

# 跨模态通道权重调整的半监督分类网络<sup>①</sup>

张力予

(南京信息工程大学 计算机学院, 南京 210044)

通信作者: 张力予, E-mail: 2315348216@qq.com



**摘要:** 地物分类是遥感图像领域的重要研究方向, 近年来高光谱图像和激光雷达数据联合分类的技术备受关注. 现有的深度学习模型的性能显著依赖于标注样本的丰富度及优质程度, 这在实际应用中常构成重大挑战. 此外, 很多模型未能有效地利用高光谱图像和激光雷达数据的信息互补性. 针对上述问题, 本文提出了一种跨模态通道权重调整的半监督双分支分类网络, 通过注意力机制, 深入剖析两种数据通道之间的相似度, 并据此自适应地调整各通道的权重. 同时, 本文结合一致性正则化与伪标签的半监督方法, 有效地利用了未标记样本的信息. 在针对 Houston 和 MUUFL 这两个标志性的联合数据集进行高光谱图像与激光雷达数据联合分类的实验中, 所提方法展现出相较于现有分类模型的显著优势, 有效提高了分类精度与效率.

**关键词:** 深度学习; 高光谱图像; 激光雷达数据; 半监督学习

引用格式: 张力予. 跨模态通道权重调整的半监督分类网络. 计算机系统应用, 2025, 34(3): 189–200. <http://www.c-s-a.org.cn/1003-3254/9786.html>

## Semi-supervised Classification Network with Cross-modal Channel Weight Adjustment

ZHANG Li-Yu

(School of Computer Science, Nanjing University of Information Science & Technology, Nanjing 210044, China)

**Abstract:** Terrain classification is a crucial research direction in remote sensing imagery. The technology of joint hyperspectral images and LiDAR data classification has drawn much attention in recent years. The classification performance of existing deep learning models significantly depends on the richness and quality of labeled samples, which often poses a major challenge in practical applications. In addition, many models fail to effectively utilize the information complementarity between hyperspectral images and LiDAR data. To solve the above problems, this study proposes a semi-supervised double-branch classification network with cross-modal channel weight adjustment. Through the attention mechanism, the similarity between two data channels is analyzed deeply, and the weight of each channel is adaptively adjusted accordingly. At the same time, the semi-supervised method of consistency regularization and pseudo-labeling is combined to effectively utilize the information of unlabeled samples. In the experiment of joint classification of hyperspectral images and LiDAR data on the two iconic joint datasets of Houston and MUUFL, the proposed method shows significant advantages over existing classification models, effectively improving classification accuracy and efficiency.

**Key words:** deep learning; hyperspectral image (HSI); LiDAR data; semi-supervised learning

利用遥感影像数据进行地物分类是一项具有广泛应用前景的重要研究领域, 它在城市规划<sup>[1]</sup>、矿物勘

探<sup>[2]</sup>、环境监测<sup>[3]</sup>等领域具有重要意义. 随着遥感技术的发展, 可以通过一系列传感器同时收集潜在土地覆

<sup>①</sup> 收稿时间: 2024-08-02; 修改时间: 2024-08-27, 2024-09-19, 2024-09-24; 采用时间: 2024-09-30; csa 在线出版时间: 2025-01-21  
CNKI 网络首发时间: 2025-01-22

盖目标的描述,如高光谱(HS)<sup>[4]</sup>、激光雷达(LiDAR)<sup>[5]</sup>、合成孔径雷达(SAR)<sup>[6]</sup>、多光谱(MS)<sup>[7]</sup>等。通过分析遥感影像数据,可以识别和分类不同类型的地物,如建筑物、道路、水体、植被等。起初,研究人员尝试使用单源传感器数据的特定优势对地物分类,例如,高光谱图像的光谱信息经常被用于根据反射率值来识别材料<sup>[8]</sup>。然而,高光谱图像很难准确区分具有相同材料但不同高度组成的地物,比如房屋屋顶和水泥路面<sup>[9]</sup>。所以,研究人员引入能够提供高程信息的激光雷达数据。因此,如何高效地整合高光谱图像和激光雷达数据,一直是地物分类领域的热门问题<sup>[10]</sup>。

在研究的早期阶段,处理高光谱和激光雷达数据融合问题常见的方法包括传统技术:如支持向量机<sup>[11]</sup>、条件随机场<sup>[12]</sup>和马尔可夫随机场<sup>[13]</sup>。然而,这些传统的浅层模型在挖掘多模态遥感数据的潜在信息方面表现并不理想。随着深度学习在特征提取和表示能力上的显著优势日益凸显,其在多模态处理中的应用受到广泛关注<sup>[14,15]</sup>。为了有效整合高光谱图像(hyperspectral image, HSI)与激光雷达数据(LiDAR)的特征,研究者们探索了多种策略。例如, Xu等<sup>[16]</sup>设计了一种双分支卷积神经网络架构,该架构独立地提取两种数据源的特征,以此实现特征的深度融合。为了进一步优化多源数据的融合效率, Hang等<sup>[17]</sup>构建了一个在特征级与决策级上实现了两个卷积神经网络分支的紧密融合的具有共享机制的耦合CNN模型。据此启发, Li等<sup>[18]</sup>在特征提取阶段引入了特征调制模块,有效利用了模态之间的补充信息,提升了光谱和空间特征的判别能力,促进了分类性能的显著提高。

上述方法虽已关注多模态信息互补的问题,但很多是监督学习模型,其分类性能受标注样本的数量和优质程度的影响。然而,地物的分布日益复杂,这使得仅靠有限的资源和人工去获取大量优质标注数据变得尤为困难<sup>[19]</sup>。面对这一挑战,众多学者开始探索从未经标注的样本中挖掘特征的新途径,旨在降低受标注样本的影响。就像 Chen等将自编码器模型引入于高光谱分类任务,通过数据重构有效地提取特征<sup>[20]</sup>。继而, Hong等<sup>[21]</sup>将此理念用于LiDAR和HSI联合分类领域上,设计了一种深度编码器-解码器网络架构(EndNet),该架构独立重构各模态信息,通过多个全连接层融合多模态特征。Zhang等<sup>[22]</sup>为了区别于单模态数据自身重构,设计了一种可以实现从HSI重建LiDAR的跨模态重

构网络。杭仁龙等<sup>[23]</sup>为增强特征判别能力,提出一种基于匹配学习的半监督分类模型来减少对标记样本的依赖。为了获取准确的分类结果,提取和融合多模态数据中包含的相似的高阶语义信息和互补的判别信息是很重要的。为此, Lu等<sup>[24]</sup>提出了一种用于HSI与LiDAR联合分类的基于耦合对抗学习的方法,该方法通过训练一个无监督的特征学习子网络从两种模态数据中有效的学习高阶语义特征。同时,训练一个有监督的多层特征融合分类子网络,并通过采用自适应概率融合策略进一步提高分类性能。

以上无监督模型和半监督模型解决了标注样本少的问题,但是却并没有充分利用多模态的互补信息。综合表述,我们提出一种跨模态通道权重调整的半监督双分支分类网络,该网络通过注意力机制,深入剖析高光谱和激光雷达数据通道之间的相似度,并据此自适应地调整各通道的权重,有效增强信息的互补性。本文还通过弱增强处理未标记样本以生成伪标签,并与强增强未标记样本预测结果进行一致性损失计算,有效地利用了未标注样本的信息,极大地减少了模型参数量。在针对Houston和MUUFL这两个标志性的联合数据集进行高光谱数据与激光雷达数据联合分类的实验中,我们提出的方法展现出了相较于现有分类模型的显著优势,有效提高了分类精度与效率。

## 1 研究方法

### 1.1 整体结构

图1展示了本文方法的整体流程图。假设 $M_H \in R^{C \times H \times W}$ 和 $M_L \in R^{H \times W}$ 分别代表同一区域的HSI图像和LiDAR数据,其中 $W$ 和 $H$ 分别表示它们的宽度和高度, $C$ 表示HSI的光谱波段数。由于HSI通常包含数百个连续的光谱波段,这些波段之间存在显著的高维冗余性且数据量很大。所以,本文采用主成分分析(PCA)对HSI进行降维处理<sup>[25]</sup>,通过将高维数据投影到低维空间以保留尽可能多的信息。PCA是一种经典的线性降维方法,通过将高维数据投影到一个新的正交坐标系中,来保留数据中方差最大的方向。其思想是将原始高维数据的协方差矩阵进行特征分解,选择特征值最大的前 $d$ 个主成分,形成新的低维空间。这些主成分能够最大化原始数据中的方差,确保尽可能多的信息被保留。提取前 $d$ 个主成分后的高光谱数据的表示为 $M'_H \in R^{d \times H \times W}$ 。LiDAR数据只有1个通道,不需要进行

降维处理. 从图1可以看出, 本文方法采用双分支结构, 其中每个分支都是同时接收标注数据和未标注数据进

行训练. 假设 HSI 和 LiDAR 的标注数据分别表示为  $M'_{HL}$  和  $M_{LL}$ , 未标注数据分别表示为  $M'_{HU}$  和  $M_{LU}$ .

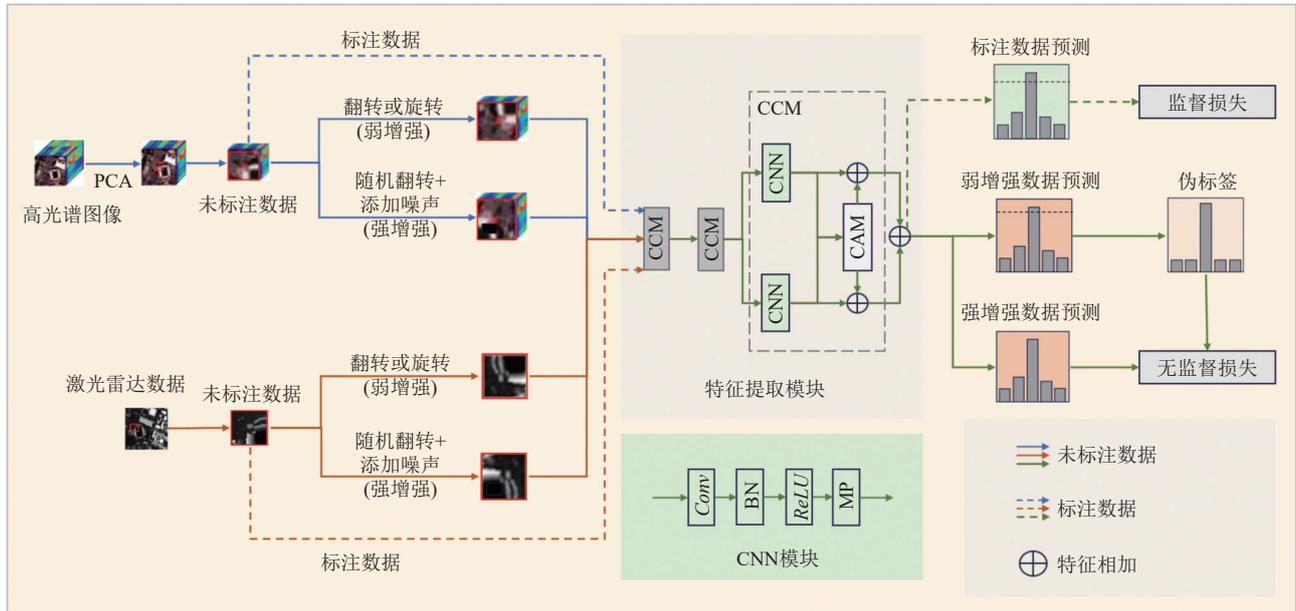


图1 整体结构图

对于标注数据, 同位置的高光谱数据  $M'_{HL}$  和激光雷达数据  $M_{LL}$  一同进入跨模态通道权重调整提取特征模块, 进行标准的监督学习, 预测值与真实值进行交叉熵计算作为监督损失. 对于同位置的高光谱和激光雷达的未标注数据  $M'_{HU}$  和  $M_{LU}$ , 我们采用两种数据增强方式, 两种增强后的未标注数据一同进入跨模态通道权重调整特征提取模块得到特征表示, 然后将两个模态的同种增强特征表示融合, 得到对应的两种增强数据预测值进行一致性正则化, 交叉熵计算正则化损失作为无监督损失. 整个训练过程的总损失为无监督损失与监督损失之和, 通过反向传播算法对模型进行迭代优化.

## 1.2 跨模态通道权重调整特征提取模块

鉴于现有模型在处理高光谱图像与激光雷达数据时, 往往忽视了各模态每个通道之间的相关性, 所以, 本文提出一种跨模态通道权重调整特征提取方法, 利用通道注意力机制剖析两种模态数据通道之间的相似度, 并据此自适应地调整各通道的权重, 增强信息互补性.

根据图1所示, 特征提取模块同样选择双分支结构, 由3个相同的CCM (CAM and CNN module) 模块构成, 每个CCM模块由双分支的CNN模块以及一个通道调和模块 (channel adjustment module, CAM) 构成.

CNN模块用于学习HSI的光谱特征和LiDAR的

高程特征. 第1个CNN利用32个 $3 \times 3$ 的卷积核提取两者的初始特征, 然后依次利用批归一化 (BN) 加快网络收敛<sup>[26]</sup>、ReLU激活增强非线性表达能力、最大池化 (MP) 扩大感受野. 为简化网络结构, 后两个CCM中的CNN共享参数.

图2描述了通道调和模块的详细结构. 在这个模块中, 我们使用了通道注意力机制来分析HSI和LiDAR两个模态特征的通道之间的相关性. 具体来说, 经过CNN模块提取的HSI特征和LiDAR特征分别表示为  $X_H \in R^{C \times H \times W}$  和  $X_L \in R^{C \times H \times W}$ , 第1个CCM模块时,  $C$  都为第1个CNN的卷积核数32.

为了进行后续的通道相关性分析, 我们将  $X_H$  和  $X_L$  分别重新调整形状为  $X'_H \in R^{C \times HW}$  和  $X'_L \in R^{C \times HW}$ ,  $X'_H$  作查询Q矩阵,  $X'_L$  作键值K矩阵, QK矩阵进行矩阵乘法操作, 得到了相似度矩阵  $X \in R^{C \times C}$ , 其中每一个元素反映了HSI和LiDAR数据中对应通道之间的相似度系数.

接下来, 我们通过计算相似度矩阵的行 (列) 的总和得  $X_R \in R^{C \times 1}$  ( $X_C \in R^{1 \times C}$ ),  $X'_R$  ( $X'_C$ ) 代表HSI (LiDAR) 数据的第  $i$  个通道和另一模态数据的所有通道的相似度之和, 此值越大, 表示HSI (LiDAR) 的第  $i$  个通道与另一个模态的所有通道相关性越高, 也就说明, HSI (LiDAR)

数据的第*i*个通道所包含的信息是相对冗余的。

根据先验知识,冗余信息可能会导致特征可分辨性的降低,进而影响后续分类的准确率。因此,我们要将含冗余信息多的通道权重降低,含冗余信息少的通道权重提高。所以,在通道调和模块中引入了一种基于通道相关性的特征掩码机制,利用一种负反馈映射函

数 $f(x)$ 将相似度系数转换为 $[0, 1]$ 范围内的通道权重。这些权重被用作特征掩码,通过与原特征数据进行逐元素相乘,实现对冗余通道的抑制和对互补通道的增强,以此来实现自适应地调整各通道的权重。经过通道调和模块的HSI和LiDAR数据表示为 $X_H'' \in R^{C \times H \times W}$ 和 $X_L'' \in R^{C \times H \times W}$ 。

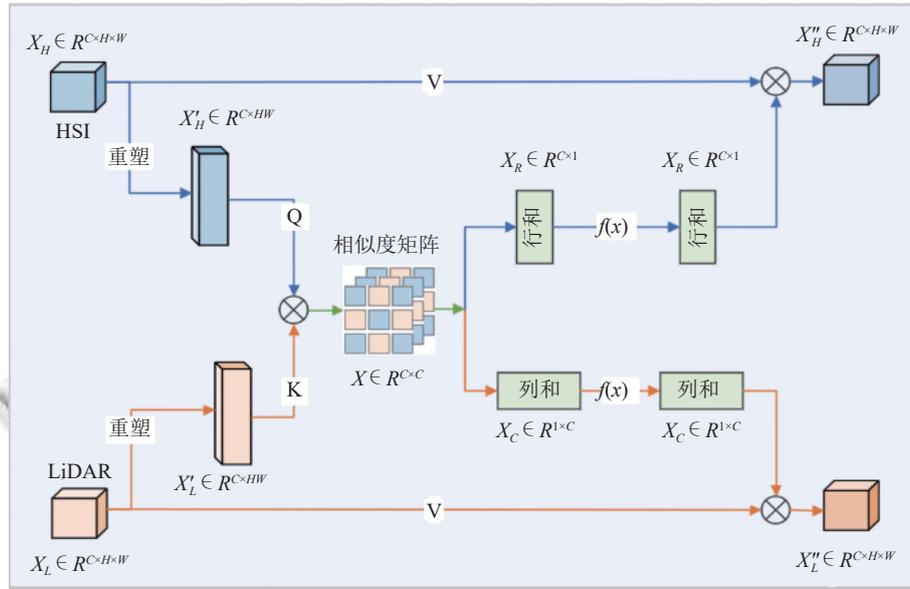


图2 CAM模块结构图

负反馈映射函数 $f(x)$ 的公式如下:

$$f(x) = 1 - \text{Sigmoid}(x) = \frac{e^{-x}}{1 + e^{-x}} \quad (1)$$

与 $\text{Sigmoid}(x)$ 函数不同,它是一个负反馈映射,即输入越大,映射值越小,反之亦然。

假设第*i* ∈ {0, 1, 2}层特征表示为 $F_j^i$ ,  $j \in \{H, L\}$ ,则经过CCM模块后,第*i* + 1层特征为:

$$F_j^{i+1} = \text{CAM}(\text{Conv}(F_j^i)) \quad (2)$$

其中,  $\text{CAM}(\cdot)$ 表示通道调和模块的操作,  $\text{Conv}(\cdot)$ 代表 $3 \times 3$ 的卷积,  $F_j^0$ 代表输入图像块。通过相加 $F_H^3$ 和 $F_L^3$ 融合两个模态的特征,然后经过两个神经元为128和2的全连接层得到弱增强数据的预测结果和强增强数据的预测结果。上述流程可简化为以下表达式:

$$y_w = \text{Softmax}\{f_2(\text{ReLU}(f_1([F_H^3; F_L^3]; W_1)); W_2)\} \quad (3)$$

其中,  $[\cdot; \cdot]$ 表示特征相加,  $W_1$ 和 $W_2$ 为2个全连接层对应的连接权重,  $f_1(\cdot)$ 和 $f_2(\cdot)$ 分别表示2个全连接层,  $y_w$ 为弱增强数据的预测结果,同理,  $y_s$ 和 $y$ 分别为强增强

数据的预测结果和标记数据预测结果。

### 1.3 一致性正则化和伪标签

为了解决标记样本不足的问题,我们采用一致性正则化<sup>[27]</sup>和伪标签<sup>[28]</sup>两种方法相结合的方式,应用于两种模态的联合分类任务上,如图1右侧所示。

我们对高光谱图像的未标注数据 $M'_{HU}$ 和激光雷达数据的未标注数据 $M'_{LU}$ 分别进行了两种不同的数据增强,一种是对数据在空间维度上进行翻转或旋转的弱增强方式,生成高光谱数据和激光雷达数据的弱增强数据分别表示为 $M'_{HW}$ 和 $M'_{LW}$ ,另一种是对数据随机地进行翻转并添加噪声的强增强方式,生成高光谱数据和激光雷达数据的强增强数据分别表示为 $M'_{HS}$ 和 $M'_{LS}$ 。根据一致性正则化的理论,这两种增强后的数据的模型预测结果应该是相似的。根据先验知识,模型对弱增强的输入预测会更准确,所以用两种弱增强数据 $M'_{HW}$ 和 $M'_{LW}$ 经过特征提取模块融合后的预测结果 $y_w$ 来生成伪标签 $\hat{y}_w$ ,与强增强数据的预测结果,  $y_s$ 进行交叉熵损失计算得无监督训练损失 $l_u$ 。

而使用伪标签就会涉及正确性和覆盖率的问题. 伪标签作为由模型初步预测生成的标签, 其准确性无法完全保证, 常伴随着误差. 这些含有错误的伪标签在后续的训练过程中, 若未经适当处理即被用作监督信号, 可能会误导模型的优化. 所以, 我们设置置信度阈值 $\tau$ 来选择可靠的伪标签, 若 $y_w \geq \tau$ , 则用作伪标签 $\hat{y}_w$ .

#### 1.4 模型损失

本文作为一个半监督学习方法, 损失函数 $l$ 由两个损失函数监督损失 $l_s$ 和无监督损失 $l_u$ 组成. 具体来说, 监督损失 $l_s$ 由标记数据进行正常的监督学习得到:

$$l_s = \frac{1}{B} \sum_{b=1}^B H(\hat{y}, y) \quad (4)$$

其中,  $\hat{y}$ 为真实标签,  $y$ 为标记数据的预测结果,  $B$ 为标记样本批量大小,  $H(\cdot, \cdot)$ 是交叉熵损失函数. 无监督训练损失 $l_u$ 可以表示为

$$l_u = \frac{1}{\omega B} \sum_{b=1}^{\omega B} (\max(y_w) \geq \tau) H(\hat{y}_w, y_s) \quad (5)$$

其中,  $H(\cdot, \cdot)$ 是交叉熵损失函数,  $\omega$ 为网络输入的无标记样本与标记样本的比例系数,  $B$ 为标记样本批量大小,  $\tau$ 是一个代表阈值的标量超参数,  $\hat{y}_w$ 为伪标签,  $y_s$ 为增强数据的预测结果.

所以, 本文的总损失函数如式(6):

$$l = l_s + \lambda l_u \quad (6)$$

其中,  $\lambda$ 表示无监督损失的相对权重, 是固定的标量超参数, 本文设置为1.

## 2 实验

### 2.1 数据集

本文选取了两个常用的高光谱和激光雷达联合数据集来进行实验, 分别为 Houston 数据集<sup>[29]</sup>和 Muufl-Gulfport (MUUFL) 数据集<sup>[30]</sup>.

Houston2013 数据集源自 2013 年 IEEE GRSS 数据融合大赛, 覆盖休斯顿大学校园及周边, 由 2012 年资助的国家机载激光测绘中心收集. 此数据集的 HSI 和 LiDAR 均具有 2.5 m 的空间分辨率, 数据都由 349×1905 像素组成, 该数据集描述了 15 种不同的地物类别, HSI 尤其丰富, 包含 144 个光谱波段. 对于同一区域, 激光雷达数据只提供一个频带. 表 1 详细列举了此数据集的样本分布. 此外, 图 3(a)–(e) 表示 HSI 的伪彩

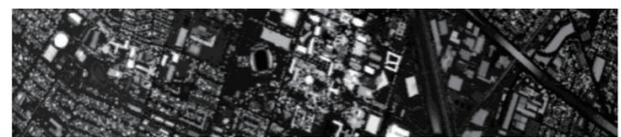
色复合图、LiDAR 的灰度图、标记集样本、测试集样本分布图和各类标签对应的颜色标记.

表 1 Houston 数据集的样本分布

类别编号	地物类别	标记样本	无标记样本	测试样本
1	健康草坪	10	1251	1053
2	非健康草坪	10	1254	1064
3	合成草坪	10	697	505
4	树木	10	1244	1056
5	土壤	10	1242	1056
6	水体	10	325	143
7	居民区	10	1268	1072
8	商业区	10	1244	1053
9	道路	10	1252	1059
10	高速公路	10	1227	1036
11	铁路	10	1235	1054
12	停车场1	10	1233	1041
13	停车场2	10	469	285
14	网球场	10	428	247
15	跑道	10	660	473
总样本数		150	15029	12197



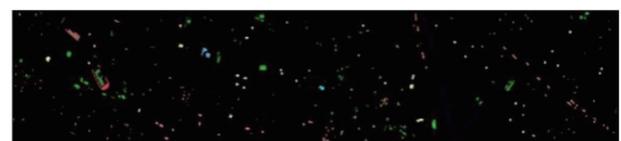
(a) HSI 伪彩色复合图



(b) LiDAR 的灰度图



(c) 标记集样本



(d) 测试集样本分布图



(e) 各类标签对应的颜色标记

图 3 Houston 数据集的可视化

MUUFL 数据集采集于 2010 年 11 月, 覆盖南密西西比州大学的 Gulfport 校区. HIS 和 LiDAR 的空间分

分辨率均为 1 m、图像尺寸为 325×220 像素、包含 11 种地物类别。HSI 有 72 个光谱波段, 由于噪声过大, 第 1 个和后 8 个频谱波段被移除。表 2 详细列举了此数据集的样本分布。图 4(a)–(e) 表示 HSI 的伪彩色图、LiDAR 的灰度图、标记集样本、测试集分布图和各类标签对应的颜色标记。

表 2 MUUFL 数据集样本分布

类别编号	地物类别	标记样本	无标记样本	测试样本
1	树木	10	23 246	23 096
2	大多数草	10	4 270	4 120
3	复合地层	10	6 882	6 732
4	尘土和沙子	10	1 826	1 676
5	道路	10	6 687	6 537
6	水体	10	466	316
7	建筑物阴影	10	2 233	2 083
8	建筑	10	6 240	6 090
9	人行道	10	1 385	1 235
10	黄色路沿石	10	183	33
11	棉布板	10	269	119
总样本数		150	53 687	52 037

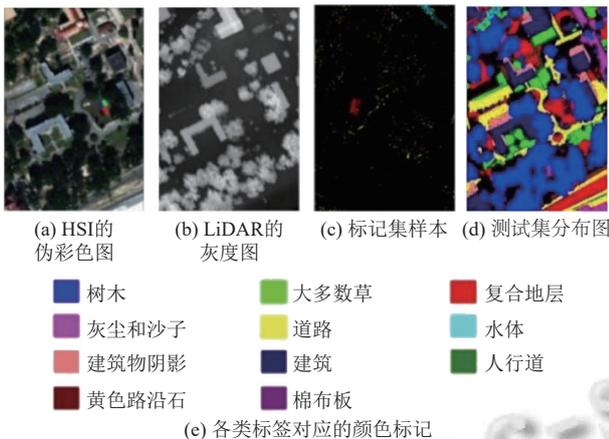


图 4 MUUFL 数据集的可视化

## 2.2 实验设置

在模型训练阶段, 用标记集及其对应标签和无标记集同时训练。在推理阶段, 使用测试集及对应的标签信息评估模型性能。分类指标包括平均精度 (AA)、总体精度 (OA) 和 Kappa 系数<sup>[23]</sup>。

本文在 PyTorch 框架上<sup>[31]</sup>进行实验, 使用配置了一张 NVIDIA GeForce RTX 3090 显卡的 Ubuntu 20.04 系统服务器。为了优化网络, 本文选择了 Adam 优化器作为初始优化器。在网络训练过程中, 两个数据集批量大小 (batch size) 均为 32, 学习率分别为 0.001 和 0.01, 训练轮数均设置为 200。

## 2.3 超参数分析

超参数的选取对模型性能有着很大影响, 本文深入探讨了若干关键超参数的优化选择: 构造伪标签时的置信度阈值  $\tau$ 、PCA 中主成分个数  $d$ 、网络输入的无标记样本与标记样本的比例系数  $\omega$ 、输入图像块大小  $s \times s$ 。

### 2.3.1 构造伪标签的置信度阈值 $\tau$

为了分析构造伪标签时置信度阈值  $\tau$  对网络分类性能产生的影响, 固定其余超参数, 从候选集 {0.88, 0.9, 0.92, 0.95, 0.97} 中逐一选取其值。图 5 显示了不同  $\tau$  值对 OA 值的影响。图 5(a) 和 (b) 所示,  $\tau$  从 0.88 增加到 0.95 时, OA 值一直都是提升的, 而从 0.95 增加到 0.97 时, OA 值则会下降。这说明  $\tau$  设置成 0.95 时, OA 值达到最高。因此, 在 Houston 和 MUUFL 数据集上, 置信度阈值  $\tau$  均设置为 0.95, 此时模型最优。

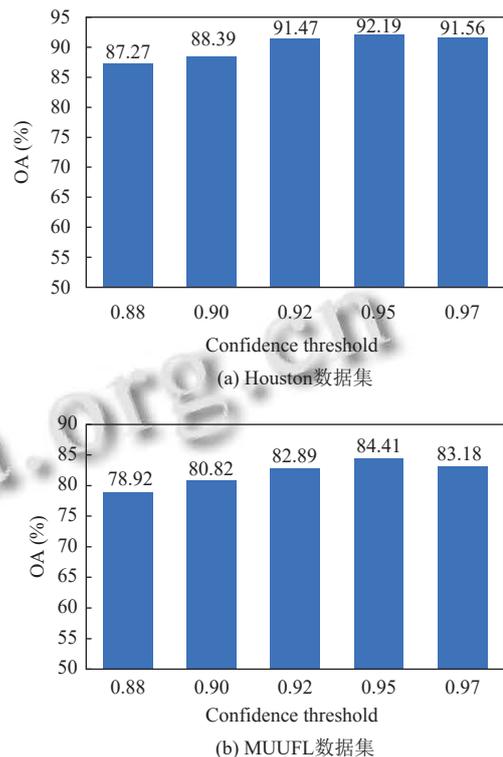


图 5 在 Houston 数据集和 MUUFL 数据集上不同构造伪标签的置信度阈值的精度对比

### 2.3.2 主成分个数 $d$ 的分析

PCA 的碎石图展示了每个主成分对数据方差的贡献比例。观察图 6 碎石图, 我们可以发现当主成分数量达到第 5 个时, 曲线趋于平缓, 这意味着前 5 个主成分已经基本包含了数据中的主要信息。鉴于 HSI 数据的

高光谱维与潜在冗余,我们选择从第5个主成分开始考虑,以减少数据量同时保留关键信息.然而,考虑更多主成分可能有细节或额外特征,我们选取候选集合{1, 5, 10, 15, 20, 25, 30, 35}作为主成分数量的评估范围,以探索主成分个数对实验的影响.

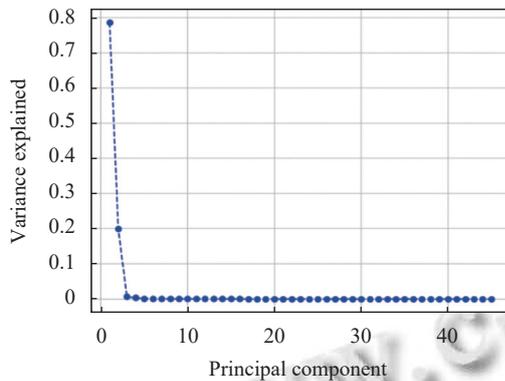


图6 Houston数据集的主成分碎石图

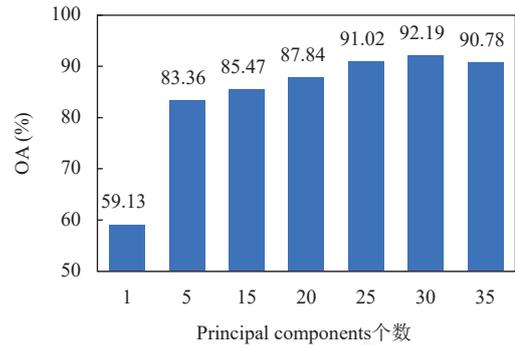
同对 $\tau$ 的分析,固定其他超参数,从候选集合{1, 5, 10, 15, 20, 25, 30, 35}中逐一选取 $d$ 值分析其影响.图7显示了不同 $d$ 值对OA值的影响.随着 $d$ 值增加到5时,OA变化极大.如图7(a)所示,Houston数据集上所体现的 $d$ 增加到30时,OA值是最高的为92.19%,继续增加 $d$ 值,OA值则会降低.同样如图7(b)所示,在MUUFL数据集上 $d$ 增加到30时,OA值是最高的为84.41%.因此,在两个数据集上的最优 $d$ 值均为30.

### 2.3.3 网络输入的无标记样本与标记样本比例系数 $\omega$

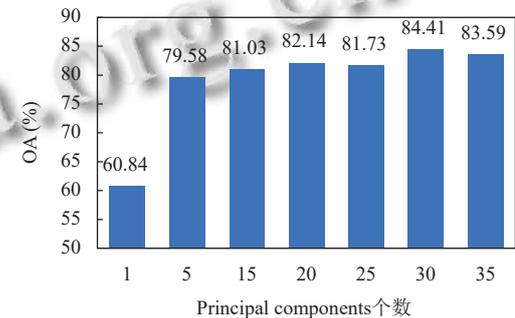
网络中输入的无标记样本与标记样本的比例为 $2\omega+1$ .同对 $\tau$ 的分析,固定其他超参数,从候选集合{5, 6, 7, 8, 9}中逐一选取 $\omega$ 值以作分析.图8显示了不同 $\omega$ 值对OA值的影响.如图8(a)所示, $\omega$ 的取值增加到7时,OA值为最高为92.19%. $\omega$ 继续增加,OA值会持续下降.如图8(b)所示, $\omega$ 的取值增加到8时,OA值为最高为84.41%.因此,在两个数据集上网络输入的无标记样本与标记样本最优比例 $\omega$ 系数分别为7和8.

### 2.3.4 图像块大小 $s \times s$ 的分析

同对 $\tau$ 的分析一样,固定其他超参数,从候选集合{ $9 \times 9$ ,  $11 \times 11$ ,  $13 \times 13$ ,  $15 \times 15$ }中逐一选取 $s$ 值以作分析.图9显示了不同 $s$ 值对OA值的影响.从图9(a)和(b)可分析,当 $s$ 的值为11时,两个数据集的OA值都是最高.所以,在两个数据集上最合适的输入图像块大小均为11.

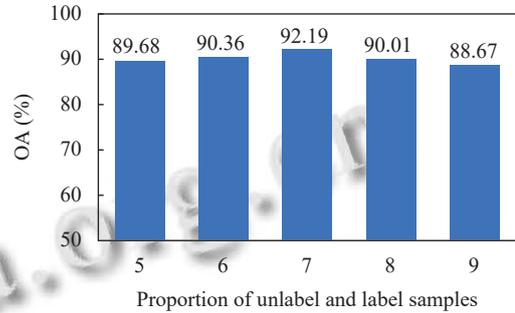


(a) Houston数据集

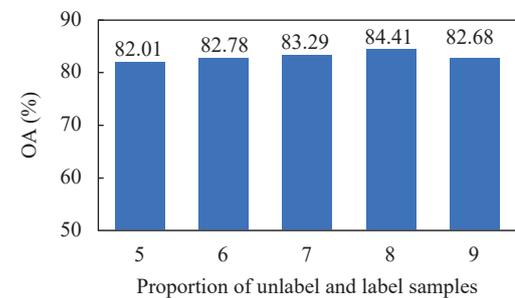


(b) MUUFL数据集

图7 在Houston数据集和MUUFL数据集上不同主成分数的精度对比



(a) Houston数据集



(b) MUUFL数据集

图8 在Houston数据集和MUUFL数据集上不同输入无标记样本与标记样本的比例的精度对比

## 2.4 模型对比

为验证方法的有效性,本文与多个HSI和LiDAR

联合分类模型在 Houston 和 MUUFL 数据集上进行对比,选取的对比模型有 2 个代表性的监督模型(即 EMFNet 和 HCT)、5 个无监督模型(即 CAE\_H、CAE\_HL、IP\_CNN、PToPCNN 和 SS\_MAE) 和 1 个半监督模型(即 CALC). 对以上模型的说明如下.

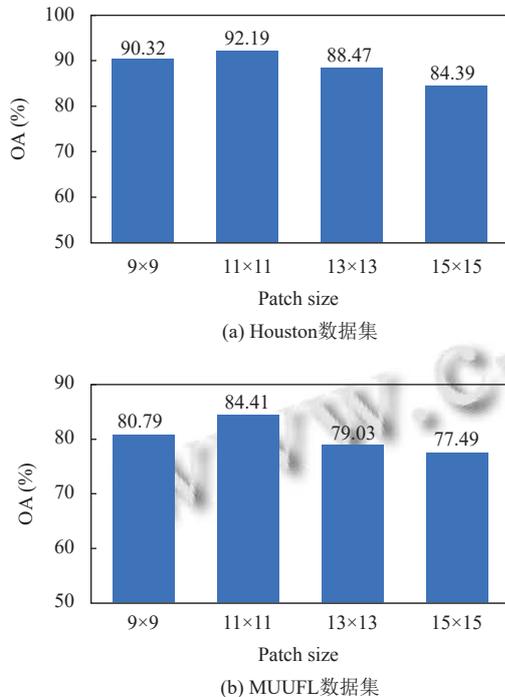


图9 在 Houston 和 MUUFL 数据集上输入图像块大小的精度对比

(1) EMFNet: 该网络设计融合多模态互补信息,增强高程及空间特征判别能力.

(2) HCT: 该网络提出了一种由分层卷积神经网络和 Transformer 网络组成的双分支方法,来融合数据特征.

(3) CAE\_H: 该架构以卷积自动编码器为主,经过 3 个卷积层和反卷积层,分别构造编码器和解码器.

(4) CAE\_HL: 网络架构沿用 CAE\_H, 输入调整为 HSI 和 LiDAR 图像块的串联结果.

(5) IP\_CNN: 为确保实验公平性, 本文遵循 Zhang 等<sup>[32]</sup>提出的架构.

(6) PToPCNN: 它以 HSI 图像作为输入数据, 经过重构分支获取 HSI 和 LiDAR 的准确表示, 再将融合特征用于分类<sup>[22]</sup>.

(7) SS\_MAE: 该网络提出了一种用于高光谱与激光雷达数据联合分类的空间光谱掩码自编码器. 为保证实验公平性, 本文与 Lin 等<sup>[33]</sup>提出结构一致.

(8) CALC: 通过训练无监督网络提取高级语义特征, 并通过监督网络进行融合分类.

表 3 与表 4 分别量化对比了 Houston 和 MUUFL 数据集上所选对比模型的实验结果, 加粗字体表示最高分类精度. 图 10 与图 11 则直观展示了分类结果的对比, 用颜色区分地物类型.

表 3 Houston 数据集上各模型的性能 (%)

Class No.	EMFNet	HCT	CAE_H	CAE_HL	IP_CNN	PToPCNN	SS_MAE	CALC	Ours
1	79.68	93.83	82.94	81.99	<b>98.21</b>	92.87	85.09	84.71	83.09
2	87.69	91.82	74.61	92.03	86.45	82.06	<b>99.72</b>	57.61	99.62
3	95.05	97.62	94.85	92.32	84.79	83.33	<b>100.00</b>	92.47	<b>100.00</b>
4	79.43	93.65	99.22	99.41	97.72	97.12	99.05	78.98	<b>100.00</b>
5	95.15	<b>100.00</b>	94.79	99.81	92.03	98.58	99.43	93.65	99.34
6	64.34	94.40	79.02	85.17	88.36	74.35	95.80	71.33	<b>100.00</b>
7	80.04	80.04	81.47	89.48	95.30	94.49	77.23	84.98	<b>95.61</b>
8	78.77	84.33	76.05	86.82	73.07	84.32	87.74	91.64	<b>92.88</b>
9	60.30	74.69	33.28	44.04	<b>85.07</b>	73.78	75.73	76.01	79.04
10	44.81	<b>93.92</b>	47.70	53.15	69.35	75.56	72.39	74.03	89.09
11	84.63	78.46	62.69	83.32	76.23	82.97	90.70	90.51	<b>94.12</b>
12	71.03	63.98	28.11	37.10	71.39	78.20	78.67	49.95	<b>78.77</b>
13	81.54	79.65	86.53	88.42	73.78	<b>94.58</b>	90.53	74.74	92.98
14	87.45	95.14	52.47	66.72	87.97	<b>100.00</b>	<b>100.00</b>	85.42	98.38
15	88.29	<b>100.00</b>	98.60	99.70	96.25	97.23	<b>100.00</b>	79.70	<b>100.00</b>
OA	77.68	86.71	70.71	78.52	84.03	86.36	88.15	78.87	<b>92.19</b>
AA	78.55	88.10	72.82	79.97	85.06	87.30	90.14	79.05	<b>93.53</b>
Kappa	75.83	85.58	68.28	76.71	82.67	85.19	87.14	77.08	<b>91.52</b>

基于上述数据, 可得出以下结论.

对于两个监督特征提取模型, HCT 在两个数据集

上的性能指标均显著优于 EMFNet, 其优势主要归因于 HCT 不仅有 Transformer 提取全局特征的优势, 又

保留了 CNN 提取细节特征的能力,可以更好地融合两种模态数据. 由于我们的模型有效利用了未标注样本

的信息,在实验结果上看,我们的性能上要比 EMFNet 和 HCT 要好很多.

表 4 MUUFL 数据集上各模型的性能 (%)

Class No.	EMFNet	HCT	CAE_H	CAE_HL	IP_CNN	PToPCNN	SS_MAE	CALC	Ours
1	84.60	86.61	75.28	83.96	91.02	<b>97.68</b>	87.24	89.93	93.47
2	67.39	<b>81.50</b>	63.51	74.81	49.14	51.48	70.61	57.81	71.82
3	42.18	57.03	32.30	30.89	63.95	<b>64.29</b>	60.84	40.55	60.75
4	69.99	<b>95.64</b>	86.07	89.43	61.12	48.93	87.41	81.14	93.43
5	73.29	68.75	80.93	83.84	<b>96.17</b>	88.99	91.85	64.55	79.10
6	82.41	96.83	<b>100.00</b>	<b>100.00</b>	18.59	36.72	<b>100.00</b>	98.10	99.37
7	74.82	66.59	78.72	<b>90.95</b>	59.86	49.90	80.89	74.99	78.78
8	85.49	93.96	89.88	93.05	88.14	95.92	<b>96.60</b>	96.06	95.70
9	32.96	15.79	24.86	34.48	28.42	26.37	38.38	41.05	<b>50.85</b>
10	57.58	12.12	36.36	47.27	0.33	1.07	3.03	24.24	<b>72.72</b>
11	81.34	96.64	<b>99.16</b>	<b>99.16</b>	39.72	22.20	97.48	97.48	95.80
OA	74.31	78.84	70.66	76.81	74.08	75.37	82.82	76.51	<b>84.41</b>
AA	68.37	70.13	69.73	75.26	54.22	53.05	74.03	69.63	<b>81.07</b>
Kappa	66.75	72.52	63.37	70.34	66.42	68.43	77.61	69.24	<b>79.47</b>

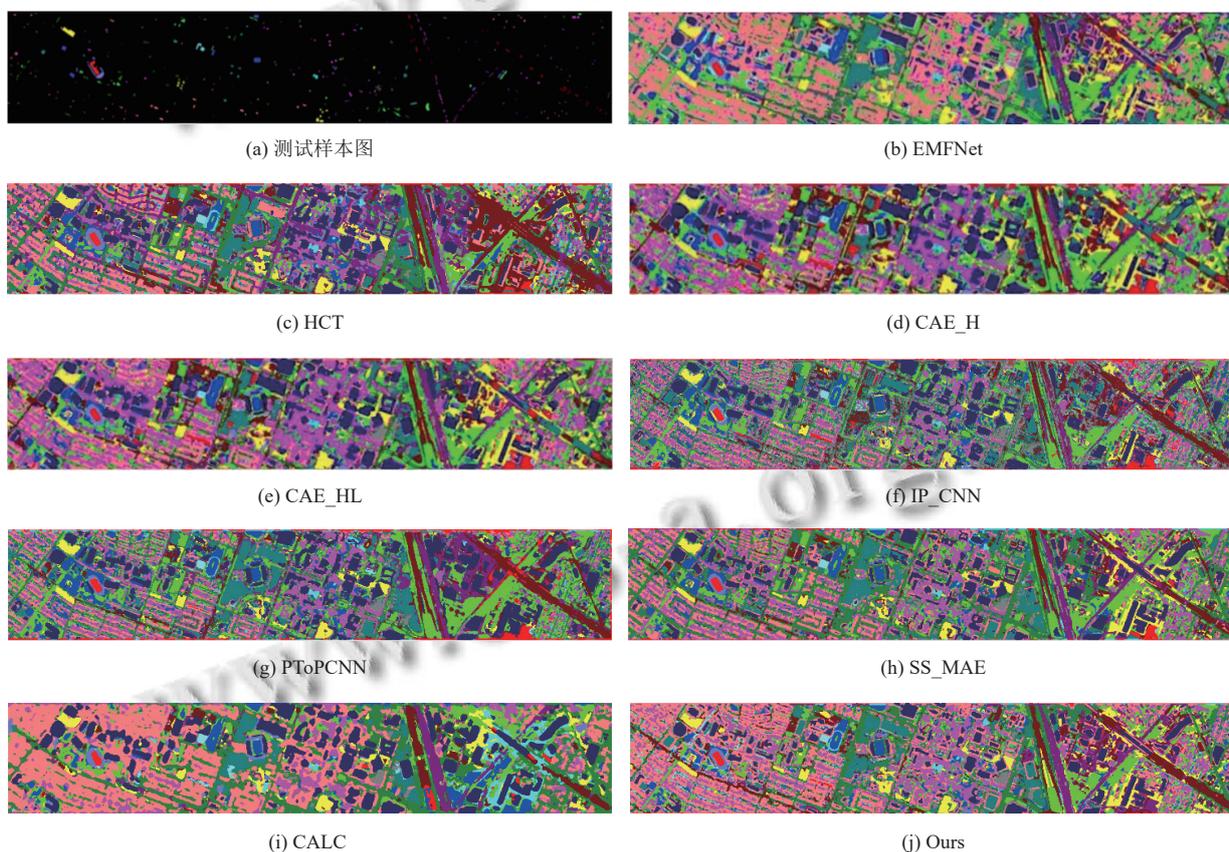


图 10 Houston2013 数据集上不同模型的性能结果

对于其他模型来说, CAE\_HL 的精度要比 CAE\_H 高得多, 这证明了激光雷达数据在土地分类上起了非常重要的作用. SS-MAE 提出的光谱和通道掩码重建模型在 MUUFL 数据集上的精度有了很大提升, OA、AA 和 Kappa 值分别为 82.82%、74.03% 和 77.61%. 另

外, 拥有重构数据分支和融合的 PToPCNN<sup>[22]</sup>比 IP-CNN 模型的精度更高, OA 值分别提升了 2.33% 和 1.29%. 从总体上看, 比起其他无监督模型和半监督模型, 我们的模型由于充分利用了多模态间的互补信息, 所以获得了更高的分类精度.

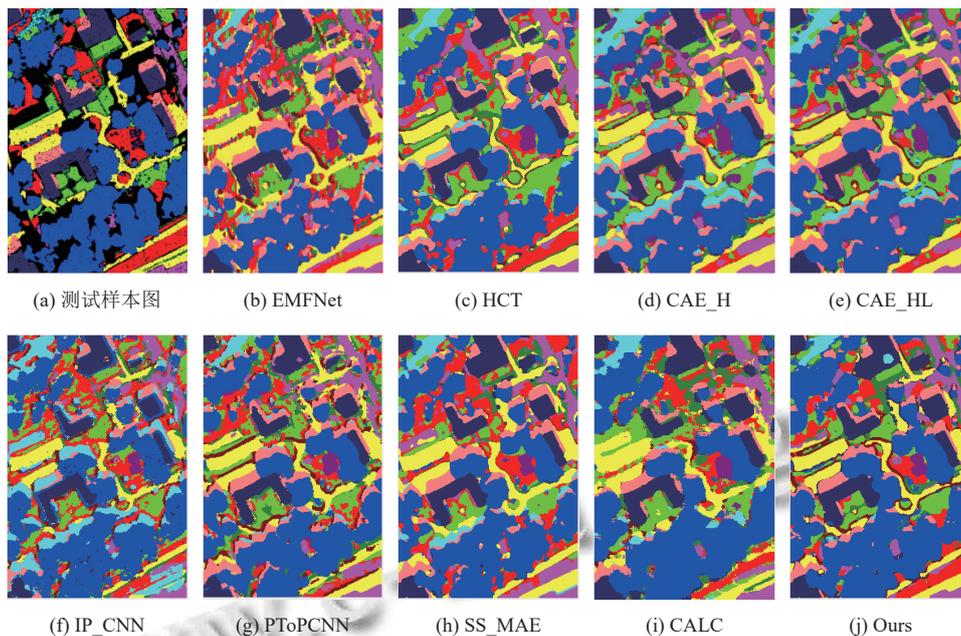


图 11 MUUFL 数据集上各模型的分类结果

为了更全面地评估我们的模型,我们在 Houston 数据集上特别关注了模型在训练过程中所需的参数量。如表 5 所示,与那些依赖于数据重构的无监督模型相

比,我们的模型所需参数量大幅度减少,这一优势在对比监督模型和半监督模型时同样显著。这一显著的优势,凸显了我们模型设计的精巧性和高效性。

表 5 Houston 数据集训练时不同模型的参数量

指标	EMFNet	HCT	CAE_H	CAE_HL	IP_CNN	PToPCNN	SS_MAE	CALC	Ours
Params	198.7k	429.83k	256.87k	300.45k	4.5M	21.16M	4.72M	336.26k	103.79k

### 2.5 训练收敛曲线

为了更全面地评估提出模型的有效性,我们对模型的训练过程进行了详细的分析,通过监控损失收敛曲线和分类精度收敛曲线来评估模型的收敛性及泛化能力。

图 12 展示了 Houston 数据集训练过程的损失收敛曲线,可以观察到,损失在前 25 轮快速下降,随后趋于平稳。图 13 展示了 Houston 数据集的分类精度的收敛曲线,可以观察到,随着训练轮数的增加,训练精度逐渐提升,并在训练后期趋于稳定,表明模型已达到其泛化能力极限。综合来看,两个收敛曲线都在后期表现趋于稳定收敛,验证了模型有效性。

### 2.6 消融实验

本文所提出的方法是设计了一种跨模态通道权重调整的特征提取网络模块,同时利用一致性正则化和伪标签结合的半监督模型。为了测试每个模块的效果,我们在本节中进行了消融实验。

首先,为了构建本文的基线模型,我们采用了少量的标记样本进行监督训练。具体来说,我们选用了 3 层卷积神经网络 (CNN) 作为模型架构,但并未引入跨模

态通道权重调整模块。这一基线模型被命名为 **Baseline**,旨在为后续的实验和性能比较提供一个基准点。

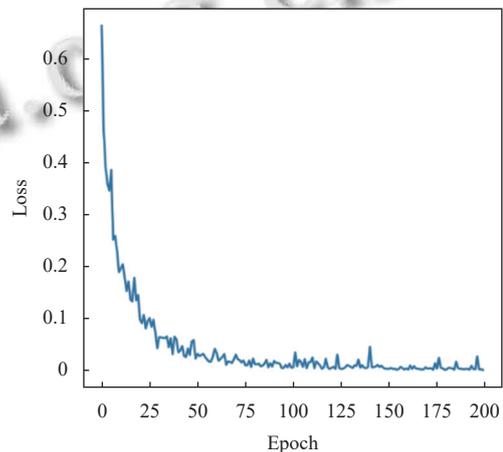


图 12 Houston 数据集上训练的损失收敛曲线

其次,我们仍然采用少量的标记样本进行监督训练,但在 **Baseline** 的基础上引入了跨模态通道权重调整模块,该模型命名为 **BasePre**。

最后,为了充分利用未标记样本和少量标记样本,在训练过程中,我们结合一致性正则化和伪标签方法

进行半监督训练,同时利用标记样本进行监督训练.此外,为处理不同模态数据互补信息,同样引入跨模态通道权重调整模块,该模型命名为 Ours.

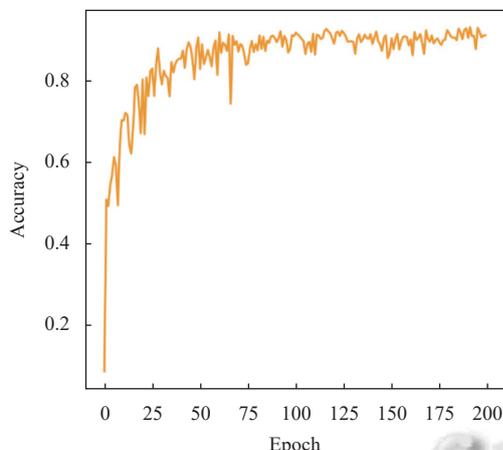


图 13 Houston 数据集上分类精度收敛曲线

图 14 展示了在 Houston 和 MUUFL 数据集上 3 个模型的 OA 值对比结果.与 Baseline 相比, BasePre 的分类精度有很大提升,在 Houston 数据集上提升 1.35%,在 MUUFL 数据集上提升了 1.22%.这表明,本文提出的跨模态通道权重调整模块,可以充分利用两个模态数据通道的相似性,以此实现提升分类精度的目的.继续观察图 14,当使用 Ours 时,在两个数据集上,网络总分类精度均达峰值,较 BasePre 分别提升 3.06% 和 1.86%.这充分说明了利用无标记样本实现一致性正则化和伪标签相结合进行半监督的方法是非常有效的,对本文的分类精度提升做了很大的贡献.

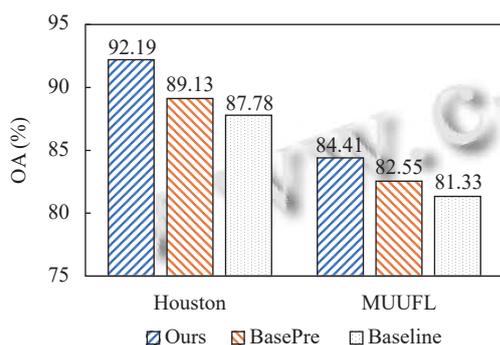


图 14 在 Houston 和 MUUFL 数据集上的消融分析

### 3 结论

监督学习广泛应用于高光谱数据与激光雷达数据联合分类领域,但是其参数优化常需大量标记样本,且它们未能有效的利用多模态数据的信息互补性.为了解决该问题,本文提出了一种跨模态通道权重调整的

半监督双分支分类网络,通过注意力机制,深入剖析高光谱和激光雷达数据通道之间的相似度,并据此自适应地调整各通道的权重,同时利用一致性正则化和伪标签相结合的方法有效地利用了未标记样本的信息.为验证方法有效性,本文在两个常用数据集上比较多个相关模型,并分析了超参数对实验结果的影响.未来我们会致力于优化网络结构,进一步发展半监督模型.

### 参考文献

- Chen C, Yan JN, Wang LZ, *et al.* Classification of urban functional areas from remote sensing images and time-series user behavior data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 1207–1221. [doi: [10.1109/JSTARS.2020.3044250](https://doi.org/10.1109/JSTARS.2020.3044250)]
- Zhang B, Sun X, Gao LR, *et al.* Endmember extraction of hyperspectral remote sensing images based on the ant colony optimization (ACO) algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, 49(7): 2635–2646. [doi: [10.1109/TGRS.2011.2108305](https://doi.org/10.1109/TGRS.2011.2108305)]
- Su HJ, Yao WJ, Wu ZY, *et al.* Kernel low-rank representation with elastic net for China coastal wetland land cover classification using GF-5 hyperspectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 171: 238–252. [doi: [10.1016/j.isprsjprs.2020.11.018](https://doi.org/10.1016/j.isprsjprs.2020.11.018)]
- Yao J, Meng DY, Zhao Q, *et al.* Nonconvex-sparsity and nonlocal-smoothness-based blind hyperspectral unmixing. *IEEE Transactions on Image Processing*, 2019, 28(6): 2991–3006. [doi: [10.1109/TIP.2019.2893068](https://doi.org/10.1109/TIP.2019.2893068)]
- Zhao XD, Tao R, Li W, *et al.* Joint classification of hyperspectral and LiDAR data using hierarchical random walk and deep CNN architecture. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(10): 7355–7370. [doi: [10.1109/TGRS.2020.2982064](https://doi.org/10.1109/TGRS.2020.2982064)]
- Geng J, Deng XY, Ma XR, *et al.* Transfer learning for SAR image classification via deep joint distribution adaptation networks. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(8): 5377–5392. [doi: [10.1109/TGRS.2020.2964679](https://doi.org/10.1109/TGRS.2020.2964679)]
- Vali A, Comai S, Matteucci M. Deep learning for land use and land cover classification based on hyperspectral and multispectral Earth observation data: A review. *Remote Sensing*, 2020, 12(15): 2495. [doi: [10.3390/rs12152495](https://doi.org/10.3390/rs12152495)]
- Li Y, Zhang HK, Shen Q. Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sensing*, 2017, 9(1): 67. [doi: [10.3390/rs9010067](https://doi.org/10.3390/rs9010067)]
- 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. *遥感学报*, 2019, 23(5): 892–903. [doi: [10.11834/jrs.20197512](https://doi.org/10.11834/jrs.20197512)]
- Jahan F, Zhou J, Awrangjeb M, *et al.* Inverse coefficient of

- variation feature and multilevel fusion technique for hyperspectral and LiDAR data classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13: 367–381. [doi: [10.1109/JSTARS.2019.2962659](https://doi.org/10.1109/JSTARS.2019.2962659)]
- 11 Dalponte M, Bruzzone L, Gianelle D. Fusion of hyperspectral and LIDAR remote sensing data for classification of complex forest areas. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46(5): 1416–1427. [doi: [10.1109/TGRS.2008.916480](https://doi.org/10.1109/TGRS.2008.916480)]
- 12 夏梦, 曹国, 汪光亚, 等. 结合深度学习与条件随机场的遥感图像分类. *中国图象图形学报*, 2017, 22(9): 1289–1301. [doi: [10.11834/jig.170122](https://doi.org/10.11834/jig.170122)]
- 13 Ni L, Gao LR, Li SS, *et al.* Edge-constrained Markov random field classification by integrating hyperspectral image with LiDAR data over urban areas. *Journal of Applied Remote Sensing*, 2014, 8(1): 085089. [doi: [10.1117/1.JRS.8.085089](https://doi.org/10.1117/1.JRS.8.085089)]
- 14 赵伍迪, 李山山, 李安, 等. 结合深度学习的高光谱与多源遥感数据融合分类. *遥感学报*, 2021, 25(7): 1489–1502. [doi: [10.11834/jrs.20219117](https://doi.org/10.11834/jrs.20219117)]
- 15 Wang XH, Feng YN, Song RX, *et al.* Multi-attentive hierarchical dense fusion net for fusion classification of hyperspectral and LiDAR data. *Information Fusion*, 2022, 82: 1–18. [doi: [10.1016/j.inffus.2021.12.008](https://doi.org/10.1016/j.inffus.2021.12.008)]
- 16 Xu XD, Li W, Ran Q, *et al.* Multisource remote sensing data classification based on convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(2): 937–949. [doi: [10.1109/TGRS.2017.2756851](https://doi.org/10.1109/TGRS.2017.2756851)]
- 17 Hang RL, Li Z, Ghamisi P, *et al.* Classification of hyperspectral and LiDAR data using coupled CNNs. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(7): 4939–4950. [doi: [10.1109/TGRS.2020.2969024](https://doi.org/10.1109/TGRS.2020.2969024)]
- 18 Li CX, Hang RL, Rasti B. EMFNet: Enhanced multisource fusion network for land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 4381–4389. [doi: [10.1109/JSTARS.2021.3073719](https://doi.org/10.1109/JSTARS.2021.3073719)]
- 19 刘伟权, 王程, 臧戎, 等. 基于遥感大数据的信息提取技术综述. *大数据*, 2022, 8(2): 28–57. [doi: [10.11959/j.issn.2096-0271.2022014](https://doi.org/10.11959/j.issn.2096-0271.2022014)]
- 20 Chen YS, Lin ZH, Zhao X, *et al.* Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, 7(6): 2094–2107. [doi: [10.1109/JSTARS.2014.2329330](https://doi.org/10.1109/JSTARS.2014.2329330)]
- 21 Hong DF, Gao LR, Hang RL, *et al.* Deep encoder-decoder networks for classification of hyperspectral and LiDAR data. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 5500205. [doi: [10.1109/LGRS.2020.3017414](https://doi.org/10.1109/LGRS.2020.3017414)]
- 22 Zhang MM, Li W, Du Q, *et al.* Feature extraction for classification of hyperspectral and LiDAR data using patch-to-patch CNN. *IEEE Transactions on Cybernetics*, 2020, 50(1): 100–111. [doi: [10.1109/TCYB.2018.2864670](https://doi.org/10.1109/TCYB.2018.2864670)]
- 23 杭仁龙, 孙瑜, 刘青山. 模态间匹配学习的高光谱和激光雷达联合分类. *遥感学报*, 2024, 28(1): 154–167.
- 24 Lu T, Ding KX, Fu W, *et al.* Coupled adversarial learning for fusion classification of hyperspectral and LiDAR data. *Information Fusion*, 2023, 93: 118–131. [doi: [10.1016/j.inffus.2022.12.020](https://doi.org/10.1016/j.inffus.2022.12.020)]
- 25 Abbasi AN, He MY. Convolutional neural network with PCA and batch normalization for hyperspectral image classification. *Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium*. Yokohama: IEEE, 2019. 959–962. doi: [10.1109/IGARSS.2019.8899329](https://doi.org/10.1109/IGARSS.2019.8899329).
- 26 Chen YS, Li CY, Ghamisi P, *et al.* Deep fusion of remote sensing data for accurate classification. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(8): 1253–1257. [doi: [10.1109/LGRS.2017.2704625](https://doi.org/10.1109/LGRS.2017.2704625)]
- 27 Bachman P, Alsharif O, Precup D. Learning with pseudo-ensembles. *Proceedings of the 27th International Conference on Neural Information Processing Systems*. Montreal: MIT Press, 2014. 3365–3373.
- 28 Lee DH. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. *Proceedings of the 2013 International Conference on Machine Learning, Workshop on Challenges in Representation Learning*. Atlanta: OpenReview.net, 2013. 896–901.
- 29 Debes C, Merentitis A, Heremans R, *et al.* Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, 7(6): 2405–2418. [doi: [10.1109/JSTARS.2014.2305441](https://doi.org/10.1109/JSTARS.2014.2305441)]
- 30 Gader P, Zare A, Close R, *et al.* MUUFL Gulfport hyperspectral and lidar airborne data set. Gainesville: University of Florida, 2013. 570.
- 31 Paszke A, Gross S, Massa F, *et al.* PyTorch: An imperative style, high-performance deep learning library. *Proceedings of the 33rd International Conference on Neural Information Processing Systems*. Vancouver: Curran Associates Inc., 2019. 721.
- 32 Zhang MM, Li W, Tao R, *et al.* Information fusion for classification of hyperspectral and LiDAR data using IP-CNN. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5506812. [doi: [10.1109/TGRS.2021.3093334](https://doi.org/10.1109/TGRS.2021.3093334)]
- 33 Lin JY, Gao F, Shi XC, *et al.* SS-MAE: Spatial-spectral masked autoencoder for multisource remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5531614. [doi: [10.1109/TGRS.2023.3331717](https://doi.org/10.1109/TGRS.2023.3331717)]

(校对责编: 张重毅)