

课程论文题目

周宇鑫

摘要

遥感传感器的快速发展使得高光谱图像的采集、分析和应用越来越广泛。然而，有限的样本集、高维特征、高度相关的波段以及混合的光谱信息给 HSI 的分类带来了很大的挑战。本文提出了一种基于卷积神经网络的无监督多尺度多元特征学习 (UMsDFL) HSI 分类方法，该方法深入考虑了 HSI 的空间光谱特征。具体来说，采用简单非迭代聚类 (SNIC) 算法并启发式计算超像素大小后，将 hsi 分割成超像素进行特征学习。无监督网络采用卷积编码器和解码器、附加聚类分支和多层特征融合设计，增强了特征学习的可分辨性和特征映射的可重用性。然后，通过无监督网络协同学习大尺度和小尺度背景下的空间关系和对象属性，利用互补的多尺度特征。此外，通过无监督网络同时学习高光谱信息和非下采样 contourlet 变换 (NSCT) 纹理的不同特征，以缓解几何表示的不足。最后，在 UMsDFL 的基础上，采用随机森林 (random forest, RF) 作为综合分类器，并采用超像素正则化对分类结果进行优化。在三个真实 HSI 数据集上进行了一系列实验，以证明我们的 UMsDFL 方法的有效性。实验结果表明，当每类只有 5 个样本进行训练时，所提出的 UMsDFL 在 Houston、Pavia 和 Dioni 数据集上的整体准确率分别为 79.23%、96.49% 和 77.26%。

关键词：卷积神经网络；特征融合；高光谱图像；超像素分割；无监督特征学习

1 引言

随着遥感观测技术的快速发展，高光谱图像的采集、分析和应用越来越广泛。隐藏在 HSI 狭窄光谱范围内的地物的许多特征正在逐渐被发现。与多光谱图像相比，HSI 可以获得数百个连续光谱波段的丰富信息，增强了特征提取和目标识别能力^[1]。高光谱遥感在土地利用与土地覆盖 (LULC) 分类、目标探测、农业监测、矿产测绘、环境管理和国防等领域发挥着重要作用^[2]。然而，有限的样本集、高维特征、高度相关的波段以及混合的光谱信息给 HSI 分类带来了巨大的挑战。迄今为止，人们已经开发了多种 HSI 特征提取和土地覆盖分类方法。早期基于像素的方法使用光谱信息和简单特征进行识别，使用分类器如 k 最近邻 (KNN)^[3]、极限学习机 (ELM)^[4] 和支持向量机 (SVM)^[5]。随后，提出了基于图嵌入^[6]、稀疏表示^[7] 和低秩表示^[8] 的方法来增强高光谱识别能力。但是，存在着相似物体不同光谱和不同物体相似光谱的现象，仅利用光谱信息很难区分混淆的物体。因此，提出了将光谱值与空间结构相结合的空间光谱方法。在空间处理单元上，采用图构造^[9]、形态分割^[10]、超像素分割^[11] 等方法将图像分割为有意义的斑块，用于后续分析。

在本文中，我们提出了一种无监督多尺度多元特征学习 (UMsDFL) 的 HSI 分类方法，该方法深入考虑了通过 CNN 获取的空间光谱特征，如图 1 所示。在我们的 UMsDFL 中，HSI 首先使用简单非迭代聚类 (SNIC) 算法分割成超像素，这是对简单线性迭代聚类 (SLIC) 算法的改进。SNIC 具有计算复杂度低、分割效果好等优点，并对其进行了改进以兼容高光谱输入。另一方面，将主成分分析 (PCA) 和非下采样 contourlet 变换 (NSCT) 依次用于 HSI 的降维和几何纹理提取。将 SNIC 分割边界应用于 HSI 和 NSCT 图像，得到 HSI 超像素和 NSCT 超像素，用于不同的深度特征学习。在分割超像素的基础上，

综合设计了卷积编码器和解码器、附加聚类分支和多层特征图组合,构建了一个有效的无监督空间光谱 CNN 网络,用于特征学习。解码和 k 均值聚类的两个分支分别用于图像重建和特征识别,然后进行后向误差传播。通过全局平均池化 (GAP) 和自适应加权拼接实现多层特征图组合,提高特征的重用性和完整性。对于多尺度特征学习,我们提出通过无监督网络学习小尺度和大尺度背景下超像素的对象属性和空间关系。多尺度特征学习在 HSI 超像素和 NSCT 超像素上并行进行,有利于在多环境下表示对象和模式的互补信息。对于多元特征学习,我们设计通过无监督网络学习 HSI 超像素的光谱属性和 NSCT 超像素的几何结构。contourlet 纹理的深层特征强调物体不同层次、不同方向的轮廓进行几何表示,弥补了 CNN 的不足。最后,考虑到随机森林 (random forest, RF) 算法在训练样本不平衡、训练样本有限的情况下仍能获得稳定的性能,将其作为 UMsDFL 之后的综合分类器。采用超像素正则化对像素分类结果进行优化,提高了对应于超像素分割和基于超像素的特征学习的边界连续性。

2 相关工作

2.1 深度学习应用于 HSI 分类

深度学习 (DL) 方法已被广泛研究并用于 HSI 分类,如卷积神经网络 (cnn),循环神经网络 (RNNs),堆叠自编码器 (sae),图卷积网络 (GCNs) 和生成对抗网络 (GANs)。深度学习方法可以通过网络逐层自动学习 HSI 中从具体到抽象的深层特征,从而增强语义表达能力。其中,CNN 方法由于其固有的卷积和特征图聚合的特征学习能力,应用最为广泛。Yu 等人改进了 CNN 中的参数优化,并将其应用到 HSI 分类中,缓解了高相关性波段和训练样本不足的问题。此外,Mei 等提出了一种使用三维卷积自编码器的无监督空间光谱特征学习方法,获得了比其他监督算法更好的分类精度。为了促进样本选择,Hu 等人引入了主动学习策略,构建有价值的样本集用于 CNN 训练,促进特征提取。为了结合不同 DL 模型的优点,Yue 等分别将 SAE 和 CNN 获得的光谱特征和空间特征合并,并采用空间金字塔池化方法接受不一致尺度的输入。同样,Hao 等人采用了堆叠去噪自编码器和 CNN 分别对特征进行编码,并通过自适应类权重融合多个分支的特征。CNN 方法在 HSI 分类中的另一种扩展是设计不同的有效网络结构来强调信息提取,提高识别精度,如残差 CNN、基于注意的 CNN、密集连接的 CNN。然而,基于 cnn 的 HSI 分类方法通常是有监督的训练,需要大规模的样本来优化大量的参数。人工标记样本的获取需要花费大量的时间和专业知识,导致样本集小,网络训练不足的问题。

2.2 自编码器无监督学习特征

自编码器的无监督方法通过图像编码和重建从未标记的样本中学习特征,而最大化分类判别的目标不明确,降低了表达能力。此外,CNN 中从低到高的多层特征映射利用不足,损失了部分 HSI 特征,限制了学习性能。从上下文角度看,邻域的光谱属性和空间关系对地物识别很重要,多尺度上下文反映了不同层次的结构信息。小尺度和大尺度背景分别侧重于互补的内部特征和外部关系,但无监督的多尺度特征学习存在挑战。另一方面,纹理信息是 HSI 分类的导电性特征描述,可以反映地物的空间分布和变化规律。利用光谱波段的 CNN 方法提取的是光谱属性的高级语义特征,而不是几何纹理,不足以进行全面和判别特征表示。通过深度网络整合多种特征学习对于提高性能具有重要意义,但如何在无监督框架下协同完成纹理学习和其他特征学习仍有待研究。

3 本文方法

3.1 本文方法概述

本文提出了一种有效的无监督空间光谱 CNN 框架，用于 HSI 特征学习和土地覆盖分类。通过解码器和聚类两个分支，在图像重建和伪标签分类的误差反馈下，对网络进行迭代训练和优化，从未标记样本中学习特征，提高类别间的特征识别能力。通过 GAP 和自适应加权拼接，将多层特征图信息进行可区分聚类，重用低、中、高层特征，增强内在内聚性。

为了利用互补的多尺度特征，本文设计了通过无监督空间光谱 CNN 学习小尺度和大尺度背景下的目标属性和空间关系。小尺度特征主要描述物体的高光谱特性和局部结构，大尺度特征表征物体之间的空间关系和分布模式。无监督多尺度深度特征学习有效地提取了不同规模、不同视角的无标记样本中丰富的上下文信息。

为了缓解几何表示的不足，本文提出通过无监督空间光谱 CNN 学习高光谱信息和轮廓波纹理的深层特征。光谱特征侧重于物体的光学吸收和反射特征，纹理特征侧重于空间结构和分布的几何模式。无监督多样性特征学习有利于提高 HSI 特征表达的全面性，挖掘无标记样本的内涵。

如图 1 所示，我们的 UM sDFL 方法有五个主要步骤。首先，利用改进的 SNIC 算法将 HSI 分割成大小相对均匀、形状规则的超像素，作为后续特征学习的基本单元。其次，构建具有聚类分支和多层组合的高效无监督空间光谱 CNN 进行超像素特征学习，提高特征的可分辨性；第三，通过无监督 CNN 学习对象属性和空间关系在小尺度和大尺度下的多尺度深度特征。第四，以无监督的方式分别提取高光谱信息和轮廓波纹理的不同深度特征。最后，基于多尺度、多样化的无监督特征学习，利用有限的标记样本训练射频分类器，获得分类图。采用超像素正则化对分类结果进行优化，取得更好的性能。

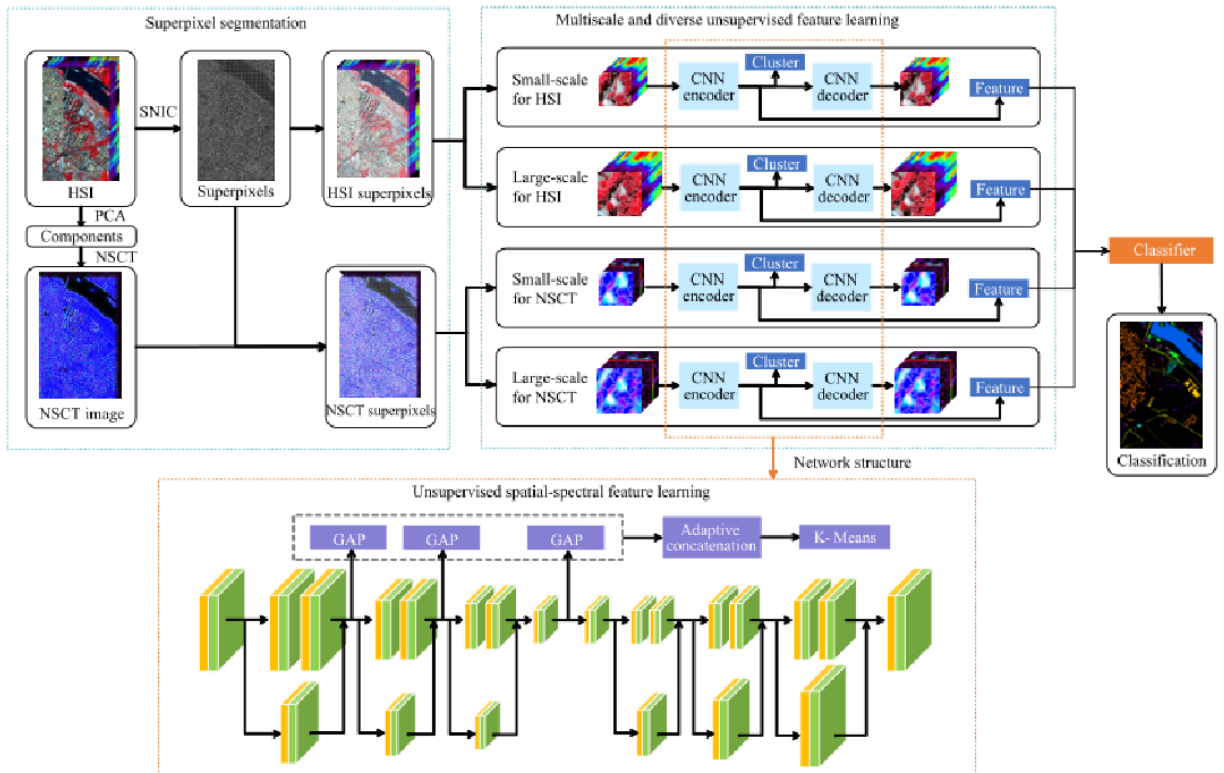


图 1: UM sDFL 方法流程图

3.2 超像素分割

SNIC 算法是 SLIC 算法的改进，不仅继承了 SLIC 容易实施、计算高效且能控制超像素紧密程度和数量的优点，而且还是非迭代的，拥有更快，占用内存更少的特点。此外，SNIC 在一开始就强制了连接性，这也适合对 HSI 进行处理。针对 HSI 的特点，作者将每个像素对中心的归属度进行了修改，从原来的空间坐标改成了空间-光谱坐标，如下式所示

其中， X 和 S 分别为空间坐标向量和光谱向量，而 w_x 和 w_s 分别代表空间和光谱距离的权重。首先，从初始中心点开始，SNIC 算法利用优先队列选择下一个加入聚类的像素。优先队列中充满了相互连接的候选像素，然后弹出距离最小的像素。然后，根据新加入的像素，对相应的中心进行在线更新。由于影像具有局部自相似性，通过简单迭代能有效实现在线更新。在 SNIC 分割后，HSI 被分割成拥有相对均匀大小和整齐形状的超像素，作为接下来特征学习和分类结果正则化的基本单元。

3.3 非监督空谱特征学习

主要由 4 部分构成：卷积编码器，解码器，聚类分支和自适应权重融合。

3.3.1 编码器与解码器

兼顾考虑学习能力和模型的简单程度，编码器由 5 个块构成，分别拥有 1, 2, 2, 2, 1 个卷积层。值得注意的是，作者这里用的是可分离卷积。其中第 2, 3, 4 块后接最大池化和残差连接结构，以压缩特征图并提高学习效率。在第五个块使用全局均值池化（GAP）将特征图编码为特征向量，作为潜在空间表达。而解码器结构与编码器对称，其最后一层的卷积数设为输入 HSI 的维度，通过重建误差反向传播进行训练。

3.3.2 聚类分支

为进一步增加潜在空间特征的区分度，利用 K-means 分支，对 GAP 编码的特征向量进行非监督聚类，然后将聚类结果作为伪标签，通过伪标签分类误差进一步优化编码器参数。详细来说，将特征向量进行聚类后，将聚类类别作为每个超像素的伪标签，通过一个分类器对特征向量进行分类，用分类结果与伪标签的差异作为损失函数，利用梯度反向传播调整网络参数。在训练过程中，利用 Adam 优化器对学习率进行调整。

3.3.3 自适应特征融合

在非监督网络中，随着编码器层数加深，提取到的特征由具体变得抽象。虽然抽象的特征一般更具判别性，但对于一些需要需要细微特征辨别的地物目标（如颜色或纹理），具体特征更为重要。由此可见，多层次特征图是相互补充的，从不同角度对地物进行了描述。因此作者将第 2, 3, 5 个块的输出特征图结合起来，分别包含了编码器中低、中、高层次的信息。考虑到可能存在表达冗余和高维空间中的聚类困难，通过 GAP 将特征图进行压缩，再将多层特征图连接起来，作为 k-means 分支的输入。由于对于不同地物类别，不同层特征的重要性并不相同，因此通过可学习权重自适应地调整不同层特征的重要程度，将组合的多层特征向量输入聚类分支，通过误差反向传播，对编码器参数进行优化。在训练结束后，只提取第 5 个块的特征图作为最终分类器的输入。

3.4 多尺度特征学习

传统的 CNN 利用固定大小的接受域和滑动卷积核逐层提取图像深度特征。然而, HSI 地物在大小、形状和空间关系上表现出较大的异质性, 具有明显的多尺度特征。在 CNN 中, 固定大小的接受域限制了观察环境, 无助于捕捉尺度相关的信息, 从而降低了性能。因此, 有必要整合空间域中的多个上下文信息, 进行多尺度无监督深度特征学习。如图 1 所示, 我们设计了小分支和大分支来提取不同上下文下的对象属性和空间模式的无监督特征, 由于更多上下文下的特征冗余和模型复杂性, 不采用更多的尺度分支。对于小尺度分支, 无监督的空间光谱 CNN 主要学习物体的内部高光谱和结构特征。对于大型分支, 该网络主要提取对象的外部环境和关联关系。在本研究中, 将两个分支的上下文大小分别设置为 32×32 和 64×64 , 经实践验证, 该方法适用于多尺度特征提取。 32×32 和 64×64 背景分别有利于提取地物的局部属性和分布模式, 对空间特征表示具有互补作用。更精确地说, 对于类边界内的超像素, 邻域信息是相对统一和一致的, 特征学习需要一个大范围的上下文来实现更好的识别。对于跨类边界的超像素, 邻域更加复杂和不一致, 需要一个小范围的上下文来避免引入干扰噪声。因此, 多尺度特征学习通过无监督的空间-光谱 CNN 获取超像素在不同背景下的深度特征, 利用未标记样本的丰富信息, 在小样本集的情况下提高性能。

3.5 多样化特征学习

为了提高无监督学习的全面性和多样性, 设计了高光谱信息和几何纹理的深度特征进行学习和融合。考虑到直接堆叠多个输入的执行时间较短, 但不利于有针对性的信息提取和参数训练, 为了更好的区分和分类, 在两个分支中分别学习不同的特征。对于具有典型高光谱特征的物体, 光谱分支的贡献更大; 对于几何图案显著的物体, 纹理分支的贡献更大。Contourlet 纹理是一个多方向、多尺度的变换, 它结合了拉普拉斯金字塔和框架金字塔中的定向滤波器组。随后, Cunha 等提出了一种名为 NSCT 的超完备 contourlet 方法, 该方法具有实现速度快、平移不变性、多尺度多方向扩展等优点。NSCT 适用于 HSI 处理, 提取和突出地物的几何结构, 其次是无监督的深度纹理特征学习。由于小波变换和 Gabor 变换在多方向和多尺度条件下的表现不充分, 因此它比小波变换和 Gabor 变换更有效、更合适。NSCT 可以在 HSI 中使用更少的系数捕获更多的边缘轮廓, 这对纹理描述和目标识别具有重要意义。具体而言, NSCT 的核心是一个不可分离的双通道非下采样滤波器组 (NSFB), NSCT 可分为两个移不变部分, 即多尺度非下采样金字塔 (NSP) 和非下采样定向滤波器组 (NSFB)。

Contourlet 变换利用多尺度、多方向子带提取几何信息, 逼近 HSI, 有利于纹理和边缘的检测。本研究采用 HSI 的 NSCT 分解作为几何表达式, 除高光谱波段外, 通过无监督 CNN 进行深度特征学习, 实现多样化表示。为了突出主要几何结构, 避免噪声干扰, 首先采用主成分分析得到降维后的数据。PCA 算法从原始数据空间中依次寻找一组相互正交的坐标轴, 减少 HSI 频带, 最大限度地减少信息损失。本研究采用 PCA 得到的前三个主成分进行 NSCT 分解和无监督特征学习, 保留了 98% 以上的原始 HSI 信息。平衡 HSI 表达的 NSCT 转换的模型复杂性和纹理冗余, 对每个主成分分别进行适当的层次分解, 得到分解后的图像。本研究将三主成分共 48 张分解图像进行叠加, 输入到无监督 CNN 中进行深度几何特征学习。此外, 基于 SNIC 算法利用 HSI 波段分割出的超像素提取 NSCT 训练补丁。同样, NSCT 分解图像也存在几何结构的多尺度特征。

为了综合高光谱和空间信息的优势, 将从 HSI 和 NSCT 斑块中获取的无监督深度特征以及原始光

谱值进行集合，得到多尺度、多样的特征，表示为其中为融合特征，用于从小范围和大范围上下文的多个视角描述第 i 个超像素，以及光谱和纹理表示，为原始光谱值。和分别表示 HSI 和 NSCT 超像素的多尺度特征。多尺度、多样化的特征学习通过无监督的空间光谱 CNN 获取超像素丰富的深度特征，利用未标记样本的潜在内容来提高性能。

4 复现细节

4.1 与已有开源代码对比

超像素分割代码是在 SNIC 算法开源代码的基础上做了针对高光谱图像的改进，在距离计算公式中加入更多通道的数据。非下采样轮廓波变换的代码也针对高光谱图像进行了改进，再提取出图像的纹理特征。原论文作者给出了深度聚类网络的核心代码，实验中参考了这部分代码，并且在特征融合部分进行了改动。

4.2 实验环境

操作系统为 64 位 windows10，开发平台为 pycharm 2022 x64，使用 pycharm 远程服务器功能连接至服务器，服务器 python 环境 3.8，显卡配置 P100，神经网络训练使用 tensorflow2.0 框架实现，使用 scikit-learn 工具包版本 1.02。部分实验如超像素分割和非下采样轮廓波变换使用 matlab2021b 实现。使用 arcgis 软件和 envi 软件对遥感图像进行处理。

5 实验结果分析

本部分对实验所得结果进行分析，详细对实验内容进行说明，实验结果进行描述并分析。对高光谱图像进行超像素分割，结果如图 2 所示。对 HSI 图像用 PCA 进行降维，再用 NSCT 算法提取纹理特征，结果如图 3 所示。

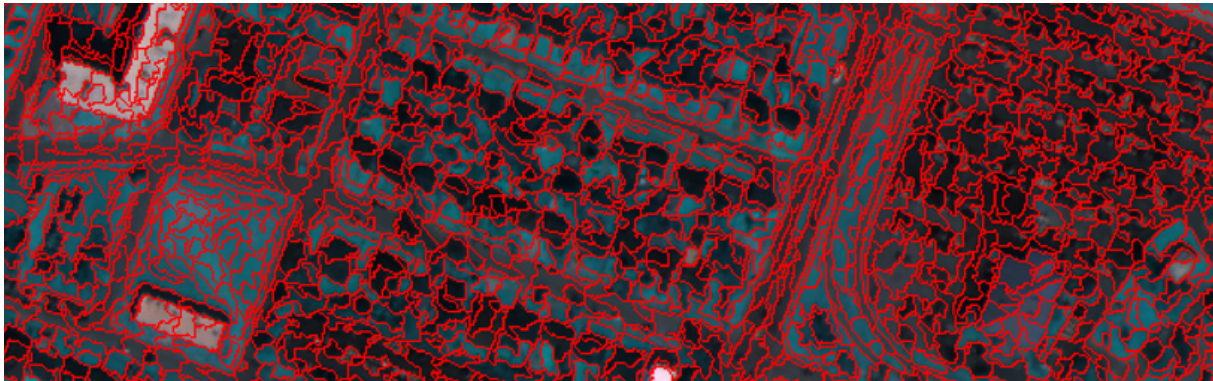


图 2: 超像素分割结果

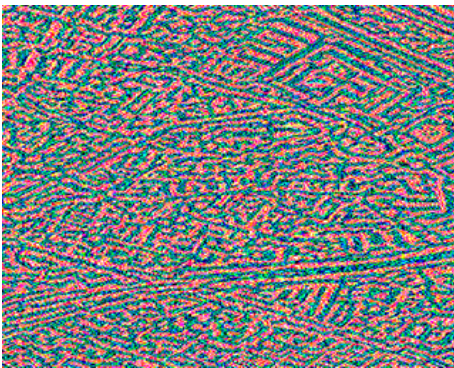


图 3: 提取的纹理特征

在自动编码器训练完成后，利用编码器将输入编码为特征向量，将多尺度和多样化特征向量拼接起来，通过随机森林分类器进行分类。为了解决预测图中椒盐效应的问题，利用超像素正则化，将整个超像素中类别最多的标签作为超像素的标签。训练后分别得到 hsi32,hsi64,nsct32,nsct64 四种数据，加上光谱特征数据，将这五种数据拼接起来放入随机森林分类器进行训练，随机森林设置 300 棵树。测试集为全图数据。将预测结果进行超像素正则化，再计算精度。以 Houston 数据集为例，OA 能达到 89.41%，Kappa 能达到 88.56%。

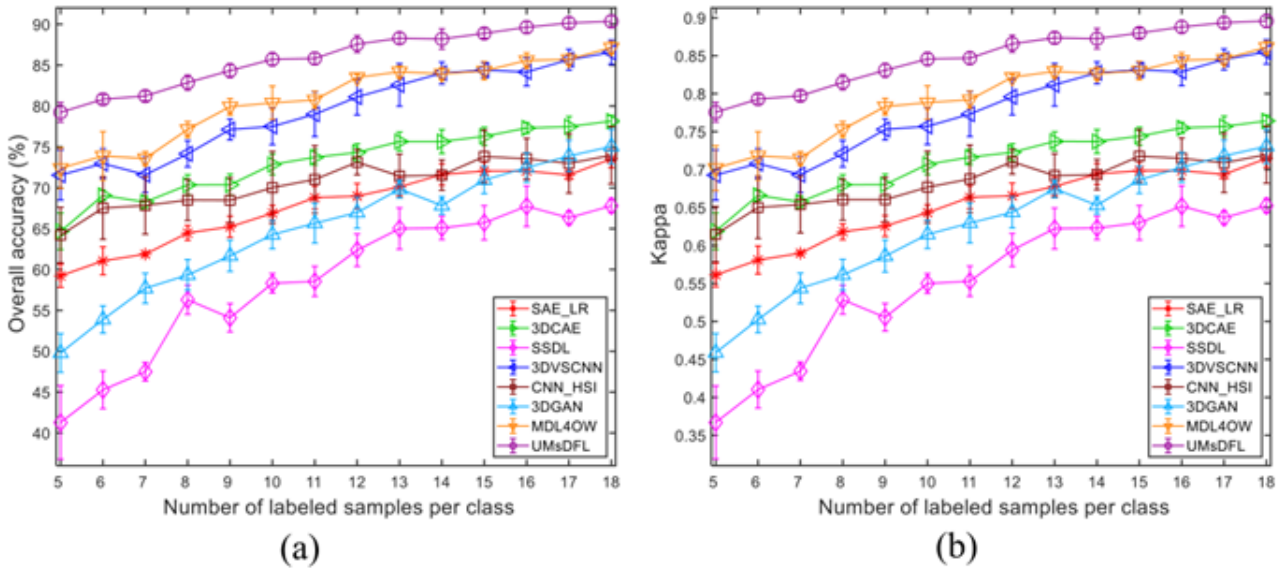


图 4: Houston 数据结果

6 总结与展望

本文提出了一种用于 HSI 分类的 UMDFL 方法。详细地说，在将改进的 SNIC 应用于 HSI 和超像素大小的启发式计算之后，通过无监督的空间光谱 CNN 来学习超像素的深度特征。该网络由卷积编码器和解码器、聚类分支和多层特征组合组成。然后，通过无监督 CNN 提取多尺度背景下的对象特征和空间关系，并协同提取高光谱信息和 NSCT 纹理的各种特征。最后，我们利用 RF 分类器融合多尺度和多样性特征，并使用小样本集获得综合分类图。采用超像素正则化来优化像素分类结果并获得良好的性能。实验结果一致地表明，UMDFL 的无监督空间光谱特征学习可以有效地提高特征表达和 HSI 分类精度。无监督 CNN 显示出优异的特征提取能力，多尺度和多样的特征学习有助于提高性能。

在未来进一步的研究中，考虑到多种特征是不同的和相互关联的，可以进一步探索注意力机制，以充分利用互补优势。对于综合分类，可以探索改进特征融合策略，来提升最终的分类结果。

参考文献

- [1] AUDEBERT N, LE SAUX B, LEFÈVRE S. Deep learning for classification of hyperspectral data: A comparative review[J]. IEEE geoscience and remote sensing magazine, 2019, 7(2): 159-173.
- [2] GHAMISI P, PLAZA J, CHEN Y, et al. Advanced spectral classifiers for hyperspectral images: A review [J]. IEEE Geoscience and Remote Sensing Magazine, 2017, 5(1): 8-32.
- [3] MA L, CRAWFORD M M, TIAN J. Local manifold learning-based k -nearest-neighbor for hyperspectral

- image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2010, 48(11): 4099-4109.
- [4] ZHOU Y, PENG J, CHEN C P. Extreme learning machine with composite kernels for hyperspectral image classification[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 8(6): 2351-2360.
- [5] MELGANI F, BRUZZONE L. Classification of hyperspectral remote sensing images with support vector machines[J]. IEEE Transactions on geoscience and remote sensing, 2004, 42(8): 1778-1790.
- [6] SHI L, ZHANG L, YANG J, et al. Supervised graph embedding for polarimetric SAR image classification[J]. IEEE Geoscience and Remote Sensing Letters, 2012, 10(2): 216-220.
- [7] TANG Y Y, YUAN H, LI L. Manifold-based sparse representation for hyperspectral image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(12): 7606-7618.
- [8] SUN W, YANG G, DU B, et al. A sparse and low-rank near-isometric linear embedding method for feature extraction in hyperspectral imagery classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(7): 4032-4046.
- [9] MA L, CRAWFORD M M, YANG X, et al. Local-manifold-learning-based graph construction for semisupervised hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 53(5): 2832-2844.
- [10] TARABALKA Y, CHANUSSOT J, BENEDIKTSSON J A. Segmentation and classification of hyperspectral images using watershed transformation[J]. Pattern Recognition, 2010, 43(7): 2367-2379.
- [11] FANG L, LI S, KANG X, et al. Spectral-spatial classification of hyperspectral images with a superpixel-based discriminative sparse model[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(8): 4186-4201.