

基于对比学习的高光谱和 LiDAR 地物分类方法

李峙含^{1,2,3,4}, 花海洋^{1,2*}, 张浩^{1,2,3,4}

¹中国科学院沈阳自动化研究所, 辽宁 沈阳 110016;

²中国科学院光电信息处理重点实验室, 辽宁 沈阳 110016;

³中国科学院机器人与智能制造创新研究院, 辽宁 沈阳 110169;

⁴中国科学院大学, 北京 100049

摘要 为了提高在少标签数据条件下的地物分类精度, 提出一种利用高光谱图像(HSI)数据与机载激光雷达(LiDAR)数据进行对比学习的地物分类方法。首先利用不带标签的HSI数据与LiDAR数据进行对比学习, 通过对比学习可以建立这两种数据的空间特征之间的联系, 实现对这两种数据的空间特征提取; 设计了卷积模块与Transformer模块相结合的网络, 使模型能够利用提取出的局部特征建立全局交互关系。在Houston 2013数据集和Trento数据集上进行对比实验, 所提方法的分类精度高于其他多源数据融合地物分类的对照方法, 在Houston 2013数据集上当每类标签样本量为5时, 所提方法的分类精度比对照方法提高20.73个百分点, 在Trento数据集上当每类标签样本量为2时, 所提方法的分类精度比对照方法提高8.35个百分点。

关键词 图像处理; 高光谱图像分类; 机载激光雷达; 地物分类; 多源数据

中图分类号 TP751

文献标志码 A

DOI: 10.3788/LOP230540

Classification Based on Hyperspectral Image and LiDAR Data with Contrastive Learning

Li Shihan^{1,2,3,4}, Hua Haiyang^{1,2*}, Zhang Hao^{1,2,3,4}

¹Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, Liaoning, China;

²Key Laboratory of Opto-Electronic Information Processing, Chinese Academy of Sciences, Shenyang 110016, Liaoning, China;

³Institute for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, Liaoning, China;

⁴University of Chinese Academy of Sciences, Beijing 100049, China

Abstract This study proposes a semi-supervised method using multimodality data with contrastive learning to improve the classification accuracy for hyperspectral images (HSI) and light and detection ranging (LiDAR) data in the case of a few labeled samples. The proposed method conducts contrastive learning using HSI and LiDAR data without labels, which helps to build the relationship between the spatial features of the two data. Thereafter, their spatial features can be extracted by the model. We designed a network combining the convolution and Transformer modules, which allows the model to extract the local features for establishing a global interaction relationship. We conducted experiments on contrastive learning on the Houston 2013 and Trento datasets. The results show that the classification accuracy of the proposed method is higher than that of other multisource data fusion classification methods. On the Houston 2013 dataset, the classification accuracy of the proposed method is 20.73 percentage points higher than that of the comparison method when the number of labeled samples is five. On the Trento dataset, the classification accuracy of the proposed method is 8.35 percentage points higher than that of the comparison method when the number of labeled samples is two.

Key words image processing; hyperspectral image classification; airborne LiDAR; land cover classification; multimodality data

收稿日期: 2023-01-30; 修回日期: 2023-03-09; 录用日期: 2023-03-13; 网络首发日期: 2023-03-26

基金项目: 中科院创新基金(E01Z040101)

通信作者: *c3i11@sia.cn

1 引言

利用遥感影像数据进行地物分类在城市规划^[1]、矿物勘探^[2]、环境监测^[3]等领域有广泛应用,因此地物分类成为遥感数据处理的一个重要课题。随着遥感技术的发展,获取同一场景的多源遥感数据成为可能^[4]。高光谱图像(HSI)数据具有光谱分辨率高的特点,其通道数可能高达几十甚至上百,因此 HSI 数据能够比较准确地反映不同地物的光谱特征,然而 HSI 数据存在同物异谱、同谱异物等问题,导致分类精度下降^[5]。与 HSI 数据不同,机载激光雷达(LiDAR)数据能够获取数字地表模型(DSM)信息,得到地物的立体高度,因此 LiDAR 数据可以作为高光谱数据的补充^[6]。融合 HSI 数据和 LiDAR 数据以提高地物分类精度成为许多学者的研究课题。

许多学者对联合 HSI 数据和 LiDAR 数据的地物分类进行了探索。Liao 等^[7]提取 LiDAR 数据和 HSI 数据的形态学轮廓特征,利用基于图的子空间嵌入方法对提取的特征进行融合。Ghamisi 等^[8]使用消光剖面对 HSI 和 LiDAR 数据进行特征提取,通过构造复合核函数对不同的数据特征进行融合,利用极限学习机进行分类。曹琼等^[9]为了能同时利用特征级融合和决策级融合的优势,提出了一种进行多级融合的方法,对从 HSI 和 LiDAR 数据提取的特征进行特征级融合,对多个分类结果利用投票方式进行决策级融合。王雷光等^[10]为了解决单一融合策略造成的分类精度低的问题,提出一种光谱空间特征融合的条件随机场分类方法。

自从 AlexNet^[11]被提出,许多学者认识到卷积神经网络(CNN)具有较强的图像特征提取能力,并将其应用于联合 HSI 数据和 LiDAR 数据的地物分类领域。为了能够分别提取多源数据的特征并对其进行融合,Xu 等^[12]提出了一种双分支的卷积神经网络,该网络能够分别提取并融合 HSI 数据和 LiDAR 数据的特征。Chen 等^[13]使用卷积神经网络分别对 HSI 和 LiDAR 数据进行特征提取,利用全连接层进行特征融合及分类,相比传统方法,提高了分类精度。为了解决卷积神经网络提取的不同数据源的特征不能够得到充分融合的问题,Mohla 等^[14]提出了交叉注意力机制,对 HSI 和 LiDAR 数据经卷积神经网络提取的空间特征进行融合,通过构建自注意力机制对 HSI 的光谱信息进行特征提取融合。为了更高效地对不同数据源数据进行融合,Hang 等^[15]在特征层面和决策层面对两个卷积神经网络分支进行特征融合,同时令两个网络分支共享部分网络参数。为了提高多源数据的融合效率,Wu 等^[16]提出了一种即插即用的跨通道特征融合模块,对提取出的不同数据源的特征进行拼接,经过特征融合模块融合重建后对重建结果与原特征进行比较,使模型学习到更高效和紧凑

地实现 HSI 数据和 LiDAR 数据的特征融合,Hong 等^[17]提出了一种利用编码器-解码器结构的特征融合方法。为了提取不同传感器之间的多模态特征,Wang 等^[18]提出了一种三支的网络结构和模态注意力模块,并采用分层密集融合策略对多模态数据进行融合。为了更加关注 HSI 数据的通道特征以及增强 HSI 和 LiDAR 特征的互补性,Wang 等^[19]设计了光谱空间特征互通模块,增加了高层语义信息的关联性。为了克服形态学属性剖面的局限性并利用其优点,Roy 等^[20]提出了一种可训练的空间形态学卷积模块,提高了地物分类精度。为了解决卷积操作会弱化中心像素权重以及 HSI 与 LiDAR 数据特征不能很好融合的问题,Dong 等^[21]设计了自引导和交叉引导模块,以更好地融合 HSI 与 LiDAR 数据信息,并提出了光谱补充模块以增加特定像素的区分度。

为了更有效地提取图像全局语义特征,Dosovitskiy 等^[22]提出了基于 Transformer^[23]的 ViT 模型,该模型的精度在当时的图像分类任务中超过了主流的卷积神经网络。为了提取 HSI 数据更精确的高层语义信息,金传等^[24]提出了一种将 CNN 与 Transformer 相结合的网络结构,提高了遥感图像分类精度。为了捕获远距离高光谱序列之间的关系,陈禹汗等^[25]设计了基于 Swin Transformer 的多尺度混合光谱注意力模型。为了充分挖掘和利用 HSI 与 LiDAR 数据特征以及更好地融合多源数据特征,Zhao 等^[26]提出了将卷积网络与 Transformer 相结合的双分支网络结构和交叉注意力融合编码器。

然而,深度学习方法往往需要较多的训练样本,获取大量标签样本需要大量的成本和时间。为了解决这个问题,本文提出一种利用 HSI 和 LiDAR 数据进行对比学习的方法。首先利用不带标签的样本进行对比学习,使模型建立这两种数据特征之间的联系,实现对数据的特征提取,然后再利用带标签的样本对模型进行微调,使模型获得地物分类能力;设计了卷积模块与 Transformer 模块相结合的网络,该网络能够更好地利用局部特征提取全局的语义信息。在 Houston 2013 数据集和 Trento 数据集对不同方法进行测试,实验结果表明,所提方法的总体分类精度高于其他对比方法。

2 基于对比学习的高光谱和 LiDAR 分类方法

2.1 方法流程

首先介绍所提方法的具体流程,采用对比学习^[27]的方法,利用无标签样本对模型进行无监督预训练,然后用带标签样本对预训练模型进行微调。图 1(a)表示无监督的对比学习预训练阶段,图 1(b)表示对预训练得到的模型用带标签样本进行微调的过程。

在无监督的对比学习预训练阶段,首先对加载的 HSI 数据进行主成分分析(PCA)降维,降维后的 HSI 数

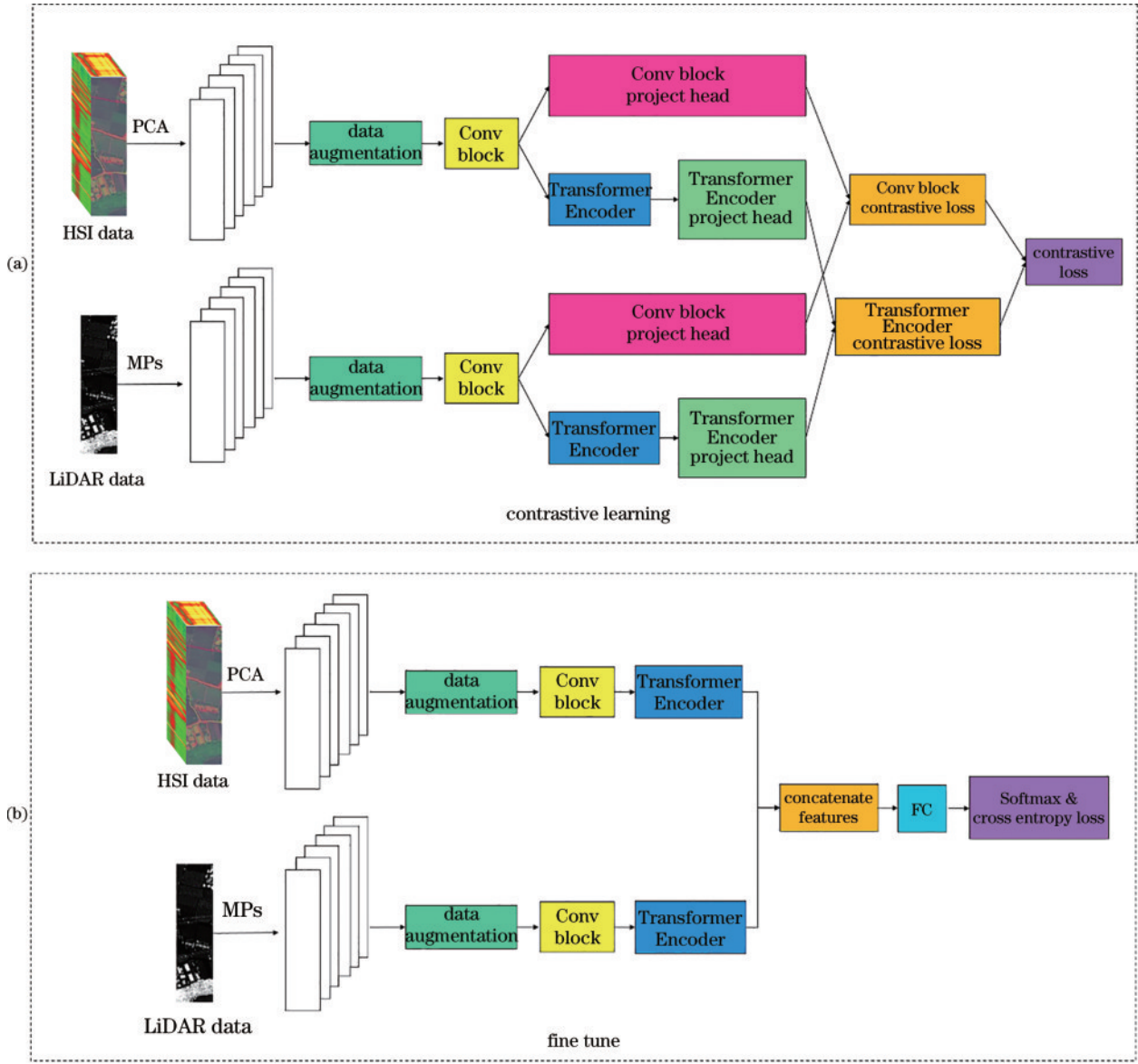


图 1 所提方法的框架。(a)对比学习;(b)微调

Fig. 1 Framework of the proposed method. (a) Contrastive learning; (b) fine tune

据光谱维度为 d_{HSI} 。为了更好地提取图像特征,对反映地面高度的 LiDAR 数据进行形态学剖面(MPs)^[28]处理, LiDAR 数据图像宽和高分别为 w_{LiDAR} 和 h_{LiDAR} , 处理后的 LiDAR 数据为 $\mathbf{x}_{\text{LiDAR}} \in \mathbf{R}^{w_{\text{LiDAR}} \times h_{\text{LiDAR}} \times d_{\text{LiDAR}}}$ 。经 PCA 处理后的 HSI 光谱维度为 d_{HSI} , 从 HSI 上随机选取 n 个样本点中心, 即正样本对数量为 n , 在样本点中心周围截取一定宽度 s 的正方形像素块 $\mathbf{x}_{\text{HSI}}^{(i)} \in \mathbf{R}^{s \times s \times d_{\text{HSI}}}$, $i = 1, 2, \dots, n$; 在 LiDAR 数据上选取对应 HSI 位置的 n 个样本 $\mathbf{x}_{\text{LiDAR}}^{(i)} \in \mathbf{R}^{s \times s \times d_{\text{LiDAR}}}$, 作为无监督对比学习的样本。之后, 为了提高模型的泛化能力, 对获取的 HSI 和 LiDAR 样本数据集进行数据增强, 采用随机旋转、灰度变换、模糊、添加噪声的方法进行数据增强 (data augmentation)。数据增强后的 HSI 样本和 LiDAR 样本分别为 $\tilde{\mathbf{x}}_{\text{HSI}}^{(i)} \in \mathbf{R}^{s \times s \times d_{\text{HSI}}}$ 和 $\tilde{\mathbf{x}}_{\text{LiDAR}}^{(i)} \in \mathbf{R}^{s \times s \times d_{\text{LiDAR}}}$ 。其中, 对应位置的样本对 $(\tilde{\mathbf{x}}_{\text{HSI}}^{(i)}, \tilde{\mathbf{x}}_{\text{LiDAR}}^{(i)})$ 作为正样本对。

在本文的对比学习中, 分别处理 HSI 和 LiDAR 数据的两个网络分支并不共享参数, 即处理 HSI 数据的网络分支与处理 LiDAR 数据的网络分支各自使用各自的参数。将数据增强后的样本 $\tilde{\mathbf{x}}_{\text{HSI}}^{(i)}$ 与 $\tilde{\mathbf{x}}_{\text{LiDAR}}^{(i)}$ 分别送入各自的卷积网络模块 (Conv block) 中。对于处理 HSI 的网络分支或处理 LiDAR 的网络分支, 卷积网络模块的输出结果分为两个部分, 一部分进入卷积网络输出的映射头 (Conv block project head), 另一部分进入 Transformer Encoder 后再进入映射头 (Transformer Encoder project head)。从两路映射头输出的结果分别进入两个对比损失函数, 即从 Conv block project head 输出的结果进入 Conv block contrastive loss, 从 Transformer Encoder project head 输出的结果进入 Transformer Encoder contrastive loss。为了提高对比学习训练时的稳定性, 这里将得到的两个对比损失函

数值按一定权重相加,结果作为最终的对比损失函数值(contrastive loss)。利用梯度下降法对计算得到的对比损失函数值进行反向传播,更新网络参数。

当无监督的对比学习预训练阶段完成后,进入监督学习的微调阶段,如图 1(b)所示。首先对输入的 HSI 数据与 LiDAR 数据进行 PCA、MPs 及数据增强处理,与无监督预训练阶段的处理一致,不同之处在于选取的样本带有标签,且样本数量少于无监督预训练阶段的样本数量,样本表示为 $(\tilde{\mathbf{x}}_{\text{HSI}}^{(j)}, \tilde{\mathbf{x}}_{\text{LiDAR}}^{(j)}, \mathbf{y}^{(j)})$, $j = 1, 2, \dots, m$, 其中 $\tilde{\mathbf{x}}_{\text{HSI}}^{(j)}$ 和 $\tilde{\mathbf{x}}_{\text{LiDAR}}^{(j)}$ 表示用于训练的 HSI 和 LiDAR 样本对, $\mathbf{y}^{(j)}$ 为样本的标签, m 为样本数。在监督学习的微调阶段,将卷积网络模块(Conv block)与 Transformer Encoder 模块的参数固定,在模型后端添加全连接层(FC)。将数据增强之后的样本 $(\tilde{\mathbf{x}}_{\text{HSI}}^{(j)}, \tilde{\mathbf{x}}_{\text{LiDAR}}^{(j)})$ 依次输入 Conv block 与 Transformer Encoder 模块,从 Transformer Encoder 模块输出的特征向量分别为 $\mathbf{v}_{\text{HSI}}^{(j)}$ 与 $\mathbf{v}_{\text{LiDAR}}^{(j)}$ 。对两个特征向量进行拼接(concatenate feature),得到 $\mathbf{v}^{(j)} = [\mathbf{v}_{\text{HSI}}^{(j)}, \mathbf{v}_{\text{LiDAR}}^{(j)}]$ 。将拼接后的向量 $\mathbf{v}^{(j)}$ 输入 FC 中,输出结果为 $(\mathbf{q}^{(1)}, \mathbf{q}^{(2)}, \dots, \mathbf{q}^{(m)})$, 其中 $\mathbf{q}^{(j)} \in \mathbf{R}^{c \times 1}$, c 为样本的类别数。经过 Softmax 和交叉熵损失(cross entropy loss)函数处理,有

$$L_{\text{cross}} = -\frac{1}{m} \sum_{j=1}^m \sum_{r=1}^c I_{j,r} \log_e \frac{e^{q_r^{(j)}}}{\sum_{k=1}^c e^{q_k^{(j)}}}, \quad (1)$$

$$I_{j,r} = \begin{cases} 1, & \mathbf{y}^{(j)} = r \\ 0, & \mathbf{y}^{(j)} \neq r \end{cases}, \quad (2)$$

式中: $q_r^{(j)}$ 为向量 $\mathbf{q}^{(j)}$ 中的第 r 个值; $I_{j,r}$ 的值为当标签值

与预测值相同时为 1, 不同则为 0。最后利用梯度下降法对计算得到的损失函数 L_{cross} 更新网络参数。对带标签样本进行监督学习训练后,对模型在测试集上进行测试,用相关指标对测试分类结果进行评估。

2.2 对比学习

在一些图像识别任务中,可能会出现未标记样本数量多于标记样本数量的情况,因此,文献[27]提出了一种利用代理任务对未标记样本进行训练从而提高模型泛化能力的对比学习框架。图 1(a)表示无监督的对比学习预训练阶段,本文对文献[27]中的框架进行了一定的改进,以适应多源数据情况下的地物分类任务。

如前面所述,数据增强后的 HSI 样本和 LiDAR 样本分别为 $\tilde{\mathbf{x}}_{\text{HSI}}^{(i)}$ 和 $\tilde{\mathbf{x}}_{\text{LiDAR}}^{(i)}$ 。样本 $\tilde{\mathbf{x}}_{\text{HSI}}^{(i)}$ 依次通过卷积网络模块、卷积网络输出的映射头后得到向量 $\mathbf{z}_{\text{HSI-C}}^{(i)}$; 样本 $\tilde{\mathbf{x}}_{\text{HSI}}^{(i)}$ 依次通过卷积网络模块、Transformer Encoder 模块、Transformer Encoder project head 后得到向量 $\mathbf{z}_{\text{HSI-T}}^{(i)}$; 同理,样本 $\tilde{\mathbf{x}}_{\text{LiDAR}}^{(i)}$ 依次通过卷积网络模块、卷积网络输出的映射头后得到向量 $\mathbf{z}_{\text{LiDAR-C}}^{(i)}$; 样本 $\tilde{\mathbf{x}}_{\text{LiDAR}}^{(i)}$ 依次通过卷积网络模块、Transformer Encoder 模块、Transformer Encoder project head 后得到向量 $\mathbf{z}_{\text{LiDAR-T}}^{(i)}$ 。其中,对应位置的样本对 $(\tilde{\mathbf{x}}_{\text{HSI}}^{(i)}, \tilde{\mathbf{x}}_{\text{LiDAR}}^{(i)})$ 作为正样本对,其他位置的样本对作为负样本对,对比损失函数计算正样本对与负样本对相似度的比值,当网络的参数更新后,降低负样本对之间的相似度,增大正样本对之间的相似度。向量 \mathbf{z}_1 和 \mathbf{z}_2 相似度计算公式为

$$f(\mathbf{z}_1, \mathbf{z}_2) = \frac{\mathbf{z}_1 \cdot \mathbf{z}_2}{\|\mathbf{z}_1\| \|\mathbf{z}_2\|}. \quad (3)$$

经过卷积模块的对比损失函数(Conv block contrastive loss)公式为

$$L_{\text{contrast-c}}(\mathbf{z}_{\text{HSI-C}}^{(i)}, \mathbf{z}_{\text{LiDAR-C}}^{(i)}) = -\log_e \frac{2 \cdot \exp\left[\frac{f(\mathbf{z}_{\text{HSI-C}}^{(i)}, \mathbf{z}_{\text{LiDAR-C}}^{(i)})}{\tau}\right]}{\sum_{i \neq t} \exp\left[\frac{f(\mathbf{z}_{\text{HSI-C}}^{(i)}, \mathbf{z}_{\text{HSI-C}}^{(t)})}{\tau}\right] + 2 \sum_{i \neq t} \exp\left[\frac{f(\mathbf{z}_{\text{HSI-C}}^{(i)}, \mathbf{z}_{\text{LiDAR-C}}^{(t)})}{\tau}\right] + \sum_{i \neq t} \exp\left[\frac{f(\mathbf{z}_{\text{LiDAR-C}}^{(i)}, \mathbf{z}_{\text{LiDAR-C}}^{(t)})}{\tau}\right]}, \quad (4)$$

式中: τ 表示一超参数。经 Transformer Encoder project head 的对比损失函数(Transformer Encoder contrastive loss)公式为

$$L_{\text{contrast-t}}(\mathbf{z}_{\text{HSI-T}}^{(i)}, \mathbf{z}_{\text{LiDAR-T}}^{(i)}) = -\log_e \frac{2 \cdot \exp\left[\frac{f(\mathbf{z}_{\text{HSI-T}}^{(i)}, \mathbf{z}_{\text{LiDAR-T}}^{(i)})}{\tau}\right]}{\sum_{i \neq t} \exp\left[\frac{f(\mathbf{z}_{\text{HSI-T}}^{(i)}, \mathbf{z}_{\text{HSI-T}}^{(t)})}{\tau}\right] + 2 \sum_{i \neq t} \exp\left[\frac{f(\mathbf{z}_{\text{HSI-T}}^{(i)}, \mathbf{z}_{\text{LiDAR-T}}^{(t)})}{\tau}\right] + \sum_{i \neq t} \exp\left[\frac{f(\mathbf{z}_{\text{LiDAR-T}}^{(i)}, \mathbf{z}_{\text{LiDAR-T}}^{(t)})}{\tau}\right]}. \quad (5)$$

将两路对比损失函数值相加,得到总对比损失函数(contrastive loss),公式为

$$L_{\text{contrast}} = \sum_{i=1}^n \left[\rho \cdot L_{\text{contrast-c}}(\mathbf{z}_{\text{HSI-C}}^{(i)}, \mathbf{z}_{\text{LiDAR-C}}^{(i)}) + (1 - \rho) L_{\text{contrast-t}}(\mathbf{z}_{\text{HSI-T}}^{(i)}, \mathbf{z}_{\text{LiDAR-T}}^{(i)}) \right], \quad (6)$$

式中: $\rho \in [0, 1]$ 为调整两路对比损失函数权重的超参数。计算对比损失函数值之后利用梯度下降方法更新

网络参数,训练完成之后实现模型对 HSI 与 LiDAR 数据的特征提取。

2.3 网络结构

2.3.1 卷积模块

介绍卷积网络模块(Conv block),如图 2 和图 3 所示,input patch 为输入的样本,densenet block 为密集连接的卷积网络模块^[29],Concat 代表特征图的叠加,

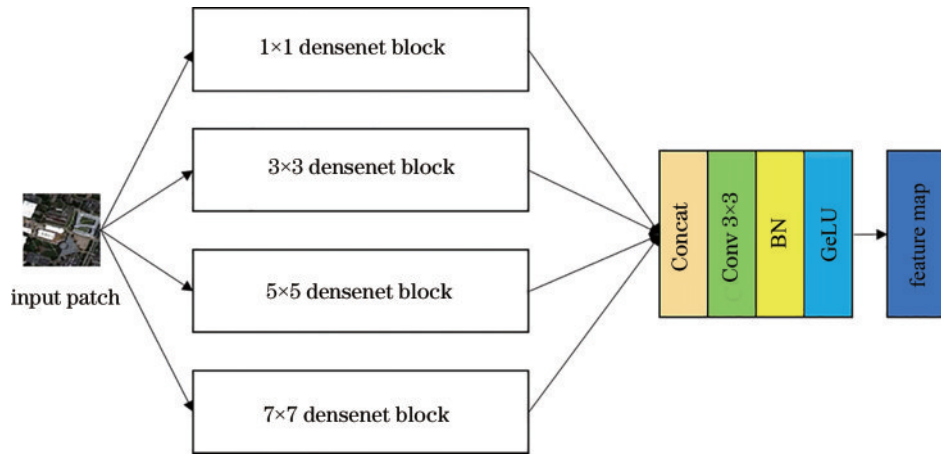


图 2 卷积网络模块

Fig. 2 Convolution network module

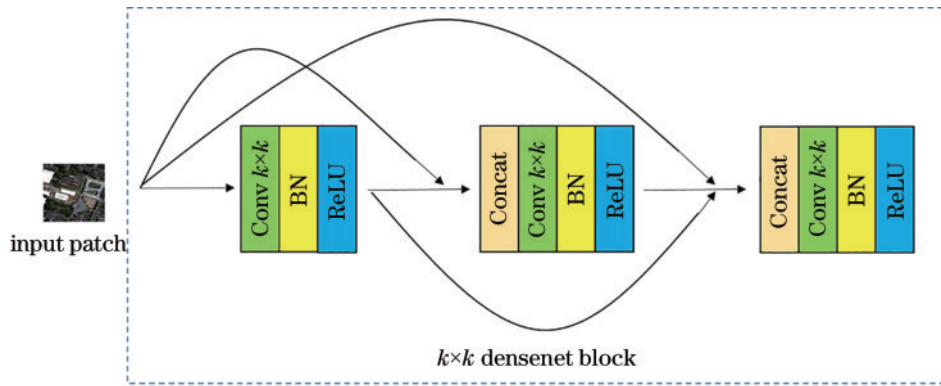


图 3 密集连接卷积网络模块

Fig. 3 Densenet block

Conv 代表卷积操作, BN 代表批归一化操作, ReLU 代表 ReLU 激活函数, GeLU 代表 GeLU 激活函数, feature map 代表特征图。将数据增强之后的样本 input patch 分别输入到卷积核大小为 (1, 1)、(3, 3)、(5, 5)、(7, 7) 的密集连接卷积网络模块中, 不同大小的卷积核能够提取出不同大小感受野的特征, 使得这些特征起到相互补充的作用。然后对提取出的各个特征图进行拼接, 再进行卷积操作, 最后输出特征图 feature map。图 3 为密集连接的卷积网络, 密集连接的卷积网络通过建立前面层与后面所有层的连接来实现特征重用, 提高特征利用率, 增强了特征传播。

2.3.2 Transformer Encoder 模块

卷积网络虽然善于提取图像的局部特征, 但却不善于获取全局特征, 由于 Transformer Encoder^[22] 能够学习到长距离的依赖关系, 因此将卷积网络与 Transformer Encoder 相结合能够更好地提取全局语义信息。图 4 为 Transformer Encoder 模块, input 代表输入, flatten 代表将二维向量重新排列为 一维向量的操作, embedding 代表编码操作, LN 代表层归一化操作,

multi-head attention 代表多头注意力机制层, 加号代表两向量对应元素相加, FC 代表全连接层, output 代表输出。

在 embedding 编码操作中, 输入为 X_c , 输出为 X_e , 计算公式为

$$X_e = [X_c \cdot W_a]^T \cdot [X_c \cdot W_b] + p, \quad (7)$$

式中: W_a 、 W_b 和 p 为可学习的参数。

在自注意力机制中, W_q 、 W_k 、 W_v 分别代表可学习的参数, 输入为 X_L , 输出为 B , 查询矩阵 $Q = X_L \cdot W_q$, 键矩阵 $K = X_L \cdot W_k$, 值矩阵 $V = X_L \cdot W_v$, 公式为

$$B = \text{Soft max} \left(\frac{QK^T}{\sqrt{d}} \right) V, \quad (8)$$

式中: d 为矩阵 Q 的行向量的维度。多头注意力机制相当于将自注意力机制分开运算合并后与另一矩阵相乘的过程, 如图 5 所示, mul 代表矩阵相乘, 分别将矩阵 Q 、 K 、 V 分成 h 份, 对每一份按式 (8) 进行运算, 对各个运算结果 $B_i (i = 1, 2, \dots, h)$ 进行合并, 再与可学习的参数 W_o 进行矩阵相乘, 输出 X_s 。

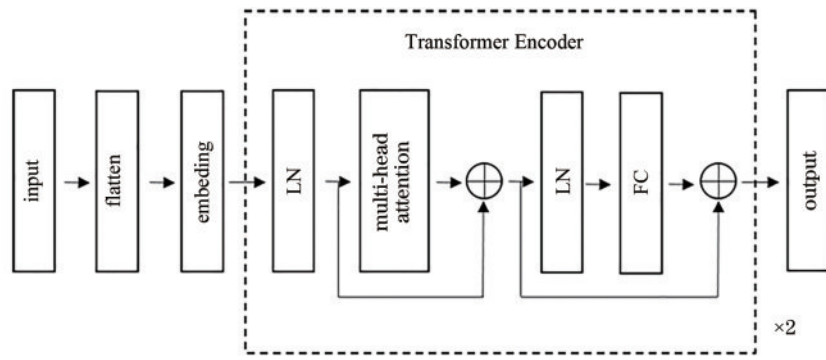


图 4 Transformer Encoder 模块
Fig. 4 Transformer Encoder block

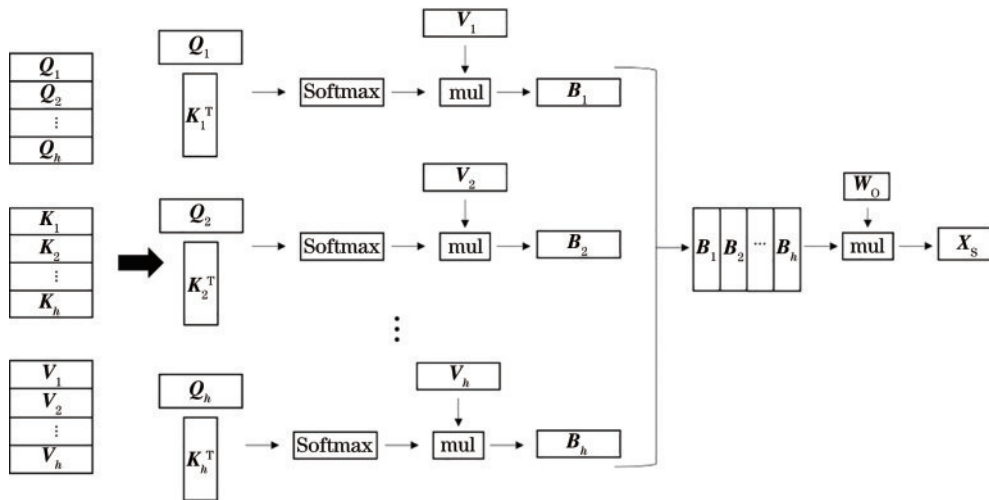


图 5 多头注意力机制
Fig. 5 Multi-head attention mechanism

3 实验分析与讨论

3.1 实验数据与设置

实验采用两组公开的实验数据集。第一组数据集是 Houston 2013 数据集^[30], 如图 6 所示, 由美国休斯顿

大学机载激光测绘中心提供, 用于 2013 IEEE GRSS 数据融合比赛, 数据由机载光谱成像仪拍摄得到, 包含高光谱图像(HSI)和机载激光雷达(LiDAR)数据, 所有数据图像大小均为 349×1905 像素, 高光谱数据包含 144 通道, 波长范围为 0.38~1.05 μm。地物类别共有

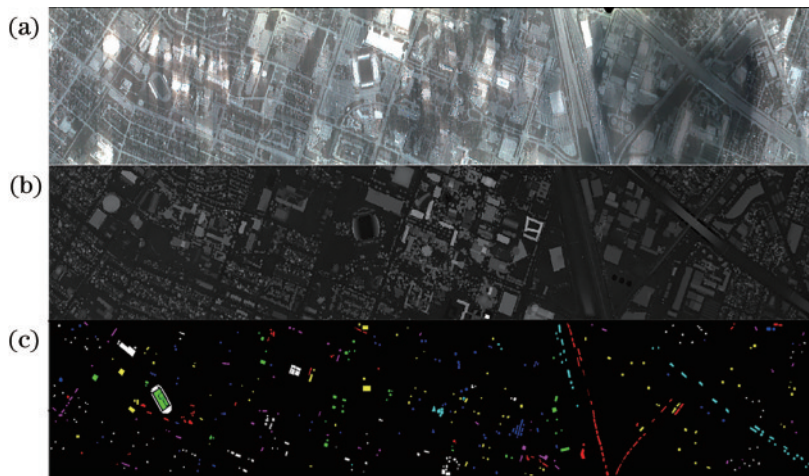


图 6 Houston 2013 数据集。(a) HSI 数据; (b) LiDAR 数据; (c) 真实标签
Fig. 6 Houston 2013 dataset. (a) HSI data; (b) LiDAR data; (c) ground-truth label

15类,包括树木、土壤、水域、道路、商业区等。在无监督预训练阶段,不带标签的样本数为12288,在监督训练的微调阶段,每个类别带标签的样本数分别为5、10、15、20、25、50、100,测试样本量为15029。第二组数据集是 Trento 数据集^[31],如图7所示,此数据集拍摄于意大利城市特伦托南部的农村地区,包括高光谱图像和机载激光雷达数据,其中高光谱图像数据包含63

通道,波长范围为0.42~0.99 μm ,机载激光雷达数据能够反映地物的高度,两种数据的图像大小为600×166像素,地物一共包含建筑物、树林、道路等6个类别。在无监督预训练阶段,不带标签的样本数为4096,在监督训练的微调阶段,每个类别带标签的样本数分别为2、3、4、5、6、9、12,测试样本量为30214。

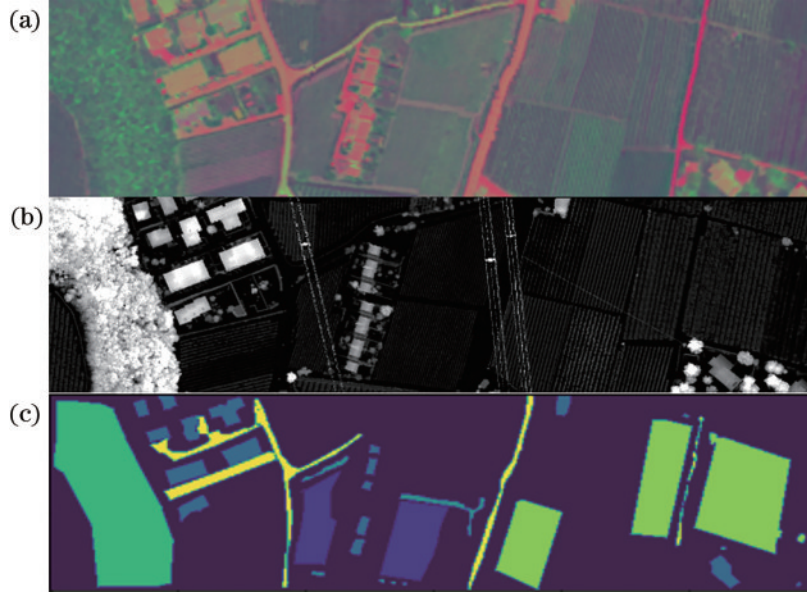


图7 Trento数据集。(a) HSI数据;(b) LiDAR数据;(c)真实标签

Fig. 7 Trento dataset. (a) HSI data; (b) LiDAR data; (c) ground-truth label

为了验证所提方法在少标签地物分类上的有效性,对所提方法与多源数据融合的地物分类方法 EndNet^[17]、CCRNet^[16]、two brance CNN^[12]、MAHiDFNet^[18]进行对比。实验硬件环境为 Intel(R) Core™ i7-3470 处理器和 NVIDIA GeForce RTX3090 显卡。所提方法的实验在 PyTorch 深度学习框架下搭建。相关参数设置为:学习率设为0.0003;样本尺寸设为11×11;在 Houston 2013 数据集训练过程中对比学习训练总轮数(epoch)设为650,批处理大小为1024,监督训练总轮数设为100;在 Trento 数据集训练过程中对比学习训练总轮数设为300,批处理大小为1024,监督训练总轮数设为100;HSI数据经过PCA降维后维度为7;优化方法采用 Adam 优化算法;超参数 τ 设为0.035,超参

数 ρ 设为0.65。

实验评价指标采用总体分类精度(OA),公式为

$$S_{OA} = \frac{\sum_{i=1}^c a_{ii}}{M}, \quad (9)$$

式中: S_{OA} 为总体分类精度; M 为测试的样本数量; a_{ii} 为标签类别为*i*,预测类别为*i*的样本数。

3.2 实验结果与分析

3.2.1 结果与分析

在 Houston 2013 数据集上各个方法的实验结果如表1所示,在 Trento 数据集上各个方法的实验结果如表2所示。 u 表示各个类别带标签样本的数量,ours为所提方法,最优结果通过加粗表示。

表1 在 Houston 2013 数据集上的总体分类精度

Table 1 Overall classification accuracy on the Houston 2013 dataset

unit: %

Algorithm	$u = 5$	$u = 10$	$u = 15$	$u = 20$	$u = 25$	$u = 50$	$u = 100$
EndNet	54.38	68.94	74.12	77.33	80.66	85.54	91.36
CCRNet	47.97	65.87	77.95	82.28	84.38	93.24	94.25
two brance CNN	57.74	75.25	77.60	81.60	84.58	89.89	93.80
MAHiDFNet	55.12	75.00	82.93	86.08	85.67	89.95	91.21
ours	78.47	88.87	90.60	91.36	93.75	95.26	96.77

表 2 在 Trento 数据集上的总体分类精度

Table 2 Overall classification accuracy on the Trento dataset

unit: %

Algorithm	$u=2$	$u=3$	$u=4$	$u=5$	$u=6$	$u=9$	$u=12$
EndNet	58.70	73.02	81.11	76.63	79.68	85.29	88.20
CCRNet	67.80	72.00	70.50	78.90	83.64	86.24	93.94
two brance CNN	78.06	89.21	90.73	94.63	96.34	96.72	95.48
MAHiDFNet	79.77	84.96	92.80	95.87	98.13	99.09	98.98
ours	88.12	97.1	97.19	97.54	98.30	99.09	99.23

在 Houston 2013 数据集中当 u 分别等于 5, 10, 15 时, 在 Trento 数据集中当 u 分别等于 2, 3, 4 时, 所提方法的总体分类精度要明显高于其他方法的总体分类精度; 随着样本数的增加, 所提方法的总体分类精度先快速提升到较高值, 之后缓慢提升, 说明随着带标签样本数的增多, 单个带标签样本对总体分类精度的作用下降; 在 Houston 2013 数据集中当 u 分别等于 25, 50, 100 时, 在 Trento 数据集中当 u 分别等于 6, 9, 12 时, 所提方法的总体分类精度与其他方法相当。因此, 所提方法可以在一定程度上解决带标签样本的数量减少时带来的性能下降问题。分类结果曲线如图 8 和图 9 所示。在 Houston 2013 数据集上当 u 分别等于 5, 10, 15 时, 分类结果分别如图 10~12 所示, 在 Trento 数据集上当 u 分别等于 2, 3, 4 时, 分类结果分别如图 13~15 所示, 可以看出 EndNet 和 CCRNet 处理的图像存在较为明显的噪声现象, 而 two brance CNN 和 MAHiDFNet 处理的图像中虽然噪声相对较少, 但容易出现相似地物错分的情况, 所提方法处理的图像中噪声现象较少, 分类情况与真实标签比较一致。

表 3 为各个方法的每轮训练的平均用时, 其中 contrast learning 表示所提方法对比学习阶段的平均用

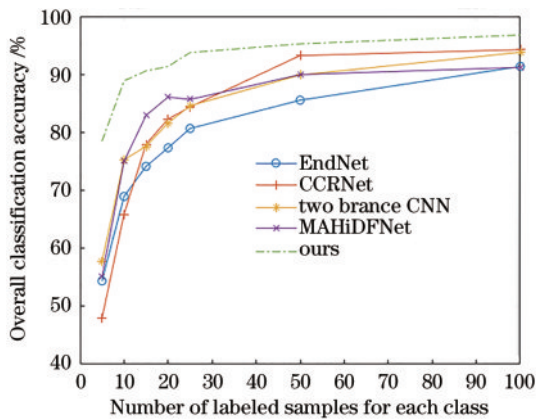


图 8 Houston 2013 数据集上的分类结果

Fig. 8 Classification results on the Houston 2013 dataset

表 3 每轮训练的平均用时

Table 3 Average time of each epoch of training

unit: s

Dataset	EndNet	CCRNet	two brance CNN	MAHiDFNet	contrast learning	fine tune	ours
Houston 2013	0.13	0.40	36.45	6.80	3.82	0.44	3.37
Trento	0.20	0.20	26.28	2.06	0.94	0.39	0.72

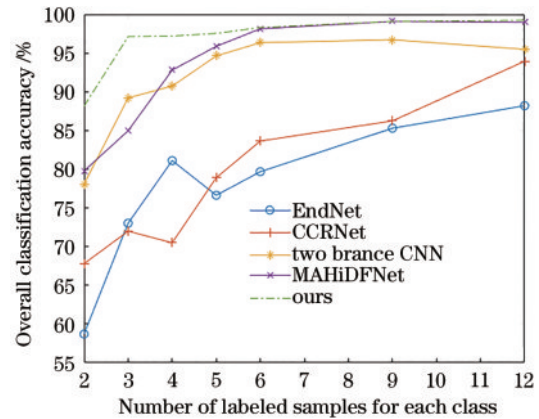


图 9 Trento 数据集上的分类结果

Fig. 9 Classification results on the Trento dataset

时, fine tune 表示所提方法微调阶段的平均用时, ours 表示所提方法对比学习阶段加微调阶段的平均用时。从表 3 可以看出所提方法训练用时主要在对学习阶段。

3.2.2 消融实验

为了验证利用 HSI 数据和 LiDAR 数据进行对比学习的有效性, 进行消融实验。实验结果如表 4 和表 5 所示, 其中 without contrast learning 表示不进行对比学习预训练, 单独利用带标签样本对模型进行训练, with contrast learning 表示训练过程包含对比学习预训练及利用带标签样本进行微调。从表 4 和表 5 可以看出: 在 Houston 2013 数据集中, 当 u 分别等于 5, 10, 15, 20, 25, 50, 100 时, 训练过程包含对比学习预训练的总体分类精度要比没有对比学习预训练的总体分类精度分别高 14.85 个百分点, 16.18 个百分点, 9.02 个百分点, 6.22 个百分点, 8.04 个百分点, 2.13 个百分点, 3.36 个百分点; 在 Trento 数据集中, 当 u 分别等于 2, 3, 4, 5, 6, 9, 12 时, 训练过程包含对比学习预训练的总体分类精度要比没有对比学习预训练的总体分类精度分别高 4.59 个百分点, 3.76 个百分点, 4.38 个百分点, 2.00 个百分点, 2.92 个百分点, 0.87 个百分点, 0.78 个百分点。即利用 HSI 数

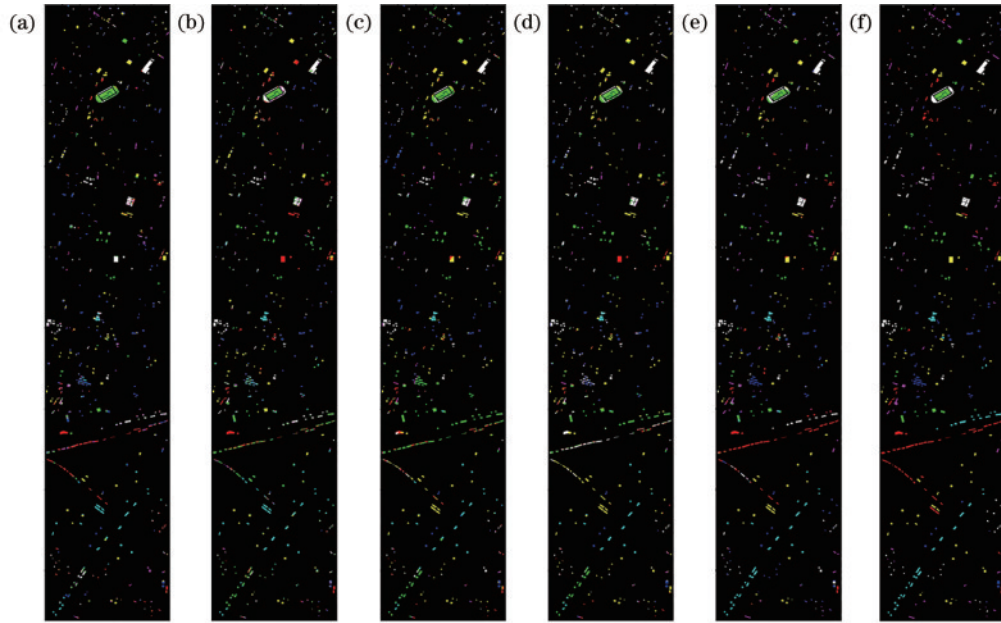


图 10 当 $u = 5$ 时, Houston 2013 数据集上的分类结果。(a) EndNet; (b) CCRNet; (c) two branch CNN; (d) MAHiDFNet; (e) 所提方法; (f) 真实标签

Fig. 10 Classification results on the Houston 2013 dataset when $u = 5$. (a) EndNet; (b) CCRNet; (c) two branch CNN; (d) MAHiDFNet; (e) ours; (f) ground-truth label

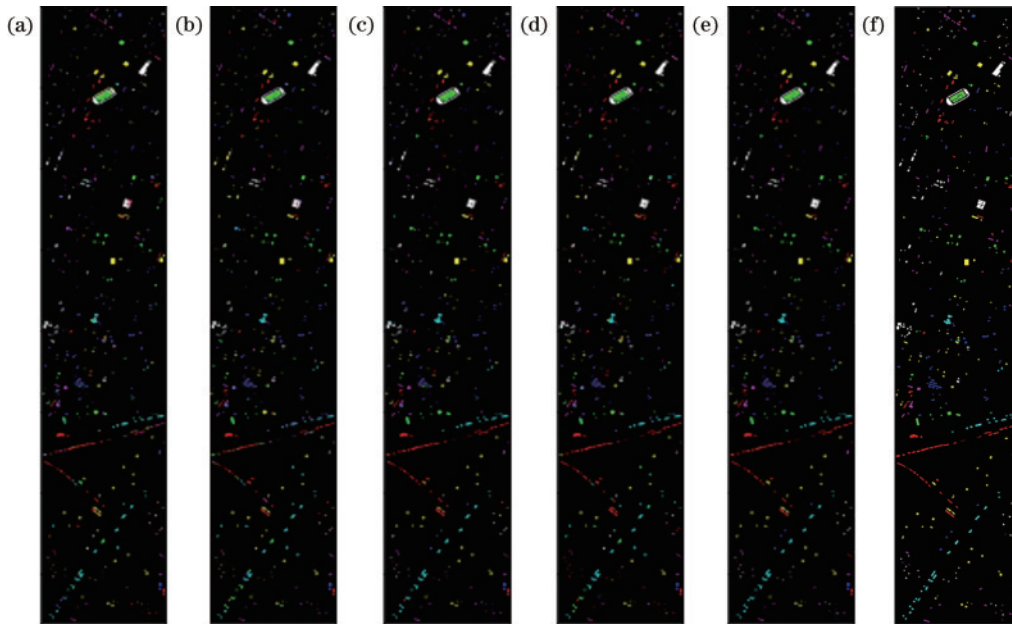


图 11 当 $u = 10$ 时, Houston 2013 数据集上的分类结果。(a) EndNet; (b) CCRNet; (c) two branch CNN; (d) MAHiDFNet; (e) 所提方法; (f) 真实标签

Fig. 11 Classification results on the Houston 2013 dataset when $u = 10$. (a) EndNet; (b) CCRNet; (c) two branch CNN; (d) MAHiDFNet; (e) ours; (f) ground-truth label

据和 LiDAR 数据进行对比学习预训练得到的地物分类精度要高于没有进行预训练的地物分类精度。结果表

明, 利用 HSI 数据和 LiDAR 数据进行对比学习预训练能够实现模型对这两种数据的特征提取。

表 4 在 Houston 2013 数据集上进行对比学习消融实验的总体分类精度

Table 4 Overall classification accuracy of contrastive learning ablation experiments on the Houston 2013 dataset unit: %

Condition	$u = 5$	$u = 10$	$u = 15$	$u = 20$	$u = 25$	$u = 50$	$u = 100$
without contrast learning	63.62	72.69	81.58	85.14	85.71	93.13	93.41
with contrast learning	78.47	88.87	90.60	91.36	93.75	95.26	96.77

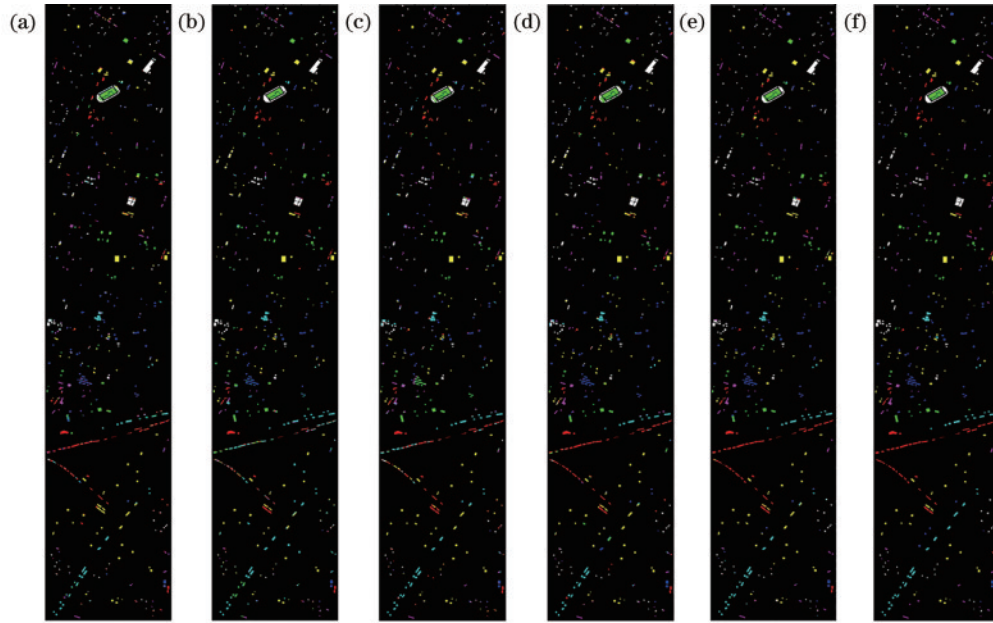


图 12 当 $u = 15$ 时, Houston 2013 数据集上的分类结果。(a) EndNet; (b) CCRNet; (c) two branch CNN; (d) MAHiDFNet; (e) 所提方法; (f) 真实标签

Fig. 12 Classification results on the Houston 2013 dataset when $u = 15$. (a) EndNet; (b) CCRNet; (c) two branch CNN; (d) MAHiDFNet; (e) ours; (f) ground-truth label

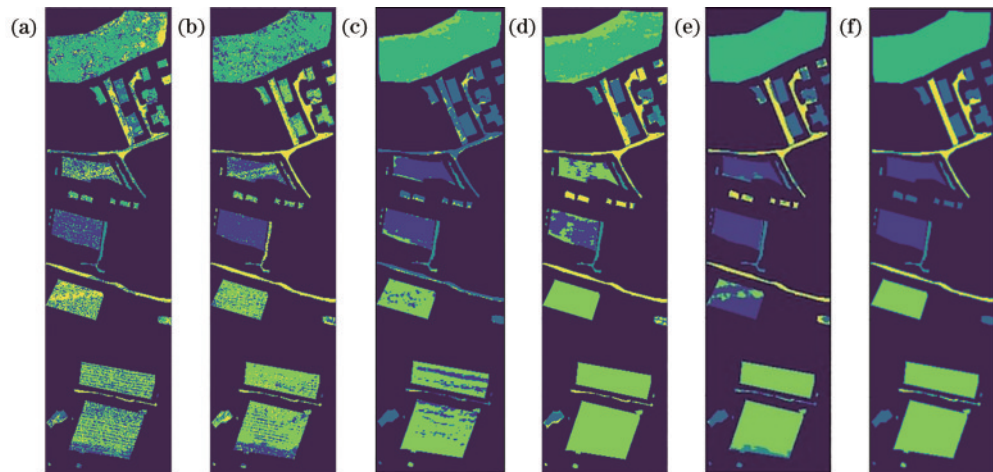


图 13 当 $u = 2$ 时, Trento 数据集上的分类结果。(a) EndNet; (b) CCRNet; (c) two branch CNN; (d) MAHiDFNet; (e) 所提方法; (f) 真实标签

Fig. 13 Classification results on the Trento dataset when $u = 2$. (a) EndNet; (b) CCRNet; (c) two branch CNN; (d) MAHiDFNet; (e) ours; (f) ground-truth label

表 5 在 Trento 数据集上进行对比学习消融实验的总体分类精度

Table 5 Overall classification accuracy of contrastive learning ablation experiments on the Trento dataset unit: %

Condition	$u = 2$	$u = 3$	$u = 4$	$u = 5$	$u = 6$	$u = 9$	$u = 12$
without contrast learning	83.53	93.36	92.81	95.54	95.38	98.22	98.45
with contrast learning	88.12	97.12	97.19	97.54	98.30	99.09	99.23

为了验证多种数据源能够提高模型的地物分类精度,进行不同数据源下的消融实验。实验结果如表 6 和表 7 所示,其中 HSI 表示单独使用 HSI 数据进行训练, LiDAR 表示单独使用 LiDAR 数据进行训练, HSI+LiDAR 表示同时使用 HSI 和 LiDAR 数据进行训练。从表 6 和表 7 可以看出:在 Houston 2013 数据集

中,当 u 分别等于 5, 10, 15, 20, 25, 50, 100 时,同时使用 HSI 和 LiDAR 数据进行训练要比单独使用 HSI 数据进行训练的总分类精度分别高 18.68 个百分点, 16.11 个百分点, 11.96 个百分点, 8.55 个百分点, 9.64 个百分点, 4.62 个百分点, 2.82 个百分点, 比单独使用 LiDAR 数据进行训练的总分类精度分别高 33.06 个

表 6 在 Houston 2013 数据集上不同数据源的总体分类精度

Table 6 Overall classification accuracy on the Houston 2013 dataset with different data modality unit: %

Data modality	$u = 5$	$u = 10$	$u = 15$	$u = 20$	$u = 25$	$u = 50$	$u = 100$
HSI	59.79	72.77	78.63	82.81	84.10	90.63	93.95
LiDAR	45.42	58.07	62.95	67.38	70.32	78.52	85.22
HSI+ LiDAR	78.47	88.87	90.60	91.36	93.75	95.26	96.77

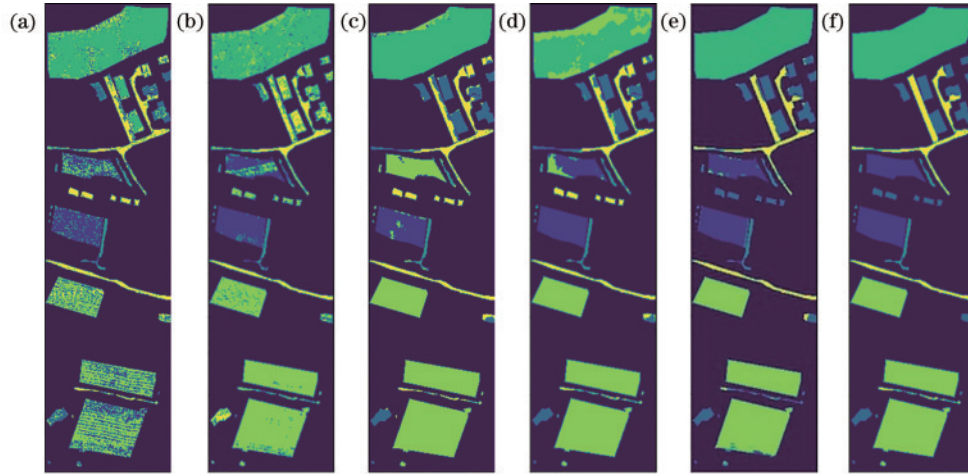


图 14 当 $u = 3$ 时, Trento 数据集上的分类结果。(a) EndNet; (b) CCRNet; (c) two branch CNN; (d) MAHiDFNet; (e) 所提方法; (f) 真实标签

Fig. 14 Classification results on the Trento dataset when $u = 3$. (a) EndNet; (b) CCRNet; (c) two branch CNN; (d) MAHiDFNet; (e) ours; (f) ground-truth label

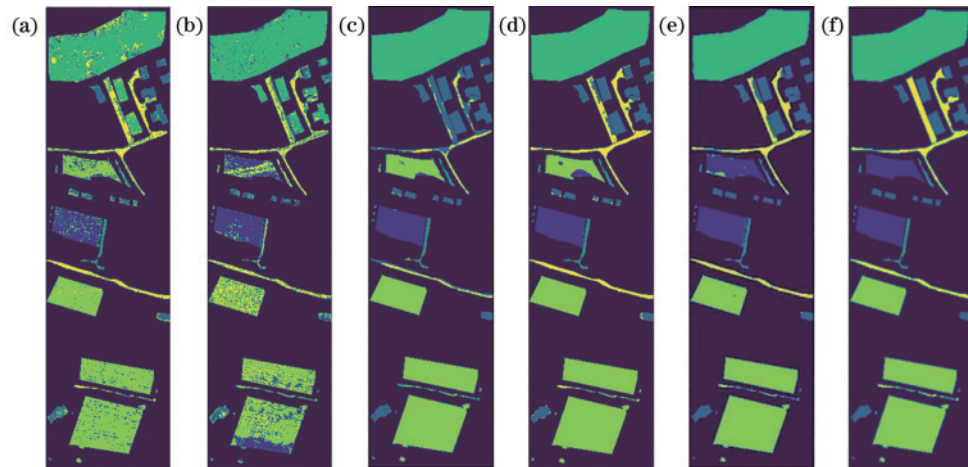


图 15 当 $u = 4$ 时, Trento 数据集上的分类结果。(a) EndNet; (b) CCRNet; (c) two branch CNN; (d) MAHiDFNet; (e) 所提方法; (f) 真实标签

Fig. 15 Classification results on the Trento dataset when $u = 4$. (a) EndNet; (b) CCRNet; (c) two branch CNN; (d) MAHiDFNet; (e) ours; (f) ground-truth label

表 7 在 Trento 数据集上不同数据源的总体分类精度

Table 7 Overall classification accuracy on the Trento dataset with different data modality unit: %

Data modality	$u = 2$	$u = 3$	$u = 4$	$u = 5$	$u = 6$	$u = 9$	$u = 12$
HSI	91.39	85.97	87.62	92.29	91.86	94.44	95.42
LiDAR	78.69	81.63	85.35	85.83	83.65	87.17	87.21
HSI+ LiDAR	88.12	97.12	97.19	97.54	98.30	99.09	99.23

百分点, 30.80 个百分点, 27.65 个百分点, 23.98 个百分点, 23.43 个百分点, 16.73 个百分点, 11.55 个百分点; 在 Trento 数据集中, 当 u 分别等于 2, 3, 4, 5, 6, 9,

12 时, 同时使用 HSI 和 LiDAR 数据进行训练要比单独使用 HSI 数据进行训练的总体分类精度分别高一 3.26 个百分点, 11.15 个百分点, 9.58 个百分点,

5.25个百分点,6.44个百分点,4.65个百分点,3.81个百分点,比单独使用LiDAR数据进行训练的总体分类精度分别高9.44个百分点,15.50个百分点,11.84个百分点,11.71个百分点,14.65个百分点,11.92个百分点,12.01个百分点。结果表明,在大多数情况下,使用多个数据源进行训练的总体分类精度要高于使用单个数据源的总体分类精度,且单独使用HSI数据进行训练的总体分类精度一般要高于单独使用LiDAR数据进行训练的总体分类精度。

表8 不同样本尺寸的总体分类精度

Table 8 Overall classification accuracy of different sample sizes

unit: %

Dataset	$s=8$	$s=9$	$s=10$	$s=11$	$s=12$	$s=13$	$s=14$	$s=15$
Houston 2013	87.97	89.89	88.34	90.60	90.72	93.13	91.14	94.10
Trento	97.74	97.28	98.06	97.54	97.49	97.47	97.73	97.00

实验中主成分分析降维后的不同维度对实验结果也有一定影响。本文在Houston 2013数据集中各个类别带标签样本的数量为15、在Trento数据集中各个类别带标签样本的数量为5的条件下对不同降维维度进行实验。实验结果如表9所示,其中 k 表示降维后维度大小。在Houston 2013数据集中 $k=11$ 时可以取到最高总体分类精度,在Trento数据集中 $k=5$ 时可以取到最高总体分类精度,设置 $k=7$ 。

表9 不同降维维度的总体分类精度

Table 9 Overall classification accuracy of different reduction

dimensions unit: %

Dataset	$k=1$	$k=3$	$k=5$	$k=7$	$k=9$	$k=11$
Houston 2013	74.94	87.44	91.05	90.60	92.22	93.03
Trento	96.25	96.45	97.59	97.54	96.93	97.39

4 结 论

将高光谱数据与机载激光雷达数据相结合可以提升地物分类的识别精度。为了解决标签样本数减少时带来的分类精度下降问题,设计了一种利用高光谱数据与机载激光雷达数据的对比学习方法,首先利用不带标签的样本对网络进行无监督的预训练,再利用带标签样本对网络进行微调。为了证明所提方法的有效性,在两个包含高光谱数据与机载激光雷达数据的数据集上进行实验,与其他4种融合高光谱数据和机载激光雷达数据的地物分类方法进行对比。结果表明,所提方法可以在带标签数据样本量减少时获得比其他方法更高的分类精度,当标签数据样本量增多时获得与其他方法相当的分精度。说明通过对比学习,所提模型具有了对样本数据的空间特征提取能力,微调带标签的数据样本后即可获得地物分类能力。

参 考 文 献

[1] Chen C, Yan J N, Wang L Z, et al. Classification of

3.2.3 样本尺寸大小及PCA降维维度分析

实验中样本尺寸大小对实验的总体分类精度有一定的影响。本文在Houston 2013数据集中各个类别带标签样本的数量为15、在Trento数据集中各个类别带标签样本的数量为5的条件下对不同样本尺寸大小进行实验。实验结果如表8所示,其中 s 表示样本尺寸大小。在Houston 2013数据集中 $s=15$ 时可以取到最高总体分类精度,在Trento数据集中 $s=10$ 时可以取到最高总体分类精度,设置 $s=11$ 。

urban functional areas from remote sensing images and time-series user behavior data[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 14: 1207-1221.

- [2] Zhang B, Sun X, Gao L R, et al. Endmember extraction of hyperspectral remote sensing images based on the ant colony optimization (ACO) algorithm[J]. IEEE Transactions on Geoscience and Remote Sensing, 2011, 49(7): 2635-2646.
- [3] Su H J, Yao W J, Wu Z Y, et al. Kernel low-rank representation with elastic net for China coastal wetland land cover classification using GF-5 hyperspectral imagery [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 171: 238-252.
- [4] Liu J Y, Fan X, Jiang J, et al. Learning a deep multi-scale feature ensemble and an edge-attention guidance for image fusion[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(1): 105-119.
- [5] Wei W, Zhang J Y, Zhang L, et al. Deep cube-pair network for hyperspectral imagery classification[J]. Remote Sensing, 2018, 10(5): 783.
- [6] Uezato T, Fauvel M, Dobigeon N. Hyperspectral image unmixing with LiDAR data-aided spatial regularization[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(7): 4098-4108.
- [7] Liao W Z, Pižurica A, Bellens R, et al. Generalized graph-based fusion of hyperspectral and LiDAR data using morphological features[J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(3): 552-556.
- [8] Ghamisi P, Rasti B, Benediktsson J A. Multisensor composite kernels based on extreme learning machines[J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(2): 196-200.
- [9] 曹琼, 马爱龙, 钟燕飞, 等. 高光谱-LiDAR多级融合城区地表覆盖分类[J]. 遥感学报, 2019, 23(5): 892-903.
- Cao Q, Ma A L, Zhong Y F, et al. Urban classification by multi-feature fusion of hyperspectral image and LiDAR data[J]. Journal of Remote Sensing, 2019, 23(5): 892-903.
- [10] 王雷光, 耿若箴, 代沁伶, 等. 高光谱-LiDAR融合的条

- 件随机场分类方法[J]. 红外与激光工程, 2021, 50(12): 562-573.
- Wang L G, Geng R Z, Dai Q L, et al. Conditional random field classification method based on hyperspectral LiDAR fusion[J]. Infrared and Laser Engineering, 2021, 50(12): 562-573.
- [11] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [C]//Proceedings of the 25th International Conference on Neural Information Processing Systems, December 3-6, 2012, Red Hook, NY, USA. Red Hook: Curran Associates, 2012: 1097-1105.
- [12] Xu X D, Li W, Ran Q, et al. Multisource remote sensing data classification based on convolutional neural network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(2): 937-949.
- [13] Chen Y S, Li C Y, Ghamisi P, et al. Deep fusion of remote sensing data for accurate classification[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(8): 1253-1257.
- [14] Mohla S, Pande S, Banerjee B, et al. FusAtNet: dual attention based SpectroSpatial multimodal fusion network for hyperspectral and LiDAR classification[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 14-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 416-425.
- [15] Hang R L, Li Z, Ghamisi P, et al. Classification of hyperspectral and LiDAR data using coupled CNNs[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(7): 4939-4950.
- [16] Wu X, Hong D F, Chanussot J. Convolutional neural networks for multimodal remote sensing data classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5517010.
- [17] Hong D F, Gao L R, Hang R L, et al. Deep encoder - decoder networks for classification of hyperspectral and LiDAR data[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 5500205.
- [18] Wang X H, Feng Y N, Song R X, et al. Multi-attentive hierarchical dense fusion net for fusion classification of hyperspectral and LiDAR data[J]. Information Fusion, 2022, 82: 1-18.
- [19] Wang J P, Li J, Shi Y L, et al. AM³Net: adaptive mutual-learning-based multimodal data fusion network[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(8): 5411-5426.
- [20] Roy S K, Deria A, Hong D F, et al. Hyperspectral and LiDAR data classification using joint CNNs and morphological feature learning[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5530416.
- [21] Dong W Q, Zhang T, Qu J H, et al. Multibranch feature fusion network with self- and cross-guided attention for hyperspectral and LiDAR classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5530612.
- [22] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16 × 16 words: transformers for image recognition at scale[EB/OL]. (2021-06-03)[2022-04-15]. <https://arxiv.org/abs/2010.11929v2>.
- [23] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[EB/OL]. (2017-10-06)[2022-04-15]. <https://arxiv.org/abs/1706.03762>.
- [24] 金传, 童常青. 融合 CNN 与 Transformer 结构的遥感图像分类方法[J]. 激光与光电子学进展, 2023, 60(20): 2028006.
- Jin C, Tong C Q. Remote sensing image classification method based on fusion of CNN and transformer structure [J]. Laser & Optoelectronics Progress, 2023, 60(20): 2028006.
- [25] 陈禹汗, 王波, 严清赟, 等. 结合 Transformer 与多尺度残差机制的高光谱遥感分类[J]. 激光与光电子学进展, 2023, 60(12): 1228002.
- Chen Y H, Wang B, Yan Q Y, et al. Hyperspectral remote sensing classification combining Transformer and multiscale residual mechanisms[J]. Laser & Optoelectronics Progress, 2023, 60(12): 1228002.
- [26] Zhao G R, Ye Q L, Sun L, et al. Joint classification of hyperspectral and LiDAR data using a hierarchical CNN and transformer[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5500716.
- [27] Chen T, Kornblith S, Norouzi M, et al. A simple framework for contrastive learning of visual representations [C]//Proceedings of the 37th International Conference on Machine Learning, July 13-18, 2020, Virtual Event. New York: ACM Press, 2020: 1597-1607.
- [28] 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.
- [29] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [30] Debes C, Merentitis A, Heremans R, et al. Hyperspectral and LiDAR data fusion: outcome of the 2013 GRSS data fusion contest[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7(6): 2405-2418.
- [31] Rasti B, Ghamisi P, Gloaguen R. Hyperspectral and LiDAR fusion using extinction profiles and total variation component analysis[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(7): 3997-4007.