

# 多特征融合下的高光谱图像混合卷积分类

熊余<sup>1,2,3</sup>, 单德明<sup>1,2,3</sup>, 姚玉<sup>1,2,3</sup>, 张宇<sup>1,2,3</sup>

(1. 重庆邮电大学 通信与信息工程学院, 重庆 400065; 2. 重庆高校市级光通信与网络重点实验室, 重庆 400065;  
3. 泛在感知与互联重庆市重点实验室, 重庆 400065)

**摘要:** 针对现有高光谱遥感图像卷积神经网络分类算法空谱特征利用率不足的问题, 提出一种多特征融合下基于混合卷积胶囊网络的高光谱图像分类策略。首先, 联合使用主成分分析和非负矩阵分解对高光谱数据集进行降维; 然后, 将降维所得主成分通过超像素分割和余弦聚类生成一个多维特征集; 最后, 将叠加后的特征集通过二维、三维多尺度混合卷积网络进行空谱特征提取, 并使用胶囊网络对其进行分类。通过在不同高光谱数据集下的实验结果表明, 在相同 20 维光谱维度下, 所提策略相比于传统分类策略在总体精度、平均精度以及 Kappa 系数上均有明显提升。

**关键词:** 图像分类; 高光谱图像; 降维; 超像素; 混合卷积胶囊网络

中图分类号: TP391.41 文献标识码: A 文章编号: 1001-8891(2022)01-0009-12

## Hyperspectral Image Hybrid Convolution Classification under Multi-Feature Fusion

XIONG Yu<sup>1,2,3</sup>, SHAN Deming<sup>1,2,3</sup>, YAO Yu<sup>1,2,3</sup>, ZHANG Yu<sup>1,2,3</sup>

(1. School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; 2. Key Laboratory of Optical Communication and Networks in Chongqing, Chongqing 400065, China;  
3. Key Laboratory of Ubiquitous Sensing and Networking in Chongqing, Chongqing 400065, China)

**Abstract:** To address the problem of insufficient utilization of spatial-spectrum features in existing convolutional neural network classification algorithms for hyperspectral remote sensing images, we propose a hyperspectral image classification strategy based on a hybrid convolution capsule network under multi-feature fusion. First, a combination of principal component analysis and non-negative matrix decomposition is used to reduce the dimensionality of a hyperspectral dataset. Second, the principal components obtained through dimensionality reduction are used to generate a multidimensional feature set through super-pixel segmentation and cosine clustering. Finally, the superimposed feature set is used to extract spatial-spectrum features through a two-dimensional and three-dimensional multi-scale hybrid convolutional network, and a capsule network is used to classify them. We performed experiments on different hyperspectral datasets, and the results revealed that under the same 20-dimensional spectral setting, the proposed strategy significantly improves the overall accuracy, average accuracy, and Kappa coefficient compared to traditional classification strategies.

**Key words:** image classification, hyperspectral image, dimensionality reduction, superpixel, hybrid convolution capsules network

## 0 引言

高光谱遥感是指在电磁波谱的紫外、可见光、近红外和中红外区域, 利用成像光谱仪获取窄而连续的光谱图像数据<sup>[1]</sup>。由于包含丰富的光谱信息使得高光

谱图像(hyperspectral image, HSI)能够捕捉复杂的地物分布, 因此 HSI 被广泛应用于深林植被保护、水质检测和地质勘测等<sup>[2]</sup>。然而, 由于大气分子影响使图像存在椒盐噪声, 以及光谱高维性所带来的冗余, 给遥感图像实现准确的分类带来了极大挑战。

收稿日期: 2020-11-02; 修订日期: 2021-01-25。

作者简介: 熊余(1982-), 男, 研究员, 博士, 主要研究方向为教育大数据, 光网络。E-mail: xiongyu@cqupt.edu.cn。

基金项目: 国家自然科学基金资助项目(61401052); 国家留学基金委资助项目(201608500030); 重庆市教委科学技术研究资助项目(KJ1400418, KJ1500445); 重庆邮电大学博士启动基金资助项目(A2015-09)。

现有的遥感图像分类技术大致可分为基于像元的分类方法和基于联合空间特征和光谱信息分类的方法<sup>[3]</sup>。前者主要是通过同类地物光谱之间的相似性特征进行分类，一般采用支持向量机（support vector machine, SVM）、逻辑回归和反向传播神经网络等<sup>[4]</sup>。它们基于光谱特征的相似性进行特征提取、选择和分类，由于易受噪声和光谱冗余影响，分类精度普遍较低。而后者主要基于相似区域将空间信息和光谱信息进行结合从而提高分类精度，如利用形态学扩展、超像素分割和二维卷积网络（two dimensional convolutional neural network, 2DCNN）等<sup>[5]</sup>。它们主要依赖于对高光谱图像中的空间特征信息提取，并与光谱特征相结合从而提高分类精度。其中卷积网络以其自动挖掘特征的优势，摆脱了人工提取特征的困难，获得了研究者的广泛关注。如文献[6]提出了联合使用最大池和平均池的组合型二维网络以避免忽略相邻像素特征之间的相关性，有效地提高了网络的分类性能。但为了获得更好的特征提取，该方法需要堆叠深层的卷积层，增加了网络的复杂性与不可解释性。因此，文献[7]结合哈希特征来改善网络，通过使用哈希函数和二进制哈希码来改进语义特征的提取，以增强网络的可解释性。但未考虑光谱波段之间的相关性，且由于光谱信息的高维性，将导致计算量增大。为此，文献[8]提出使用三维卷积神经网络（three dimensional convolutional neural network, 3DCNN）作为分类器，以实现对空间和光谱特征的同时提取，在采用较少参数的同时显著提升了分类精度。但受限于固定大小的卷积核，未能对地物边界像元的特征进行较好提取，带来空间信息利用不足的问题。而文献[9]结合10层卷积层和1层全连接层构造了混合分类网络，通过联合使用一维、二维和三维的小卷积核来多尺度地提取光谱和空间特征，显著提升了分类结果。文献[10]设计了一种结合二维、三维卷积核的混合卷积网络，旨在提高对主成分分析（principal component analysis, PCA）降维后高光谱图像空谱特征的提取能力，同时降低了模型的复杂度，取得了较好的分类精度。但该模型需对不同数据集进行不同程度的降维，泛化能力不足，且空间和光谱特征融合仅限于三维卷积核，没有充分利用光谱信息和空间特征以及未考虑小样本对该模型分类的影响。

基于上述问题，本文提出了一种多特征融合下基于混合卷积胶囊网络（hybrid convolutional capsule network with multi-feature fusion, MFF-HCCN）的高光谱分类策略。首先，使用主成分分析和非负矩阵分解（non-negative matrix factorization, NMF）对HSI数

据集进行降维；然后，使用简单非迭代聚类（Simple Non-Iterative Clustering, SNIC）算法和余弦聚类算法进行图像分割和聚类融合；最后，将融合后的数据集通过改进的混合卷积胶囊网络（hybrid convolutional capsule network, HCCN）进行分类。

## 1 MFF 数据预处理

高光谱图像的光谱特征具有高维性和冗余性，导致计算量大、分类困难，在分类前一般采用PCA与NMF对高光谱数据集进行特征提取<sup>[11]</sup>。PCA依赖于采用最大化方差去相关性，以保留数据集信息的主要成分，但降维数据可正可负，造成PCA的基本在原始数据方面不具有直观的物理意义<sup>[12]</sup>。虽然NMF分解结果为非负，具有可解释性和明确的物理意义，但进行高光谱图像的光谱解混时，易受噪声影响陷入局部最优<sup>[13]</sup>。为此，综合考虑二者的性能后，提出了一种PCA和NMF联合的多特征融合（multi-feature fusion, MFF）算法。

### 1.1 数据降维

给定一个高光谱图像数据  $\mathbf{H}_{W \times H \times B} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ ， $\mathbf{x}_i \in \mathbb{R}^{W \times H \times B}$ ， $i=1, 2, \dots, n$ 。其中， $x_i$  为 HSI 的第  $i$  个像元， $n$  为高光谱像元数量， $W$  和  $H$  分别为高光谱图像空间维的宽与高， $B$  为光谱维波段数目。NMF 的基本思想可以简述为对于任意给定的非负矩阵，总能找到一对非负加权基向量矩阵和非负加权系数矩阵相乘的形式。即可使用公式(1)对高光谱数据集降维到三维，得到  $\mathbf{I}_{W \times H \times 3} = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ 。

$$\mathbf{H}_{(W \times H) \times B} = \mathbf{V}_{(W \times H) \times r} \mathbf{W}_{r \times B} \quad (1)$$

式中： $\mathbf{H}_{(W \times H) \times B}$  为高光谱数据； $\mathbf{V}_{(W \times H) \times r}$  和  $\mathbf{W}_{r \times B}$  分别对应加权矩阵和加权系数矩阵； $r$  为降维的特征数目。

虽然NMF能够较好地保留HSI的特征，但原始信息缺失较大，需要使用PCA对HSI进行降维。因此，使用PCA对HSI进行降维，分别得到前5主成分  $\mathbf{I}_{W \times H \times 5} = \{\mathbf{PC}_1, \mathbf{PC}_2, \dots, \mathbf{PC}_5\}$  和前16主成分  $\mathbf{I}_{W \times H \times 16} = \{\mathbf{PC}_1, \mathbf{PC}_2, \dots, \mathbf{PC}_{16}\}$ 。图1为以Pavia University数据集为例的PCA降维的各个成分之间的方差比例，数据的前5主成分和前16主成分信息分别保留在98%和99%以上。

### 1.2 超像素 SNIC 分割与聚类

超像素是由一系列位置相邻且颜色、亮度、纹理等特征相似的像素点组成的小区域，它能够提取图像局部特征，获取图像的冗余信息<sup>[14]</sup>。而为获取高光谱图像各类地物分布情况，使用文献[15]所提的SNIC分

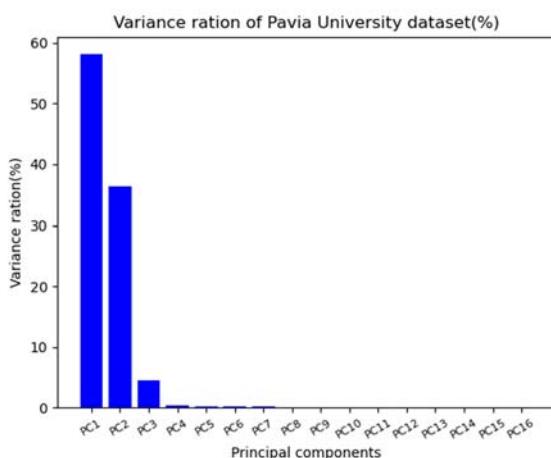


图1 Pavia University 数据集 PCA 降维成分方差比例分布图

Fig.1 The distribution of variance ratio of PCA dimension reduction components in Pavia University dataset

割算法对 NMF 所降维的三维图像进行分割。该算法是传统简单线性迭代聚类算法 (simple linear iterative clustering, SLIC) 的改进算法, 它以单次迭代和强制性连接取代了 SLIC 的 K-means 聚类分割, 能够更快、更有效地对物体边界分割。SNIC 将图像由 RGB 颜色空间转化为颜色分明的 CIELAB 颜色空间, 在迭代过程中以初始中心像元的 4 邻域或 8 邻域像元作为一个超像素, 其距离计算公式如公式(2)所示。

$$d_{i,j} = \sqrt{\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{s} + \frac{\|\mathbf{c}_i - \mathbf{c}_j\|_2^2}{m}} \quad (2)$$

式中:  $\mathbf{x}=(x, y)^T$  表示像元的空间位置;  $\mathbf{c}=(l, a, b)^T$  表示像元在 CIELAB 空间的颜色通道;  $s$  和  $m$  分别为第  $i$  个像元与第  $j$  个像元之间的空间距离和颜色距离的归一化因子。同时,  $m$  也称为紧密度参数,  $m$  值越高, 超像素越紧凑, 但边界贴合度越差, 反之亦然。

使用 NMF 分解后的三维成分进行超像素分割而非 PCA 降维后的前 3 主成分, 是因为 NMF 将数据集  $\mathbf{H}_{W \times H \times B}$  分解为三维的矩阵, 能够较好地保留个别地物的分布情况。从图 2 的 Pavia University 降维图可以看出, 使用 NMF 降维后的图 2(a)比使用 PCA 降维的图 2(b)地物边界保留更加完整、清晰。然后, 为保证分割图保留较为完整信息, 将分割后的图像与高光谱数据集  $\mathbf{I}'_{W \times H \times 5}$  按像元空间位置进行合并, 得到一个具有五维光谱通道的立方体图, 如图 2(c)所示。

图 2(c)中的超像素可表示为  $\mathbf{I}_k=\{x_1^k, x_2^k, x_3^k, x_4^k, x_5^k\}$ ,  $k=1,2,\cdots,K$ 。其中  $K$  表示超像素的个数。取每个超像素内所有像元光谱均值作为当前超像素的光谱值, 并以每个超像素作为一个节点

$$\mathbf{o}_k = \{\bar{\mathbf{x}}_1^k, \bar{\mathbf{x}}_2^k, \bar{\mathbf{x}}_3^k, \bar{\mathbf{x}}_4^k, \bar{\mathbf{x}}_5^k\}, \text{ 即: } \bar{\mathbf{x}}_i^k = \frac{1}{p} \sum_{i=0}^{p-1} \mathbf{x}_i \quad (3)$$

式中:  $i=\{0, 1, \cdots, p-1\}$  为每个超像素所包含的像元个数;  $k=\{1, 2, \cdots, K\}$  为第  $k$  个超像素。因此, 得到一个节点向量  $\mathbf{O}=\{\mathbf{o}_1^T, \mathbf{o}_2^T, \cdots, \mathbf{o}_K^T\}$ , 同时也得到了一个均值化后的 HSI 图像数据  $\mathbf{I}'_{W \times H \times 5}$ 。

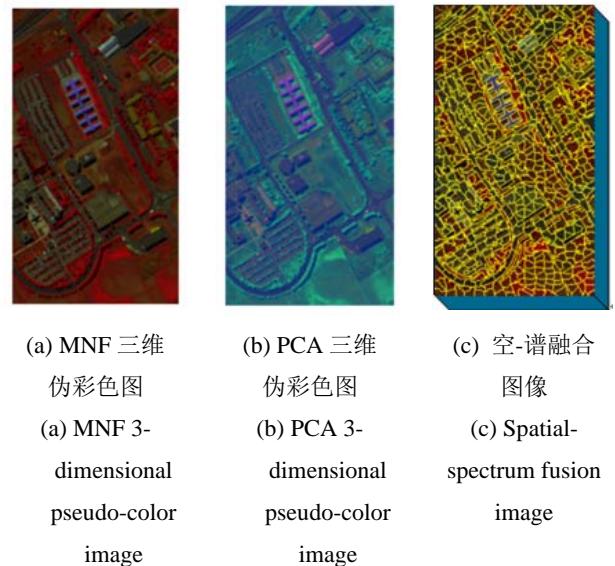


图2 Pavia University 降维图

Fig.2 Pavia University dimensionality reduction image

对于地物分布的初始聚类,  $\mathbf{I}'_{W \times H \times 5}$  拥有丰富的光谱信息。为了得到更为精确的粗分类, 以光谱波段之间的夹角作为相似性度量, 引入余弦聚类算法对  $\mathbf{I}'_{W \times H \times 5}$  进行无监督分类。余弦聚类中的距离公式如下所示:

$$r_U = \frac{1}{n} \sum_{i \in n, j \in n, i \neq j} \left( \frac{\mathbf{x}_i}{\sigma_{\mathbf{x}_i}^{(0)}} \right) \left( \frac{\mathbf{x}_j}{\sigma_{\mathbf{x}_j}^{(0)}} \right) \quad (4)$$

式中:  $\sigma_{\mathbf{x}_i}^{(0)} = \sqrt{\frac{1}{n} \sum_{i=1}^n \mathbf{x}_i^2}$ ,  $\sigma_{\mathbf{x}_j}^{(0)} = \sqrt{\frac{1}{n} \sum_{j=1}^n \mathbf{x}_j^2}$ ,  $\sigma_{\mathbf{x}_i}^{(0)}$  和  $\sigma_{\mathbf{x}_j}^{(0)}$  为两个样本波段之间的标准差, 范围在 -1 和 1 之间, 其系数的距离计算如下:

$$d_U = 1 - r_U, \quad d_U \in [0, 2] \quad (5)$$

以像元间光谱波段的夹角余弦距离度量像元的相似性, 取代传统的欧式距离度量, 并做 K-means 聚类得到的聚类结果作为图像的初始分类标签  $C_{ini} = \{L_1, L_2, \cdots, L_{ini}\}$ 。其中,  $L_i$  表示类标签,  $ini$  表示初始聚类得到标签数量。

### 1.3 多特征融合

仅使用单一的  $\mathbf{I}_{W \times H \times 16}$  虽然保留较为完整的光谱

信息并去除了一定的冗余噪声,但同谱异物、异物同谱问题仍然突出,所以这里使用 $\mathbf{I}'_{W \times H \times 5}$ 取代 $\mathbf{I}_{W \times H \times 16}$ 光谱维的前5主成分可以在一定程度上增加同类地物像元之间的相似性和异物之间的差异性。然而SNIC算法在进行图像分割时会产生一定的误差,因此在此基础上加入 $\mathbf{I}_{W \times H \times 3}$ 和标签集 $C_{ini}$ ,这将有效提高异物像元之间的差异性,具体如公式(6):

$$\mathbf{I}_{W \times H \times 20} = \mathbf{I}'_{W \times H \times 11} + \mathbf{I}'_{W \times H \times 5} + \mathbf{I}_{W \times H \times 3} + \mathbf{C}_{ini} \quad (6)$$

对所获得的 $\mathbf{I}_{W \times H \times 20}$ 进行归一化处理得到 $\mathbf{I}'_{W \times H \times 20}$ ,易于分类处理。归一化处理函数如下:

$$\mathbf{D}' = \frac{\mathbf{D}_i - \min\{\mathbf{D}\}}{\max\{\mathbf{D}\} - \min\{\mathbf{D}\}} \quad (7)$$

式中: $\mathbf{D}$ 为给定的待归一化的数据集; $\mathbf{D}_i$ 为 $\mathbf{D}$ 的元素; $\mathbf{D}'$ 为归一化后的结果。

## 2 分类网络

高光谱数据集经过以上预处理后,可以得到一个具有空谱融合特征的20维数据集。为了进一步提高对空间和光谱特征的利用率,首先使用4种不同尺度的三维卷积核对数据集进行空谱特征提取,然后使用3种不同尺度的二维卷积核进一步提取局部空间特征,最后通过具有方向特性的胶囊网络进行分类。

### 2.1 HCCN 网络

已知输入数据集为 $\mathbf{I}'_{W \times H \times 20} \in \mathbb{R}^{W \times H \times 20}$ ,从 $\mathbf{I}'_{W \times H \times 20}$ 立方体中创建3D邻域块 $\mathbf{P} \in \mathbb{R}^{Q \times Q \times 20}$ , $Q$ 为输入图像的宽和高。为了能够充分利用图像的空间信息,以重叠的3D块作为输入,并以中心像素的标签作为该立方块的类别标签。故在整个数据集 $\mathbf{I}'_{W \times H \times 20}$ 中总共生成了 $(W-Q+1)(H-Q+1)$ 个3D邻域块。其中,每一块 $\mathbf{P}$ 所对应的中心位置为 $(\alpha, \beta)$ , $\mathbf{P}$ 的空间大小为 $Q \times Q$ 。因此每一块3D邻域块可由其中心表示,即宽的范围为 $\alpha-(Q-1)/2 \sim \alpha+(Q-1)/2$ ,高的范围为 $\beta-(Q-1)/2 \sim \beta+(Q-1)/2$ ,光谱波段数为20。然后将所得的邻域块输入到三维卷积网络中。

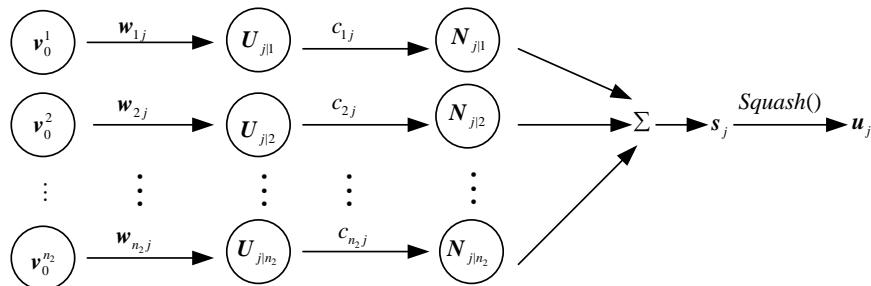


图3 胶囊网络神经元解析图

Fig.3 Analytic diagram of capsule network neurons

在三维卷积过程中,第*i*层第*j*个特征图在空间位置(*x,y,z*)处的特征值 $\mathbf{v}_{i,j}^{x,y,z}$ 可由公式(8)得到:

$$\mathbf{v}_{i,j}^{x,y,z} = f(\sum_{\tau=1}^{d_{i-1}} \sum_{\lambda=-\eta}^{\eta} \sum_{\rho=-r}^r \sum_{\sigma=-\delta}^{\delta} \mathbf{w}_{i,j,\tau}^{\sigma,\rho,\lambda} \times \mathbf{v}_{i-1,\tau}^{x+\sigma, y+\rho, z+\lambda} + b_{i,j}) \quad (8)$$

式中: $f$ 为ReLU激活函数; $b_{i,j}$ 为第*i*层第*j*个特征图的偏置; $d_{i-1}$ 为第(*i*-1)层特征图的数量和用于生成第*i*层第*j*个特征图卷积核 $\mathbf{w}_{i,j}$ 的深度, $2r+1$ 、 $2\delta+1$ 和 $2\eta+1$ 分别为卷积核 $\mathbf{w}_{i,j}$ 的宽度、高度和光谱通道数。

首先,为了加强对不同空间尺度特征提取,在三维卷积层中使用 $3 \times 2 \times 3$ 、 $2 \times 3 \times 3$ 、 $2 \times 1 \times 3$ 和 $1 \times 2 \times 7$ 四种不同大小的三维卷积核,可以得到一系列具有空谱特征的立方体特征图。然后,将所得的特征立方体图根据其光谱通道数变换为二维特征图,分别使用3种不同尺度的 $3 \times 3$ 、 $2 \times 2$ 和 $1 \times 1$ 二维卷积核进行滤波处理,如公式(9)所示。并将每个滤波器输出的二维特征图平铺为一维向量并进行叠加如公式(10)所示。

$$\mathbf{v}_{i,j}^{x,y} = f(b_{i,j} + \sum_{\tau=1}^{d_{i-1}} \sum_{\rho=-r}^r \sum_{\sigma=-\delta}^{\delta} \mathbf{w}_{i,j,\tau}^{\sigma,\rho} \times \mathbf{v}_{i-1}^{x+\sigma, y+\rho}) \quad (9)$$

$$\mathbf{v}_0 = \bigcup_{h=2,3,5} \mathbf{v}_{h \times h} \quad (10)$$

式中: $h$ 为使用的二维卷积核的宽和高大小; $\mathbf{v}_{h \times h}$ 为二维卷积层输出特征图整形后的结果。

$\mathbf{v}_0 = \{\mathbf{v}_0^1, \mathbf{v}_0^2, \dots, \mathbf{v}_0^{n_2}\}$ 表示叠加后的特征值; $n_2$ 表示输入的向量个数。

将叠加后的 $\mathbf{v}_0$ 输入到胶囊网络中进行特征提取,获得一组具有大小和方向的向量<sup>[16]</sup>。胶囊卷积是传统神经网络的一种改进形式,它具有平移不变性以及可以自动选择图像的位置、大小和方向等特征并做特征的聚类以强化相似性,输出一种更具有表达力的特征向量。因此,使用胶囊网络接收多尺度卷积层输出的特征可以有效提高对不同特征的提取和判别能力,以提升模型的分类性能。图3展示了胶囊网络内部神经元的加权组合过程。

首先, 将输入向量  $\mathbf{v}_0$  通过权重  $\mathbf{W} = \{\mathbf{w}_{1j}, \mathbf{w}_{2j}, \dots, \mathbf{w}_{n_2 j}\}$  获得一组新的输入神经元  $\mathbf{U} = \{\mathbf{U}_{j1}, \mathbf{U}_{j2}, \dots, \mathbf{U}_{jn_2}\}$  如公式(11)所示。然后, 通过公式(12)乘上耦合系数  $\mathbf{c} = \{c_{1j}, c_{2j}, \dots, c_{n_2 j}\}$  以增加异物特征的差异性并求和输出胶囊  $\mathbf{s}_j$ 。最后通过压缩激活函数获得一组压缩后的胶囊向量, 如公式(13)所示。

$$\mathbf{U}_{ji} = \mathbf{W}_{ij} \cdot \mathbf{v}_i, i=1,2,\dots,n_2, j \in N \quad (11)$$

$$\mathbf{s}_j = \sum_i c_{ij} \cdot \mathbf{U}_{ji} \quad (12)$$

$$\mathbf{u}_j = \frac{\mathbf{s}_j}{\|\mathbf{s}_j\|} \cdot \frac{\|\mathbf{s}_j\|^2}{1 + \|\mathbf{s}_j\|^2} \quad (13)$$

式中:  $i$  为图像中的特征位置;  $j$  为任意向量神经元。

$\frac{\mathbf{s}_j}{\|\mathbf{s}_j\|}$  使向量单位化, 即将  $\mathbf{s}_j$  的模长压缩为 1。 $\frac{\|\mathbf{s}_j\|^2}{1 + \|\mathbf{s}_j\|^2}$  表示  $\mathbf{s}_j$  的模长越大, 其代表的特征就越强。

对于所获得的每一个胶囊属于一个类, 使用胶囊向量的模值长度表示对应像元的类概率。因此使用公式(14)边缘损失作为整个网络的损失函数:

$$L_M = \sum_{j=1}^n \left[ T_j \max(0, a^+ - \|\mathbf{u}_j\|)^2 + \lambda (1 - T_j) \max(0, \|\mathbf{u}_j\| - a^-)^2 \right] \quad (14)$$

式中:  $a^+ = 0.9$ ,  $a^- = 0.1$ ,  $\lambda = 0.25$ , 如果像元属于第  $j$  类,  $T_j = 1$ 。参数  $a^+$  表示输出的向量  $\|\mathbf{u}_j\|$  的长度大于  $a^+$

时, 则可以确定该像元属于第  $j$  类; 参数  $a^-$  表示  $\|\mathbf{u}_j\|$  小于  $a^-$  时, 则像元不属于第  $j$  类。 $\lambda$  为控制权值大小的参数, 以达到当像元误判时停止初始化学习。

所提 HCCN 分类如图 4 以 Pavia University 数据集为例所示, 其各层参数如表 1 所示。

表 1 中卷积核的前 3 个数为核的大小, 第 4 个数为核的数量。输出大小表示每层输出特征图的大小。如第一层卷积层的卷积核的参数为(3,2,3,16), 分别表示卷积核大小为  $3 \times 2 \times 3$  和核的数量为 16。输出为  $(9,10,18)$ , 表示输出的特征图大小为  $9 \times 10 \times 18$ 。

## 2.2 MFF-HCCN 算法步骤

MFF-HCCN 算法主要分为以下几步:

步骤 1: 使用 PCA 对 HSI 数据集进行降维, 获取前 5 主成分  $\mathbf{I}_{W \times H \times 5}$  和前 16 主成分  $\mathbf{I}_{W \times H \times 16}$  并使用 NMF 获取 HSI 的前 3 成分  $\mathbf{I}_{W \times H \times 3}$ 。

步骤 2: 使用 SNIC 分割算法对  $\mathbf{I}_{W \times H \times 3}$  进行分割, 并对超像素内像元的前 5 主成分进行平均得到  $\mathbf{I}'_{W \times H \times 5}$ 。以每个超像素为结点通过公式(4)的余弦聚类得到聚类标签  $\mathbf{C}_{ini}$ 。

步骤 3: 以  $\mathbf{I}'_{W \times H \times 5}$  取代  $\mathbf{I}_{W \times H \times 16}$  的前 5 主成分, 并融合  $\mathbf{I}_{W \times H \times 3}$  和  $\mathbf{C}_{ini}$  获得  $\mathbf{I}_{W \times H \times 20}$ 。

步骤 4: 对  $\mathbf{I}_{W \times H \times 20}$  使用公式(7)进行归一化处理, 并输入到改进的混合卷积胶囊网络 HCCN 中进行分类处理。

MFF-HCCN 算法的结构图如图 5 所示。

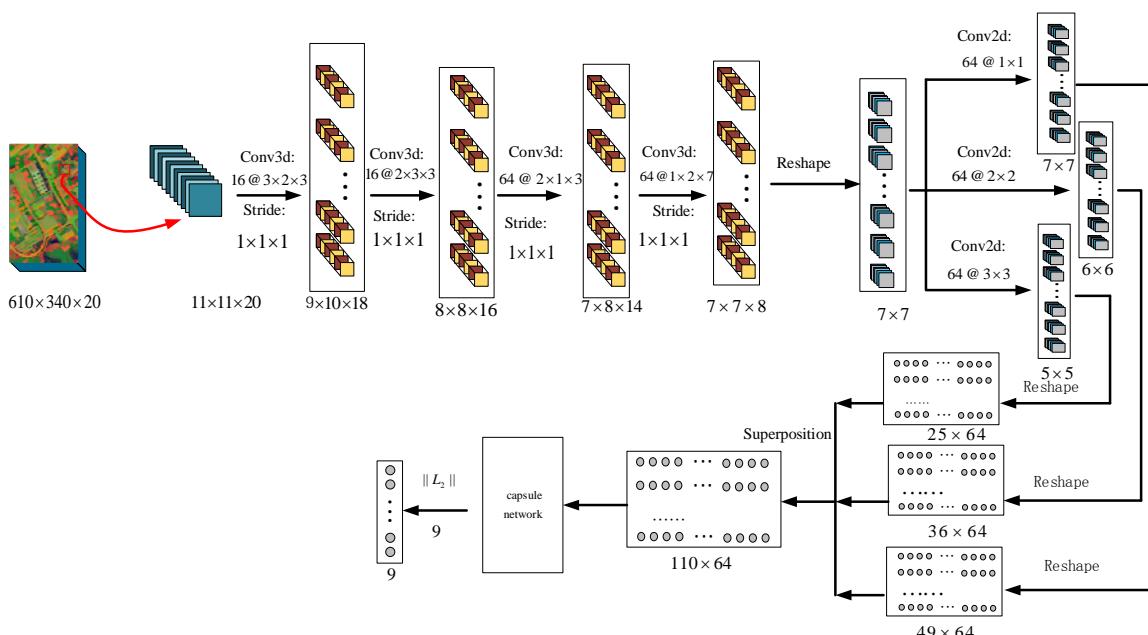


图 4 HCCN 分类示意图

Fig.4 Schematic diagram of HCCN classification

表1 Pavia University 数据集卷积分类各层的参数

Table 1 Parameters of each layer of convolutional classification of Pavia University dataset

Network layer (type)	Convolution kernel	Stride	Parameter	Output
Input layer			0	(11,11,20)
Conv3D layer1	(3,2,3,16)	(1,1,1)	304	(9,10,18)
Conv3D layer2	(2,3,3,16)	(1,1,1)	4624	(8,8,16)
Conv3D layer3	(2,1,3,64)	(1,1,1)	6208	(7,8,14)
Conv3D layer4	(1,2,7,64)	(1,1,1)	57408	(7,7,8)
Reshape1			0	(7,7)
Conv2D layer1	(1,1,64)	(1,1)	32832	(7,7)
Reshape2			0	(49,64)
Conv2D layer2	(2,2,64)	(1,1)	131136	(6,6)
Reshape3			0	(36,64)
Conv2D layer3	(3,3,64)	(1,1)	294976	(5,5)
Reshape			0	(25,64)
Concatenate			0	(110,64)
Capsule			9216	(9,16)
Output layer				9

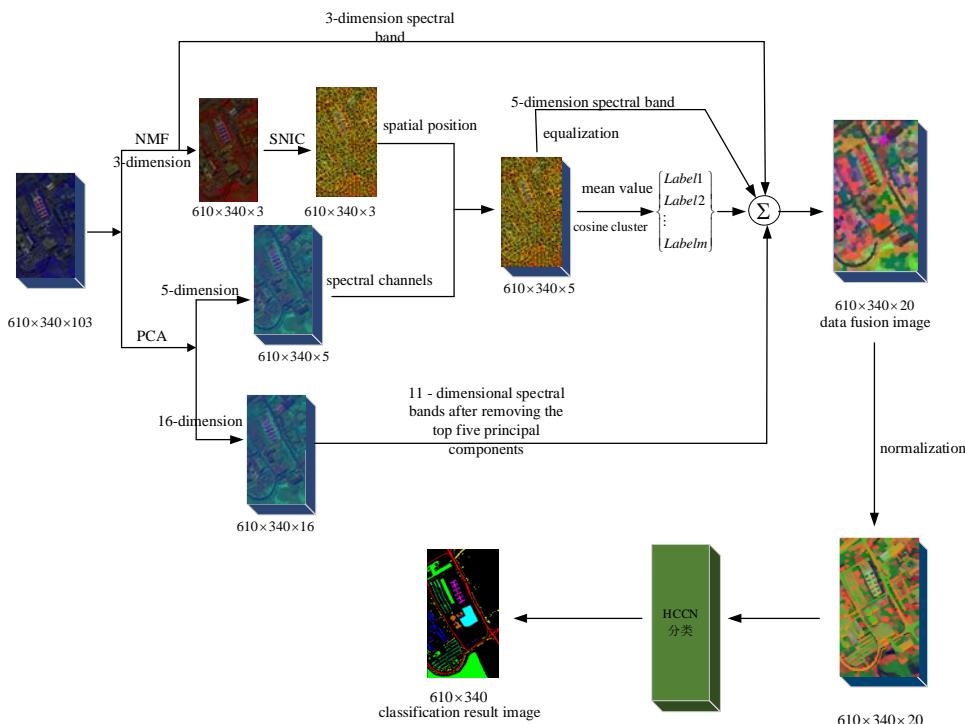


图5 MFF-HCCN 算法结构图

Fig.5 MFF-HCCN algorithm structure diagram

### 3 仿真分析与讨论

#### 3.1 实验数据

实验以3种公开的遥感数据集 Indian Pines、Pavia University 和 WHU-Hi-Longkou 来检验算法分类效果。

1) Indian Pines 数据集<sup>[7]</sup>。该数据集是通过 AVIRIS 传感器于 1992 年 6 月获取的来自美国印第安纳州北部农业区的具有 220 个波段的高光谱遥感图像，光谱波段范围为 0.4~2.5 μm。该图像大小为 145×145 像素，空间分辨率约为 20 m。该数据去除了水吸收和低信噪比波段后为 200 个波段数据，总共划分为 16 类真实地物，如表 2 所示。

2) Pavia University 数据集<sup>[7]</sup>。该数据集是通过 ROSIS 传感器于 2001 年获取的 Pavia 大学的场景图，具有 115 个光谱波段，光谱波段范围在 0.43~0.86 μm，所获图像大小为 610×340 像素，空间分辨率大约为 1.3 m，真实标记地物划分为 9 类。实验选取 103 个波段用于分类，如表 3 所示。

3) WHU-Hi-Longkou 数据集<sup>[17]</sup>。该数据集是通过搭载 Headwall Nano-Hyperspec 成像传感器的无人机于 2018 年 7 月获取的中国湖北省龙口镇农业场景图, 具有 270 个光谱波段, 光谱波段范围在 400~1000 nm, 所获图像大小为 550×400 像素, 空间分辨率约为 0.463 m, 真实标记地物划分为 9 类, 如表 4 所示。

表 2 Indian Pines 数据集的地物类别和样本数

Table 2 Land cover classes and numbers of samples in Indian Pines dataset

No.	Class name	Numbers of samples
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-min	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-grass-trees-crives	386
16	Stone-steel-towers	93
Total		10249

表 3 Pavia University 数据集的地物类别和样本数

Table 3 Land cover classes and numbers of samples in Pavia University dataset

No.	Class name	Numbers of samples
1	Asphalt	6631
2	Meadows	18649
3	Gravel	2099
4	Trees	3064
5	Painted metal sheets	1345
6	Bare soil	5029
7	Bitumen	1330
8	Self-blocking bricks	3682
9	Shadows	947
Total		42776

### 3.2 参数设置

在对高光谱图像进行分类时, 使用 SNIC 将 Indian Pines 数据集分割为 200 个超像素。考虑到没有标记

地物的影响, 初始聚类数目为 18 类。当对 Pavia University 数据集进行分类时, SNIC 将其分割为 800 个超像素, 初始聚类数目为 11 类。同样, 在对 WHU-Hi-Longkou 数据集进行分割时, 将其分割为 500 个超像素, 初始聚类数目为 11 类。SNIC 的紧密度参数  $m$  设置为 10。模型 batch size 设置为 128, 初始学习率为 0.001, 衰减速率为  $1 \times 10^{-6}$ 。网络训练迭代次数为 300 次, 网络输入大小为  $11 \times 11 \times 20$ 。

表 4 WHU-Hi-Longkou 数据集的地物类别和样本数

Table 4 Land cover classes and numbers of samples in WHU-Hi-Longkou dataset

No.	Class name	Numbers of samples
1	Corn	34511
2	Cotton	8374
3	Sesame	3031
4	Broad-leaf soybean	63212
5	Narrow-leaf soybean	4151
6	Rice	11854
7	Water	67056
8	Roads and houses	7124
9	Mixed weed	5229
Total		204542

为了验证所提策略在分类方面的高效性, 与以下策略进行对比: SVM、PCA-SVM、MFF-SVM、3DCNN<sup>[8]</sup>、PCA-3DCNN、MFF-3DCNN、PCA-HybridSN<sup>[10]</sup>。其中, 带有 PCA 的分类策略表示使用相同 20 维的 PCA 降维数据集, 用于对比所提 MFF 算法的有效性。PCA-HybridSN 和所提策略 MFF-HCCN 使用相同大小的 batch 和网络输入。SVM 使用径向基函数, 惩罚系数为 100。实验分类结果的评估准则使用总体精度(overall accuracy, OA)、平均精度(average accuracy, AA) 和 Kappa 系数进行评估。

### 3.3 仿真结果分析

在 Indain Pines 分类实验中, 随机选取数据集的 10% 和 90% 分别作为训练样本和测试样本, 仿真结果如表 5 所示。首先对比 SVM 和 PCA-SVM 以及 3DCNN 和 PCA-3DCNN 分类情况, 易见未降维的分类策略所获得的分类精度优于使用 PCA 降维的分类策略。这是由于 Indain Pines 数据集具有低的空间分辨率和高的光谱分辨率, “同谱异物, 异物同谱” 现象较为严重, 所以此时大多分类策略主要利用数据集光谱维波段的差异性来对地物像元进行精准分类。而 PCA 线性降维破坏了光谱波段间的非线性相关关系, 降低了像元间的光谱差异性, 导致分类精度下降。与此同时, 通过对比还可以发现使用 3DCNN 分类策略优于 SVM

分类策略,这是因为3DCNN在进行卷积操作时能够有效地提取局部空间特征,从而进一步提高分类精度。接着使用MFF-SVM和MFF-3DCNN分类策略与前面所述的分类策略进行对比,可以发现相比于PCA降维和未降维的分类策略,后两种分类策略的OA、AA和Kappa都获得了显著的提升。因为所提的MFF算法使用SNIC算法将相似且相邻的地物像元划分到多个小区域,并对每个超像素进行均值化从而提高了同类地物像元的相似性和异物地物像元的差异性。而且MFF算法又对均值化的图像融合多种光谱特征进一步提高了地物像元的差异性,减少了SNIC误分割带来的影响。最后将所提MFF-HCCN分类策略与最新的PCA-HybridSN分类策略相比,可见所提策略在OA、AA和Kappa精度上分别提升了2.472%、1.924%和2.814%,表明了所提策略能够有效地利用空谱特

征。这是因为相比于PCA-HybridSN分类策略,所提的MFF-HCCN分类策略在二维卷积层使用了3种不同尺度的二维卷积核。通过对三维卷积层输出的特征图多尺度的局部特征提取,从而提高了对空谱特征的利用率。与此同时,所提分类策略又使用胶囊网络作为输出层,通过对二维卷积层输出的特征转换为具有大小和方向的胶囊向量作为输出,可以提高对地物像元的判别能力。当然,在训练时间方面,所提策略相对于3DCNN和HybridSN网络使用较多的卷积层和胶囊网络,复杂度较高,所以在相同20维光谱特征下,所提分类策略训练时间较长。图6为Indian Pines数据集的伪彩色图和其真实地物标记图。图7展示了不同分类策略在10%训练样本下的分类结果,可以发现所提方法分类效果最佳,而且使用MFF算法的分类策略相比其他策略图像更为平滑。

表5 各算法在Indian Pines数据集10%训练样本下的分类结果比较

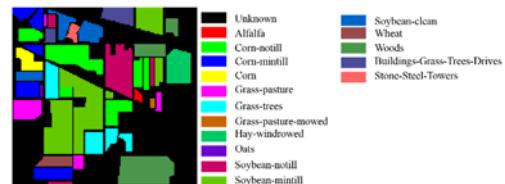
Table 5 Comparison of the classification results of each algorithm under 10% training samples of the Indian Pines dataset

	SVM	PCA-SVM	MFF-SVM	3DCNN	PCA-3DCNN	MFF-3DCNN	PCA-Hybrid SN	MFF-HCCN
OA(%)	80.369	65.431	89.442	91.09	86.101	92.926	96.758	99.230
AA(%)	75.027	54.874	86.284	90.443	80.647	94.204	95.871	97.795
Kappa×100	77.493	59.306	87.964	89.743	84.134	91.943	96.309	99.123
Train times/s				590.3	275.2	275.2	248.2	733.4



(a) 伪彩色图

(a) Pseudo-color image

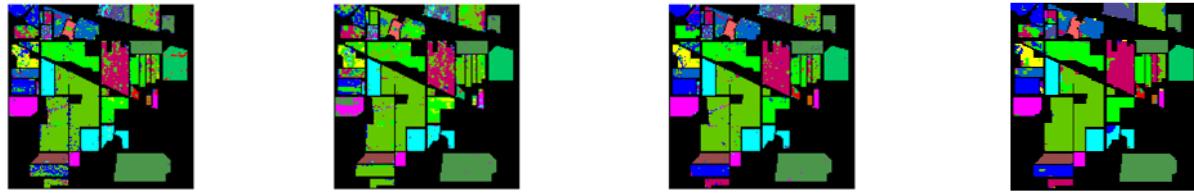


(b) 真实地物标记图

(b) Real ground object marking image

图6 Indian Pines 伪彩色图及其标记图

Fig.6 Pseudo-color map of Indian Pines and its marker map

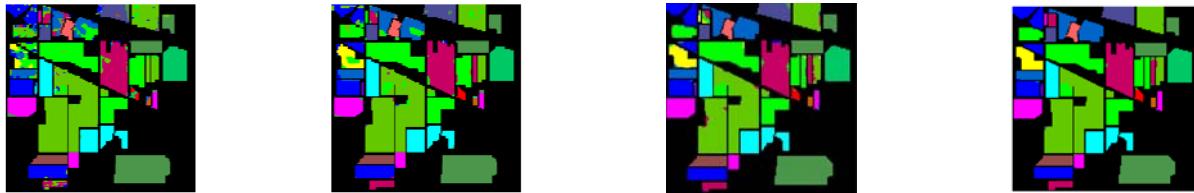


(a)SVM

(b)PCA-SVM

(c)MFF-SVM

(d)3DCNN



(e)PCA-3DCNN

(f)MFF-3DCNN

(g)PCA-Hybrid SN

(h)MFF-HCCN

图7 各算法在Indian Pines数据集10%训练样本下分类图像

Fig.7 Each algorithm classifies images under 10% of the training samples in the Indian Pines dataset

在 Pavia University 分类实验中,选取数据集的2%和98%分别作为训练集和测试集,仿真结果如表6所示。易知在 Pavia University 数据集下,所使用分类策略的精度相比于 Indian Pines 数据集时都表现出更好的分类效果,这是因为 Indian Pines 数据集存在数据不平衡问题,而 Pavia University 则有着较高的空间分辨率和较少的地物类别。首先对比 SVM、PCA-SVM、3DCNN 和 PCA-3DCNN 四种分类策略,可以发现 PCA-3DCNN 分类精度最高,PCA-SVM 分类精度最低。这是由于 3DCNN 能够同时提取空间和光谱特征,提高了地物像元类别的判断能力。然而光谱的高维性也使得了 3DCNN 的空间特征和光谱信息不能较好地结合,适当地降维提高网络对光谱特征的提取,因此 PCA-3DCNN 获得了更好的分类效果。然后

将 MFF-SVM 和 MFF-3DCNN 策略分别与 SVM 和 PCA-3DCNN 策略对比,发现使用 MFF 算法的分类策略在分类精度方面都得到显著提升,证明了基于 SNIC 分割的 MFF 算法可以有效地融合空谱特征。最后将所提 MFF-HCCN 策略与 PCA-HybridSN 策略对比,发现前者可以取得更好的分类精度,其 OA、AA 和 Kappa 系数相比于 PCA-HybridSN 策略分别提高了 1.312%、1.317% 和 1.743%。表明所提 MFF-HCCN 能够更好地对高分辨率、高光谱图像进行地物分类,其主要原因是该策略通过聚类融合、卷积提取和胶囊判别有效地挖掘了地物的空谱特征。图 8 和图 9 展示了使用不同分类策略在 2% 训练样本下的分类结果图,可以发现所提方法图像更为平滑,误分类较为集中,这是由于使用超像素分割和均值化的结果。

表 6 各算法在 Pavia University 数据集 2% 训练样本下分类结果比较

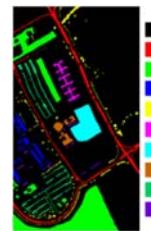
Table 6 Comparison of the classification results of each algorithm under 2% training samples of the Pavia University dataset

	SVM	PCA-SVM	MFF-SVM	3DCNN	PCA-3DCNN	MFF-3DCNN	PCA-HybridSN	MFF-HCCN
OA(%)	91.913	78.147	95.274	93.338	95.840	97.674	97.941	99.253
AA(%)	88.899	64.517	93.946	90.960	94.157	96.386	97.304	98.621
Kappa × 100	89.204	69.417	93.699	91.137	94.475	96.904	97.267	99.010
Train times/s	-	-	-	366.6	256.2	256.2	212.9	574.4



(a) 伪彩色图

(a) Pseudo-color image

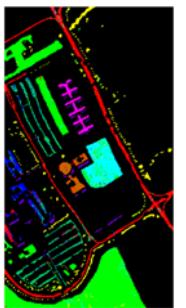


(b) 真实地物标记图

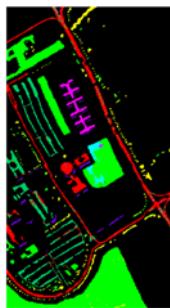
(b) Real ground object marking image

图 8 Pavia University 伪彩色图及其标记图

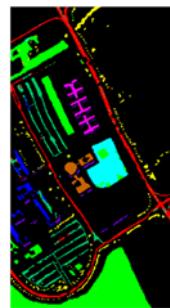
Fig.8 Pseudo-color map of Pavia University and its marker map



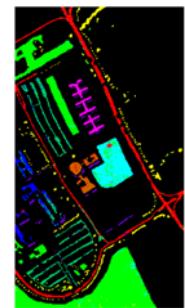
(a) SVM



(b) PCA-SVM



(c) MFF-SVM



(d) 3DCNN

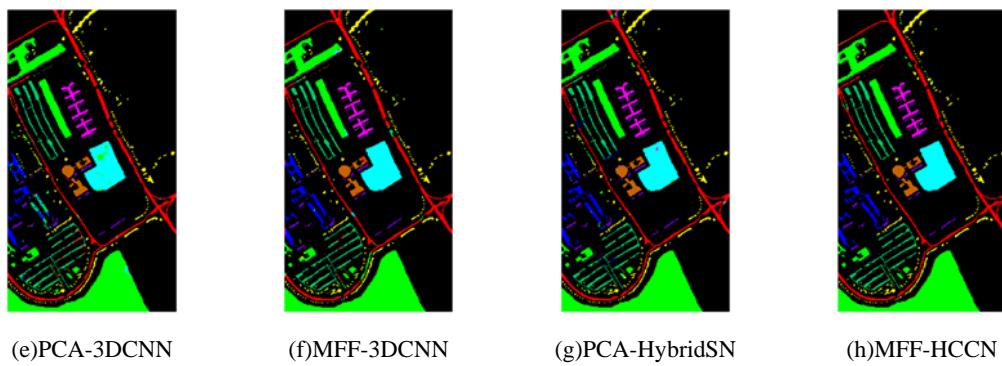


图9 各算法在 Pavia University 数据集 2%训练样本下分类图像

Fig.9 Each algorithm classifies images under 2% of the training samples in the Pavia University dataset

在 WHU-Hi-Longkou 分类试验中,随机选取数据集的 0.5% 和 99.5% 作为训练集和测试集,仿真结果如表 7 所示。由表 7 可知,所提分类策略的总体分类精度分别优于 PCA-3DCNN 和 PCA-HybridSN 分类。同时,通过对比也可以发现,WHU-Hi-Longkou 数据集是一种高空间分辨率和高光谱分辨率的影像,相比传统基于像元的 SVM 分类策略,基于卷积神经网络的分类策略在分类时由于融合了空间信息,因此分类精度普遍较高。对比 MFF-SVM 和 SVM 分类策略,可以发现,MFF 预处理方法使用 SNIC 分割方法在一定程度上融合了一定程度的空间信息,使得同类像元相似性增大,异类像元差异性加大,更易于 SVM 分类器判别。图 10 为 WHU-Hi-Longkou 数据集的伪彩色图

表7 各算法在 WHU-Hi-Longkou 数据集 0.5%训练样本下分类结果比较

Table 7 Comparison of the classification results of each algorithm under 0.5% training samples of the WHU-Hi-Longkou dataset

	SVM	PCA-SVM	MFF-SVM	3DCNN	PCA-3DCNN	MFF-3DCNN	PCA-HybridSN	MFF-HCCN
OA(%)	95.036	88.933	98.032	94.233	98.336	98.759	98.716	99.024
AA(%)	83.133	58.389	94.894	83.464	98.736	96.309	97.233	97.278
Kappa × 100	93.437	85.388	97.400	92.409	95.163	98.369	98.315	98.718
Train times/s	-	-	-	491.2	382.1	382.1	353.9	816.3

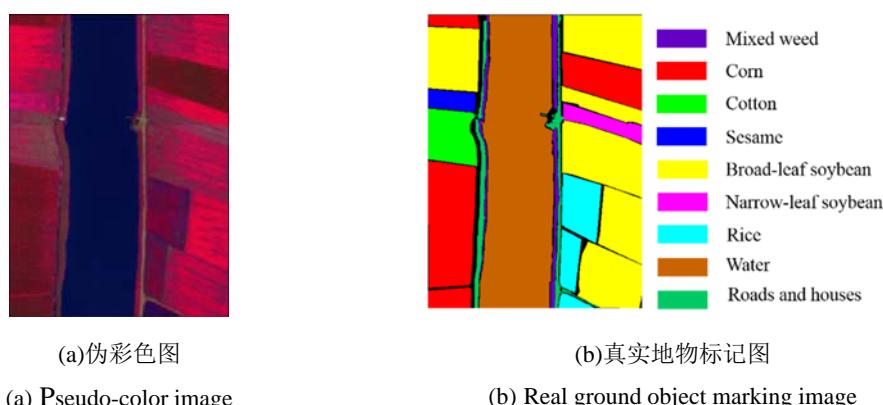


图10 WHU-Hi-Longkou 伪彩色图及其标记图

Fig.10 Pseudo-color map of WHU-Hi-Longkou and its marker map

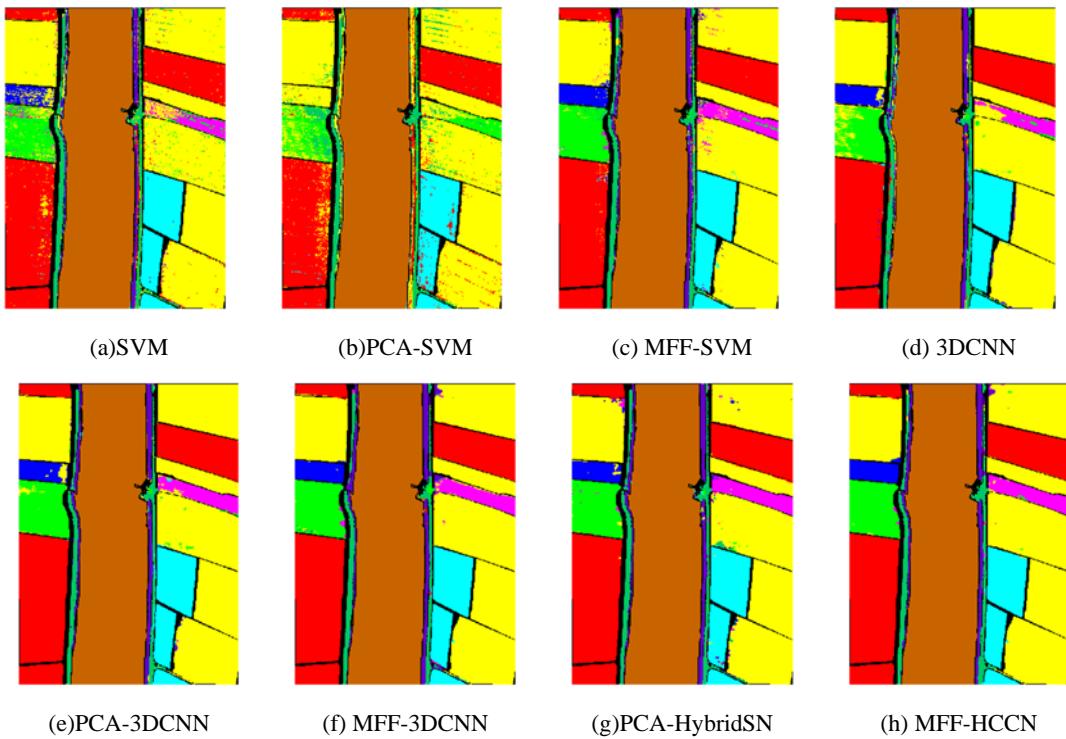


图 11 各算法在 WHU-Hi-Longkou 数据集 0.5% 训练样本下分类图像

Fig.11 Each algorithm classifies images under 0.5% of the training samples in the WHU-Hi-Longkou dataset

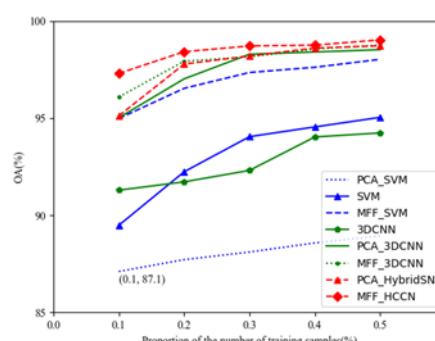
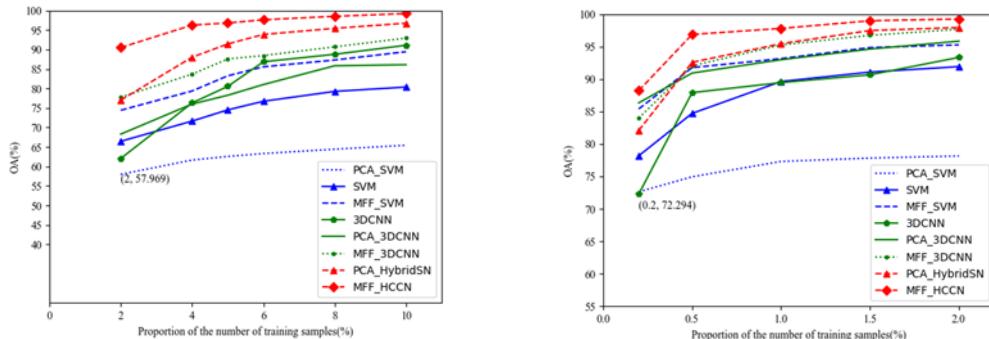


图 12 不同训练样本下的总体分类精度 OA 曲线图

Fig.12 OA curves of overall classification accuracy under different training samples

度,如在 Indian Pines 数据集下仅使用 2% 的训练样本仍能够获得 90.503% 的总体分类精度, 分别比 SVM、PCA-SVM、MFF-SVM、3DCNN、PCA-3DCNN、MFF-3DCNN 和 PCA-HybridSN 高 24.052%、32.534%、16.088%、28.452%、22.205%、12.715% 和 13.48%。与此同时,当采用 Pavia University 数据集的 0.2% 样本作为训练样本时,所提策略的总体精度为 88.217%,相比于其他分类策略分别提高了 10.041%、15.588%、2.82%、15.923%、1.886%、4.229%、6.143%。而对于 WHU-Hi-LongKou 数据集,仅使用数据集的 0.1% 作为训练样本,总体分类精度可达 97.296%,明显优于其他分类策略。可见,所提策略在小样本场景下依然适用,这是因为使用了 SNIC 分割算法、多尺度二维卷积核和胶囊网络能有效提升对图像局部空间特征的提取能力,从而能实现小样本情况下高光谱图像分类精度的提高。

## 4 结论

针对现有高光谱图像分类方法中空间和光谱利用率不足的问题,提出了一种多特征融合下基于混合卷积胶囊网络的高光谱图像分类策略。首先针对传统 PCA 降维未能利用空间特征,设计了基于超像素的 PCA 和 NMF 联合降维的 MFF 算法,以有效地将空间特征融入到降维的数据集中。同时通过将多种特征相融合来降低误分割带来的误差,以及避免超像素内像元均值化导致像元光谱维过于简单产生卷积网络过拟合的现象。然后,设计一种多核的混合卷积分类网络,通过使用多尺度二维核对三维卷积输出的特征图进行特征提取,进一步提高了对空谱特征的利用。最后,使用胶囊对卷积网络输出的局部特征进行胶囊变换,输出具有大小和方向的胶囊,以提高不同类别像元的判别精度。实验在传统的 Indian Pines 和 Pavia University 公共数据集下对比所提分类策略与传统分类策略分类情况,结果表明,该策略在总体精度、平均精度和 Kapp 系数方面均表现最佳,而且当使用较少训练样本时所提策略依然能够取得较高的分类精度,证明了所提策略利用空谱特征提高分类精度的有效性。与此同时,为了体现所提分类策略的应用价值,在近年精准农业方面的 WHU-Hi-LongKou 数据集上进行仿真,实验证明该分类策略具有良好的分类性能和泛化能力。

## 参考文献:

- [1] 蓝金辉,邹金霖,郝彦爽,等.高光谱遥感图像混合像元分解研究进展[J].遥感学报,2018,22(1): 13-27.
- [2] 徐金环,沈煜,刘鹏飞,等.联合核稀疏多元逻辑回归和 TV-L1 错误剔除的高光谱图像分类算法[J].电子学报,2018,46(1): 175-184.
- [3] XU J, SHEN Y, LIU P, et al. Hyperspectral image classification combining kernel sparse multinomial logistic regression and TV-L1 error rejection[J]. *Acta Electronica Sinica*, 2018, 46(1): 175-184.
- [4] 刘启超,肖亮,刘芳,等. SSC DenseNet: 一种空-谱卷积稠密网络的高光谱图像分类算法[J].电子学报,2020,48(4): 751-762.
- [5] LIU Q, XIAO L, LIU F, et al. SSC DenseNet: a spectral-spatial convolutional dense network for hyperspectral image classification[J]. *Acta Electronica Sinica*, 2020, 48(4): 751-762.
- [6] LI S, SONG W, FANG L, et al. Deep learning for hyperspectral image classification: an overview[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(9): 6690-6709.
- [7] Nanjun H, Paoletti M E, Mario H J, et al. Feature extraction with multiscale covariance maps for hyperspectral image classification[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2018, 57(2): 755-769.
- [8] GAO H, LIN S, LI C, et al. Application of hyperspectral image classification based on overlap pooling[J]. *Neural Processing Letters*, 2019, 49(3): 1335-1354.
- [9] YU C, ZHAO M, SONG M, et al. Hyperspectral image classification method based on CNN architecture embedding with hashing semantic feature[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12(6): 1866-1881.
- [10] YING L, Haokui Z, QIANG S. Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network[J]. *Remote Sensing*, 2017, 9(1): 67-88.
- [11] HE M, LI B, CHEN H, et al. Multi-scale 3D deep convolutional neural network for hyperspectral image classification[C]//*IEEE International Conference on Image Processing*, 2017: 3904-3908.
- [12] Roy S K, Krishna G, Dubey S R, et al. Hybrid SN: exploring 3D-2D CNN feature hierarchy for hyperspectral image classification[J]. *IEEE Geoscience and Remote Sensing Letters*, 2020, 17(2): 277-281.
- [13] WEI W, ZHANG J, LEI Z, et al. Deep cube-pair network for hyperspectral imagery classification[J]. *Remote Sensing*, 2018, 10(5): 783-801.
- [14] Baisantry M, SAO A K. Band selection using segmented PCA and component loadings for hyperspectral image classification[C]//*IEEE International Geoscience and Remote Sensing Symposium*, 2019: 3812-3815.
- [15] ZHANG W, FU K, SUN X, et al. Joint optimisation convex-negative matrix factorisation for multi-modal image collection summarisation based on images and tags[J]. *IET Computer Vision*, 2018, 13(2): 125-130.
- [16] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(11): 2274-2282.
- [17] Achanta R, Süsstrunk S. Superpixels and polygons using simple non-iterative clustering[C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 4895-4904.
- [18] Sabour S, Frosst N, Hinton G E. Dynamic routing between capsules[C]//*Proceedings of the 2017 International Conference on Neural Information Processing Systems*, 2017: 3856-3866.
- [19] ZHONG Y, WANG X, XU Y, et al. Mini-UAV-Borne hyperspectral remote sensing: from observation and processing to applications[J]. *IEEE Geoscience and Remote Sensing Magazine*, 2018, 6(4): 46-62.