16]和构建了11个卷积层的深度CNN模型^[22]进行比较。为了公平比较,所有分类算法的训练集、确认集和测试集数据量均相同。SVM以径向基函数作为核函数,并采用网格搜索法确定参数C和 γ ,搜索范围均为 $(2^{-2}, 2^{-1} \dots 2^5)$ 。2D-CNN、3D-CNN和深度CNN模型均选取 9×9 的空间邻域作为输入数据,以达到限制像素邻域大小的目的。需要说明的是,文献[22]中的深度CNN模型选取了每类200个样本用于模型训练,而本文只选取了每类180个样本,因此分类精度略低于原文献。

表6~8给出了不同分类算法的分类结果。SVM仅利用光谱信息进行分类,因此分类精度最低。EMP和2D-CNN均能够利用高光谱影像中的空间信息,因此其分类效果较SVM有一定改善。3D-CNN利用三维卷积核进行三维卷积运算,能够更好地利用影像中的空谱联合信息,因此其OA、AA和Kappa系数较EMP和2D-CNN均有提升。深度CNN模型通过构建深度网络能够利用更为抽象的深度特征,因此其OA、AA和Kappa系数进一步提高。需要强调的是,由于实验中将像素邻域的空间大小限制为 9×9 ,3种CNN模型的性能并不高。由此可以看出,传统CNN分类模型在像素邻域范围受限难以取得较高的分类精度。P-CapsNet在帕维亚大学、印第安纳松树林、萨利纳斯山谷这3个数据集上均表现出了最好的分类性能。对比3D-CNN分类模型,P-CapsNet在3个数据集上的OA、AA和Kappa系数均有明显提高,其中OA分别提高了1.83%、4.28%和2.14%。这说明较传统CNN分类模型而言,P-CapsNet能够从小邻域范围内学习到更具判别性的抽象特征,因此能够取得更高的分类精度。

图5~7展示了不同分类算法在3组数据集上

的分类图。SVM 的分类图中包含了大量的噪声点。2D-CNN 的分类图中噪声现象有所减弱,但其分类效果受像素邻域影响较大,因此在类边界处和小地物区域内仍然出现了大量的错分现象(红框标注)。3D-CNN 和深度 CNN 模型分类图中错分现象进一步减少。从类一致性的角度分析,P-CapsNet 的分类图具有最好的视觉效果。可以看出,其在类边界处和小地物区域等复杂环境下的错分现象最少,因此具有最高的分类精度。

表 6 不同算法在帕维亚大学数据集的分类精度/%
Tab.6 Classification Accuracy of Different Algorithms on Pavia University Dataset/%

类别序号	SVM	EMP	2D-CNN	3D-CNN	CNN	P-CapsNet
1	97.34	86.17	88.80	92.08	97.18	99.07
2	97.32	91.57	90.33	92.02	97.08	97.12
3	77.25	88.61	91.19	86.57	89.68	90.04
4	84.73	95.07	99.71	99.64	94.43	99.80
5	98.09	99.03	100.00	99.78	99.78	100.00
6	71.47	94.35	92.22	96.48	86.67	82.02
7	61.49	95.79	95.56	98.95	88.75	90.01
8	85.60	83.24	95.17	97.77	97.08	96.57
9	99.89	99.89	99.58	100.00	93.00	100.00
OA	88.59	91.00	92.11	93.96	94.95	95.79
AA	85.91	92.64	94.73	95.92	93.74	94.96
Kappa	85.22	88.24	0.897 2	92.12	0.933 3	93.49

表 7 不同算法在印第安纳松树林数据集的分类精度/%
Tab.7 Classification Accuracy of Different Algorithms on Indian Pines Dataset/%

类别序号	SVM	EMP	2D-CNN	3D-CNN	CNN	P-CapsNet
1	75.78	69.75	83.75	81.39	88.13	92.77
2	71.12	87.71	96.14	93.25	88.37	91.15
3	89.10	97.93	98.96	96.69	93.08	97.97
4	96.15	98.77	98.77	97.26	98.91	99.05
5	99.79	99.79	98.74	100.00	99.38	100.00
6	69.98	91.87	93.93	91.05	89.10	96.76
7	89.22	82.57	77.72	85.74	96.05	96.63
8	79.79	90.56	98.65	96.29	87.24	84.47
9	99.68	95.10	98.02	99.92	98.70	99.68
OA	84.57	87.23	90.00	91.23	92.81	95.51
AA	85.62	90.45	93.86	93.57	93.22	95.39
Kappa	82.09	85.11	88.39	89.76	91.59	94.74

深度学习模型的训练需要消耗大量的时间。表 9 给出了 3D-CNN、深度 CNN 和 P-CapsNet 在 3 组数据集上的训练和测试时间。三维卷积操作

大大增加了 3D-CNN 的时间开销,深度 CNN 因构建了 11 个卷积层的深度模型,训练时间也较长。P-CapsNet 在 3 组数据集上的训练和测试时间均为最少,表明了其在分类效率上的优越性。

表 8 不同算法在萨利纳斯山谷数据集的分类精度/%
Tab.8 Classification Accuracy of Different Algorithms on Salinas Dataset/%

类别序号	SVM	EMP	2D-CNN	3D-CNN	CNN	P-CapsNet
1	100	99.85	98.11	100.00	88.58	100.00
2	98.88	99.49	99.54	98.74	99.97	100.00
3	91.67	100.00	98.03	99.44	99.90	99.25
4	97.88	99.78	100.00	100.00	98.86	98.72
5	99.37	97.46	98.21	99.78	99.00	99.48
6	99.82	99.70	100.00	100.00	99.97	99.80
7	99.03	98.91	97.85	99.80	99.89	99.69
8	79.12	81.60	87.82	85.42	88.75	92.09
9	98.75	98.03	96.70	98.97	99.77	99.57
10	85.48	97.04	96.28	97.93	96.28	96.93
11	79.21	98.78	98.03	100.00	96.13	96.04
12	97.13	100.00	99.90	100.00	100.00	99.07
13	95.03	99.02	100.00	100.00	99.46	99.78
14	94.12	99.72	99.72	100.00	99.53	98.08
15	68.15	85.90	79.43	79.94	85.96	82.50
16	98.89	99.39	94.80	99.83	95.31	99.23
OA	89.12	93.54	93.47	93.89	94.73	96.03
AA	95.66	97.17	96.53	97.49	96.71	97.51
Kappa	87.91	92.82	92.73	93.20	94.13	96.21

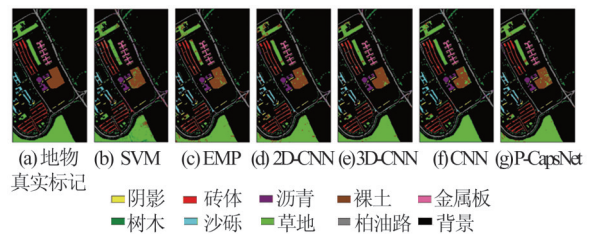


图 5 不同算法在帕维亚大学数据集上的分类图
Fig.5 Classification of Different Algorithms on Pavia University Dataset

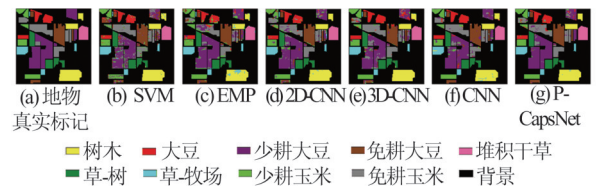


图 6 不同算法在印第安纳松树林数据集上的分类图
Fig.6 Classification of Different Algorithms on Indian Pines Dataset

为了验证胶囊网络的实例化建模能力,从帕维亚大学数据集中随机选择 4 类地物进行重构影

像的可视化展示(空间邻域 27×27)。图8第一行为P-CapsNet利用数字胶囊层中的模长最大的胶囊向量重构得到的影像,第二行为实际的输入影像。P-CapsNet在实现分类任务的同时能够准确地重构出输入影像,这表明了胶囊神经元能够对特征对象进行充分建模,从而间接表明胶囊网络具有更高层次上的抽象建模能力。

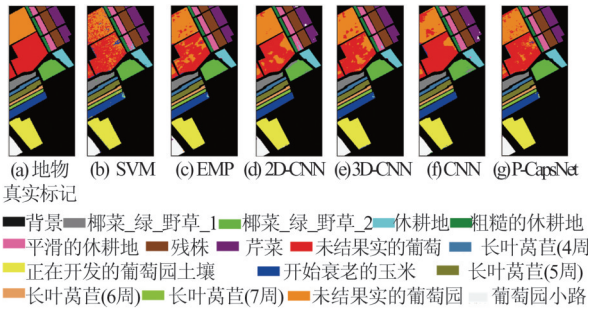


图7 不同算法在萨利纳斯山谷数据集上的分类图
Fig. 7 Classification of Different Algorithms on Salinas Dataset

2.3 训练样本数量的影响

为进一步研究P-CapsNet在小样本条件下的分类效果,从5组数据集的每类地物中分别随机选取50、100、150、200个标记样本进行实验。

图9给出了在3组数据集上不同算法总体分类精度随训练样本数量的变化情况。在帕维亚大学数据集上,当训练样本数量大于100个时,

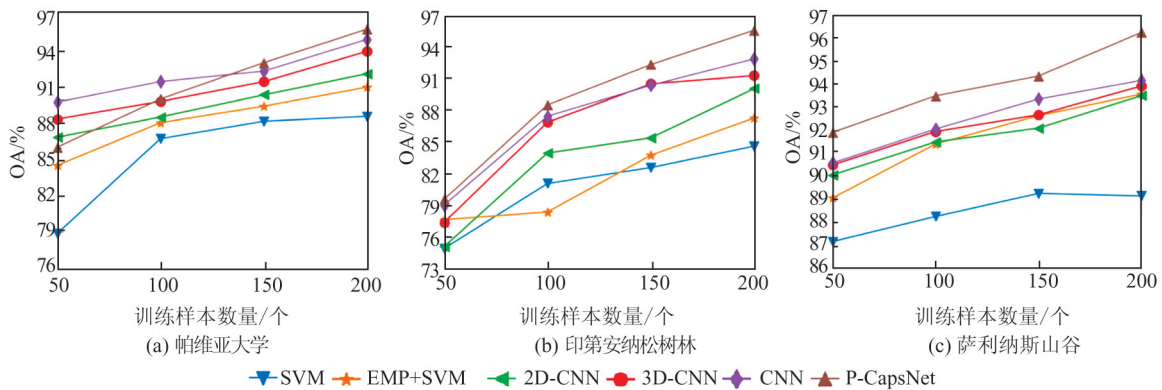


图9 不同数量的训练样本条件下的总体精度
Fig.9 Overall Accuracy with Different Numbers of Training Samples

2.4 像素邻域大小的影响

现有的2D-CNN分类模型通常通过增大像素邻域来提高影像的分类精度。这种方法能够在一定程度上达到理想的分类效果,但并不能从根本上克服CNN模型的固有缺陷。P-CapsNet同样使用二维卷积运算进行特征提取,因此有必要进行实验探究其在不同像素邻域条件下的分类性能。针对3组数据集,分别选取 3×3 、 9×9 、

15×15 、 21×21 像素的邻域作为网络输入并进行分类实验。需要说明的是,该实验中训练集、确认集和测试集的设置均与§2.2相同。

P-CapsNet的分类精度高于其他分类算法;当训练样本数量小于100个时,P-CapsNet的分类精度低于3种CNN模型但高于其他传统分类算法。在印第安纳松树林和萨利纳斯山谷数据集上,P-CapsNet的分类精度始终高于其他分类算法。可以看出,P-CapsNet需要一定数量的标记样本进行网络优化,从而保证良好的分类性能。但总体而言,在6种不同分类算法中,P-CapsNet对训练样本数量具有最好的适应性。

表9 不同算法的执行时间/s

Tab.9 Execution Time of Different Algorithms/s

数据集	训练时间/测试时间		
	3D-CNN	CNN	P-CapsNet
帕维亚大学	1 026/49.32	474/40.38	304.38/37.85
印第安纳松树林	1 569/10.68	482/9.03	205.06/7.66
萨利纳斯山谷	3 411/55.67	696/50.36	448.91/42.43

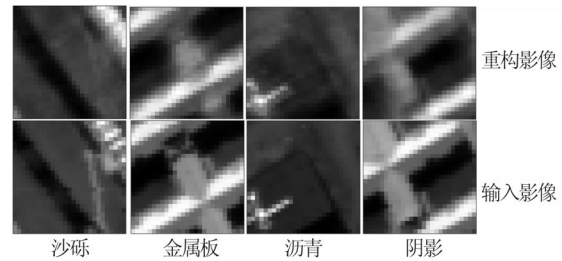


图8 输入影像与重构影像

Fig. 8 Input Images and Reconstruction Images

15×15、21×21像素的邻域作为网络输入并进行分类实验。需要说明的是,该实验中训练集、确认集和测试集的设置均与§2.2相同。

由表10~12可知,在3组数据集上,在不同邻域大小的条件下,P-CapsNet的分类精度均高于2D-CNN。在萨利纳斯山谷数据集上,2D-CNN在不同邻域大小的条件下也能够取得较好的分类精度,因此P-CapsNet的优势并不明显。在帕

维亚大学和印第安纳松树林数据集上,当像素邻域大小设置为 3×3 时,P-CapsNet 的分类精度有显著提高,较 2D-CNN 分别提高了 8.44% 和 7.26%;当像素邻域大小设置为 21×21 时,P-CapsNet 分类精度的提升并不明显。总体而言,当邻域范围较大时,2D-CNN 和 P-CapsNet 均能够取得理想的分类精度;但当像素邻域受限时,P-CapsNet 因具有更高层次上的特征表达能力,所以具有更好的分类效果。总体而言,P-CapsNet 对像素邻域的变化具有更好的适应性,且在小邻域条件下具有更明显的优势。

表 10 不同像素邻域大小对应的总体分类精度
(帕维亚大学)/%

Tab.10 Overall Classification Accuracy Corresponding to Different Sizes of Pixel Neighborhood (Pavia University)/%

分类模型	像素邻域大小			
	3×3	9×9	15×15	21×21
2D-CNN	79.55	92.11	95.18	97.09
P-CapsNet	87.99	95.79	96.76	98.05

表 11 不同像素邻域大小对应的总体分类精度
(印第安纳松树林)/%

Tab.11 Overall Classification Accuracy Corresponding to Different Sizes of Pixel Neighborhood (Indian Pines)/%

分类模型	像素邻域大小			
	3×3	9×9	15×15	21×21
2D-CNN	74.41	90.00	96.58	98.47
P-CapsNet	81.67	95.51	97.70	98.55

表 12 不同像素邻域大小对应的总体分类精度
(萨利纳斯山谷)/%

Tab.12 Overall Classification Accuracy Corresponding to Different Sizes of Pixel Neighborhood (Salinas)/%

分类模型	像素邻域大小			
	3×3	9×9	15×15	21×21
2D-CNN	89.89	93.47	97.07	98.83
P-CapsNet	90.67	96.03	97.18	98.87

总结实验结果可以发现,本文所提网络模型(P-CapsNet)在绝大多数实验条件下均取得了最高的总体分类精度。首先,P-CapsNet 以胶囊神经元为核心单元,能够对特征对象进行实例化参数建模;其次,P-CapsNet 能够利用权重矩阵建立特征对象之间的抽象空间关系。因此,P-CapsNet 能够在更高层次上对高光谱影像的特征信息进行更为抽象的表达,从而在训练样本减

少和邻域范围受限的条件下取得更高的分类精度。

3 结 语

为进一步提高高光谱影像在小邻域条件下的分类精度,本文以胶囊网络为基础,设计了一种新型网络模型 P-CapsNet。该网络模型首先利用卷积层提取影像中的低层次特征,然后利用胶囊层对高层次特征进行实例化参数建模,并利用权重矩阵编码不同特征之间的抽象空间关系,最后,利用全连接层进行影像重构,以达到网络正则化的目的。采用帕维亚大学、印第安纳松树林和萨利纳斯山谷数据集进行实验验证,结果表明:(1)在训练样本足够的条件下,P-CapsNet 的总体分类精度高于传统算法和 CNN 分类模型,且具有较低的时间开销。(2)在进一步减少训练样本和缩小像素邻域的条件下,P-CapsNet 较其他算法而言具有更好的适应性。

胶囊神经元的向量输出模式赋予了胶囊网络更高层次的特征表达能力。如何更好地适应高光谱影像的高维复杂数据结构,构建深层胶囊网络,在小样本条件下进一步提高影像分类精度,将是下一步的研究重点。

参 考 文 献

- [1] 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
- [2] Chen Y S, Lin Z H, Zhao X, et al. Deep Learning-Based Classification of Hyperspectral Data [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, 7(6): 2094-2107
- [3] Tao C, Pan H B, Li Y S, et al. Unsupervised Spectral-Spatial Feature Learning with Stacked Sparse Autoencoder for Hyperspectral Imagery Classification[J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(12): 2438-2442
- [4] Li T, Zhang J P, Zhang Y. Classification of Hyperspectral Image Based on Deep Belief Networks[C]// *IEEE International Conference on Image Processing*, Paris, France, 2014
- [5] Chen Y S, Zhao X, Jia X P. Spectral-Spatial Classification of Hyperspectral Data Based on Deep Belief Network [J]. *IEEE Journal of Selected Topics in*

- Applied Earth Observations and Remote Sensing*, 2015, 8(6): 2381-2392
- [6] Zhang X R, Sun Y J, Jiang K, et al. Spatial Sequential Recurrent Neural Network for Hyperspectral Image Classification[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018, 11(11): 4141-4155
- [7] Liu Bing, Yu Xuchu, Yu Anzhu, et al. Deep Convolutional Recurrent Neural Network with Transfer Learning for Hyperspectral Image Classification[J]. *Journal of Applied Remote Sensing*, 2018, 12(2): 026028
- [8] 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
- [9] Mei S H, Ji J Y, Bi Q Q, et al. Integrating Spectral and Spatial Information into Deep Convolutional Neural Networks for Hyperspectral Classification [C]//IEEE International Geoscience and Remote Sensing Symposium, Beijing, China, 2016
- [10] Li W, Wu G D, Zhang F, et al. Hyperspectral Image Classification Using Deep Pixel-Pair Features [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(2): 844-853
- [11] Yue J, Mao S J, Li M. A Deep Learning Framework for Hyperspectral Image Classification Using Spatial Pyramid Pooling [J]. *Remote Sensing Letters*, 2016, 7(9): 875-884
- [12] Zhang M M, Li W, Du Q. Diverse Region-Based CNN for Hyperspectral Image Classification [J]. *IEEE Transactions on Image Processing*, 2018, 27(6): 2623-2634
- [13] Zhi Lu, Yu Xuchu, Zou Bin, et al. A Multi-Layer Binary Pattern Based Method for Hyperspectral Imagery Classification Using Combined Spatial-Spectral Characteristics [J]. *Geomatics and Information Science of Wuhan University*, 2019, 44(11): 1659-1666 (职露, 余旭初, 邹滨, 等. 多层次二值模式的高光谱影像空-谱分类[J]. 武汉大学学报·信息科学版, 2019, 44(11): 1659-1666)
- [14] Li Zhuqiang, Zhu Ruifei, Gao Fang, et al. Hyperspectral Remote Sensing Image Classification Based on Three-Dimensional Convolution Neural Network Combined with Conditional Random Field Optimization[J]. *Acta Optica Sinica*, 2018, 38(8): 404-413 (李竺强, 朱瑞飞, 高放, 等. 三维卷积神经网络模型联合条件随机场优化的高光谱遥感影像分类[J]. 光学学报, 2018, 38(8): 404-413)
- [15] Gao K L, Liu B, Yu X C, et al. Deep Relation Network for Hyperspectral Image Few-Shot Classification[J]. *Remote Sensing*, 2020, 12(6): 923
- [16] Liu Bing, Yu Xuchu, Zhang Pengqiang, et al. Deep 3D Convolutional Network Combined with Spatial-Spectral Features for Hyperspectral Image Classification[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(1): 53-63 (刘冰, 余旭初, 张鹏强, 等. 联合空-谱信息的高光谱影像深度三维卷积网络分类[J]. 测绘学报, 2019, 48(1): 53-63)
- [17] Yang J, Zhao Y, Chan J C. Learning and Transferring Deep Joint Spectral-Spatial Features for Hyperspectral Classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(8): 4729-4742
- [18] Haut J M, Paoletti M E, Plaza J, et al. Active Learning with Convolutional Neural Networks for Hyperspectral Image Classification Using a New Bayesian Approach[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(11): 6440-6461
- [19] Wang L G, Hao S Y, Wang Q M, et al. Semi-supervised Classification for Hyperspectral Imagery Based on Spatial-Spectral Label Propagation [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 97: 123-137
- [20] Patrick M K, Adekoya A F, Mighty A A, et al. Capsule Networks: A Survey[J]. *Journal of King Saud University-Computer and Information Sciences*, 2019
- [21] Liu B, Yu X C, Zhang P Q, et al. Supervised Deep Feature Extraction for Hyperspectral Image Classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(4): 1909-1921
- [22] Lee H, Kwon H. Going Deeper with Contextual CNN for Hyperspectral Image Classification [J]. *IEEE Transactions on Image Processing*, 2017, 26(10): 4843-4855
- [23] Gu J X, Wang Z H, Kuen J, et al. Recent Advances in Convolutional Neural Networks[J]. *Pattern Recognition*, 2018, 77: 354-377
- [24] Sabour S, Frosst N, Hinton G E. Dynamic Routing Between Capsules[C]// Conference and Workshop on Neural Information Processing Systems, Long Beach, CA, USA, 2017
- [25] Glorot X, Bengio Y. Understanding the Difficulty of Training Deep Feedforward Neural Networks[C]// International Conference on Artificial Intelligence and Statistics, Chia Laguna Resort, Sardinia, Italy, 2010
- [26] Benediktsson J A, Palmason J A, Sveinsson J R.

Classification of Hyperspectral Data from Urban Areas
Based on Extended Morphological Profiles [J]. *IEEE*

Transactions on Geoscience and Remote Sensing,
2005, 43(3): 480-491

Hyperspectral Image Spatial-Spectral Classification Using Capsule Network Based Method

GAO Kuiliang¹ YU Xuchu¹ ZHANG Pengqiang¹ TAN Xiong¹ LIU Bing¹

¹ Information Engineering University, Zhengzhou 450001, China

Abstract: Objectives: The deep learning model such as convolutional neural network(CNN) has achieved satisfactory results in hyperspectral images classification. However, because traditional neurons can only perform scalar computation, the existing deep learning models cannot model the instantiation parameters of hyperspectral image features, so they cannot achieve satisfactory classification results under the condition of restricted neighborhood scope. Aiming at the problem, we design a new network model by introducing capsule network structure, to further improve the classification accuracy.**Methods:** The model is composed of traditional convolution layers, capsule layers and fully connected layers, which has a stronger feature representation ability. This model can further improve the classification accuracy of hyperspectral images using the vector calculation of capsule neurons and the spatial relationship between features encoded by the weight matrix. Specifically, the hyperspectral image patches are firstly processed by a convolution layer to extract local features. Next, the primary capsule layer and the digital capsule layer are used to extract the deeper abstract features at higher levels and classify the input data. In addition, the fully connected layers are used for reconstruction to further enhance the abstract modeling ability and generalization ability of the capsule network.**Results:** Three public hyperspectral images data sets including Pavia University, Indian Pines and Salinas are selected for experiments. The results show that the proposed method outperform the support vector machine(SVM)-based and the traditional CNN-based classification methods. Specifically, compared with the SVM-based methods, the proposed method improves the overall classification accuracy by about 4.7%—7.2%, 8.2%—10.9% and 2.5%—6.9% on three different datasets. Compared with the traditional CNN-based methods, the proposed method improves the overall classification accuracy by about 0.8—3.7%, 2.7—5.5% and 1.3—2.5% on three different datasets.**Conclusions:** In conclusion, the proposed network model has better classification performance than that of traditional algorithms. Under the condition of sufficient training samples, the overall classification accuracy of the proposed model is higher than the traditional SVM-based and CNN-based classification models, and it has a lower time cost. In addition, the proposed model has better adaptability under the condition of further reducing training samples and pixel neighborhoods.

Key words: hyperspectral image classification; capsule network; capsule neurons; convolutional layer; deep learning

First author: GAO Kuiliang, master, specializes in hyperspectral image processing and analysis. E-mail: gokling1219@163.com

Foundation support: The Science and Technology Plan of Henan Province(182102210148).

引文格式: GAO Kuiliang, YU Xuchu, ZHANG Pengqiang, et al. Hyperspectral Image Spatial-Spectral Classification Using Capsule Network Based Method[J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(3):428-437. DOI:10.13203/j.whugis20200008(高奎亮, 余旭初, 张鹏强, 等. 利用胶囊网络实现高光谱影像空谱联合分类[J]. *武汉大学学报·信息科学版*, 2022, 47(3):428-437. DOI:10.13203/j.whugis20200008)