

基于知识图谱语义表示的 高光谱遥感图像零样本分类

马丽^{12*}, 汤睿婕², 李彦胜³

1.海南大学生物医学工程学院, 海南三亚 572024

2.中国地质大学(武汉)机械与电子信息学院, 湖北武汉 430074

3.武汉大学遥感信息工程学院, 湖北武汉 430079

摘要: 针对高光谱遥感图像标注类别不完整条件下未知地物类别难以被传统监督学习方法有效识别的问题, 本文提出一种基于知识图谱语义表示的高光谱遥感图像零样本分类网络 (Knowledge Graph semantic representations based Zero-Shot Network for hyperspectral remote sensing image classification, KG-ZSNet). 该方法通过构建遥感领域知识图谱获取类别层面的丰富语义表示, 并在生成式零样本学习框架下引入变分自编码器, 实现视觉特征与语义特征的跨模态对齐, 进而在共享隐层空间中完成未知类别分类. 在多个公开高光谱遥感数据集上的实验结果表明, 所提出方法在零样本像素级分类任务中的分类精度优于多种典型零样本学习方法, 通过引入结构化知识图谱并结合生成式跨模态对齐机制, 在高光谱遥感图像零样本分类任务中表现出良好的有效性与稳定性.

关键词: 高光谱遥感图像; 零样本学习; 知识图谱; 变分自编码器; 未知类分类.

中图分类号: P237; TP751

收稿日期: 2026-1-14

Zero-Shot Classification of Hyperspectral Remote Sensing Images Based on Knowledge Graph Semantic Representations

Ma Li^{12*}, Tang Ruijie², Li Yansheng³

1. School of Biomedical Engineering, Hainan University, Sanya 572024, China

2. School of Mechanical Engineering and Electronic Information, China University of Geosciences, Wuhan 430074, China

3. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

Abstract: Under incomplete category annotation conditions, traditional supervised learning methods have difficulty in effectively recognizing unseen land-cover categories in hyperspectral remote sensing images. This paper proposes a Knowledge Graph semantic representations based Zero-Shot Network for hyperspectral remote sensing image classification(KG-ZSNet). We construct a remote sensing knowledge graph to obtain category-level semantic representations and integrate variational autoencoder into a generative zero-shot learning framework to achieve cross-modal alignment between visual and semantic features, thereby enabling unseen category classification in shared latent space. Experiments on multiple public hyperspectral datasets demonstrate that KG-ZSNet achieves higher accuracy than several representative zero-shot learning methods in zero-shot hyperspectral remote sensing image classification. KG-ZSNet, by incorporating structured knowledge-graph semantics and the generative cross-modal alignment mechanism, achieves strong effectiveness and stability in zero-shot hyperspectral remote sensing image classification.

Keywords: hyperspectral remote sensing images; zero-shot learning; knowledge graph; variational autoencoder; unseen class classification.

基金项目: 国家自然科学基金项目 (No.61771437) 智慧地球重点实验室开放基金 (KF2023YB04-05)

作者简介: 马丽 (1982-), 女, 教授, 博士生导师, 主要从事机器学习, 深度学习等方面的研究. E-mail: maryparisster@gmail.com. ORCID: 0000-0003-3873-5080.

*通讯作者: 马丽, E-mail: maryparisster@gmail.com.

0 引言

高光谱图像 (Hyperspectral Images, HIS) 是由搭载高光谱成像仪的航空航天飞行器捕捉到的三维立体图像, 图像中的每个像素均含有上百个不同波段的光谱反射信息。在地物分类、环境监测和资源调查等领域展现出显著优势 (陈伟涛等, 2025; 陈志行等, 2025)。其中, 高光谱遥感图像地物分类一直是高光谱遥感领域的研究重点和研究热点之一, 它利用高光谱传感器获取的多波段数据, 将图像中的每个像素点分配到不同的地物类别中。近年来, 随着成像技术的发展和数据获取能力的提升, 高光谱遥感数据规模不断扩大, 推动了深度学习模型在该领域的广泛应用 (Kumar et al., 2024)。这类方法在地物充分标注的条件下, 通常能取得较高的分类精度。然而, 在实际应用中, 遥感图像包含地物类型复杂, 人工调查范围有限, 难以保证对所有地物进行标注, 图像中可能存在未被标注的地物类别。在此背景下, 传统依赖完整标注信息的监督学习方法难以有效识别数据集中的未标注类别, 未标注类别被错误分类为已标注类别, 影响图像分类精度。

为缓解类别标注不完整问题, 零样本学习 (Zero-shot Learning, ZSL) 通过引入类别语义等辅助信息, 在训练阶段仅利用已知类别的标注样本, 实现对未标注目标类别, 也就是未知类别的识别。该思想在自然图像与计算机视觉领域已得到广泛研究。从技术路线来看, 现有零样本学习方法大致可分为嵌入式方法和生成式方法两类 (Xie et al., 2023)。嵌入式方法通常通过学习视觉特征与语义特征之间的映射关系, 将二者投射到共享嵌入空间中, 并基于相似性度量完成类别判别。该类方法的分类性能在一定程度上依赖于语义表示的表达能力以及跨模态嵌入空间的映射效果, 当语义信息区分性不足或类间分布较为接近时, 其对未知类别的泛化性能可能受到影响 (Wang et al., 2019)。生成式零样本学习方法从特征生成的角度出发, 在训练阶段学习已知类别视觉特征与语义特征之间的跨模态关联, 并利用未知类别的语义特征在共享隐层空间中生成对应的类别原型, 从而将零样本学习问题转化为可采用监督分类策略处理的问题。相较于嵌入式方法, 生成式方法通过在特征空间中学习类别条件分布, 为未知类别分类提供了可用的特征表示基础, 在缓解已知类别与未知类别样本缺失带来的不平衡问题方面具有一定优势 (Pourpanah et al., 2023), 因此近年来在零样本分类任务中受到广泛关注。其中,

基于生成对抗网络 (Generative Adversarial Network, GAN) (Goodfellow et al., 2014) 和变分自编码器 (Variational Autoencoder, VAE) (Kingma and Welling, 2022) 的方法被证明在跨模态特征建模和分布对齐方面具有良好的灵活性与稳定性。GAN 通过生成器与判别器之间的对抗训练机制, 在刻画复杂数据分布和生成具有判别性的特征表示方面展现出较强的建模能力, 已被成功应用于多种零样本和跨模态学习任务中。与此同时, 也有研究指出 (Tan et al., 2024), GAN 类方法的训练过程依赖于对抗博弈的平衡状态, 其收敛行为和生成质量在一定程度上受到网络结构设计、训练策略以及超参数设置的影响。VAE 通过最大化变分下界对隐层变量进行概率建模, 为生成过程提供了明确的优化目标和稳定的训练范式。其连续隐层空间能够在引入类别层面语义特征的前提下, 对不同类别的特征分布进行平滑表达, 从而有助于提升合成特征的多样性与整体覆盖性 (Mishra et al., 2018)。

近年来, 零样本学习逐渐被引入遥感图像分析任务中, 但现有成果多集中于高分辨率遥感场景分类 (Li et al., 2021; Liu et al., 2023) 或高分辨率遥感土地利用制图 (Li et al., 2025), 真正面向高光谱像素级分类的工作仍相对稀少。现有研究主要关注跨数据集高光谱分类问题, 在训练集与测试集类别不重叠的条件下引入语义表示并对视觉与语义信息进行对齐, 以实现对未知类别的识别。Pan et al. (2021) 较早将零样本学习引入高光谱图像像素级分类任务, 在类别不重叠的跨数据集设置下, 采用 Word2Vec 词向量作为类别语义表示, 并通过嵌入式零样本方法实现未知类别的识别, 为高光谱零样本学习奠定了基础。随后, 有研究开始在具体应用场景中借鉴零样本学习思想, Freitas et al. (2022) 在海洋垃圾检测场景中, 通过构建潜在特征空间并结合距离判别与阈值策略, 实现了对未知类别的区分, 该方法更侧重于未知类别发现, 而非依赖显式类别语义信息。总体来看, 零样本学习在高光谱遥感中的研究仍处于起步阶段, 现有工作主要集中于嵌入式零样本框架, 尚缺乏对生成式零样本学习方法的系统探索。

零样本学习的核心在于借助类别层面的语义表示建立视觉特征空间与语义空间之间的关联, 从而实现对未知类别的识别 (Chen et al., 2023)。现有零样本学习中的语义信息通常来源于类别描述, Akata et al. (2013) 通过专家经验为类别构建人工属性向量, 具有一定的语义可解释性, 但标注成本高、扩

展性有限;之后 Akata *et al.* (2015) 又引入 Word2Vec 通用词向量作为类别语义嵌入,而 Jia *et al.* (2021) 进一步采用 BERT 对类别文本进行编码,以降低对人工属性的依赖。高光谱像素级零样本分类中语义信息的选取仍以上述人工属性向量或通用文本词向量为主。此类语义表示获取相对便捷,但其主要来源于自然语言语料或人工经验,难以显示刻画高光谱遥感图像不同地物类别之间的丰富关系。知识图谱作为一种对知识进行结构化存储与表达的手段,能够描述客观世界的事物与关系。其节点表示客观世界的事物,其边代表事物之间的关系,最终形成计算机能够理解并加以计算的知识结构体系(Fensel *et al.*, 2020)。对于高光谱遥感图像,知识图谱不仅能够描述地物类别属性,还能够表达类别之间丰富的语义关系、层级关系与空间关系等。例如对“木桥”这种未知类,知识图谱不仅能够体现其与“石砖桥”、“沥青桥”这两种已知类具有较密切的语义关系,且三种桥都与“水”有空间邻接关系,同时,知识图谱还能够体现三种桥的属性差异(尺度和光谱属性)。因此,基于知识图谱提取的地物语义信息能够更全面地表征地物特性,进而建立更准确的语义-视觉映射关系,以实现未知类视觉信息的推理,提升未知类的零样本分类精度。

将知识图谱应用于深度学习网络通常包含两个步骤:知识图谱构建和知识表示学习。前者用于确定实体集合及其关系类型,从而刻画类别之间的语义关联结构;后者则依据图结构将实体与关系映射到连续向量空间,生成可被下游模型直接利用的语义向量表示,相关方法包括 TransE、RotatE、DistMult 与 ComplEx 等(Bordes *et al.*, 2013; Yang *et al.*, 2014; Trouillon *et al.*, 2016; Sun *et al.*, 2019)。在构建知识图谱的研究中,国外比较有代表性的通用领域知识图谱有 Freebase (Bollacker *et al.*, 2008)、Dbpedia (Lehmann *et al.*, 2015) 等。国内学者借鉴了国外百科知识图谱的建设经验,利用本土化的百度百科等网络资源构建并公开了 CN-Dbpedia (Xu *et al.*, 2017)、CN-Probase (Chen *et al.*, 2019) 等中文百科知识图谱,从而对中文环境下的知识进行了抽取和整理,以实现知识图谱化。由于上述提到的通用知识图谱相对更加注重知识覆盖的广度,难以针对具体领域设计合理的本体来进行领域知识表达与实体关系约束,遥感领域的学者们为了应对遥感领域内知识的特性与难点,构建了明显的遥感领域倾向性的遥感知识图谱(蒋秉川等, 2018; 张雪英等,

2020)。之后就有学者将知识图谱运用到遥感图像场景分类领域中。李彦胜等(2024)将知识图谱嵌入到遥感场景图片级分类模型中,实验结果表明从知识图谱获得的语义向量可以显著提升分类精度。Li *et al.* (2021) 则将知识图谱引入零样本遥感场景分类任务,对遥感场景中的未知类进行分类。然而,知识图谱在高光谱图像像素级零样本分类中的应用尚未展开研究。与遥感场景分类不同,高光谱图像分类属于像素级分类任务,其基本识别单元不是整幅图像,而是以像素为中心的局部邻域块。对于遥感场景分类,图像通常包含较丰富的形状、纹理、颜色以及地物之间的空间布局信息,知识图谱中类别属性和类别关系与图像内容之间具有较为直接的对应关系。相比之下,高光谱像素级分类所能利用的空间关系信息相对有限,而地物在连续光谱波段上的响应特征对类别判别更为关键。例如,植被在近红外波段通常具有较高反射率,而水体在近红外区域的反射率明显较低,这类光谱属性能够为地物识别提供重要依据。因此,已有面向场景分类任务构建的知识图谱难以直接适用于高光谱图像像素级零样本分类任务,需要结合高光谱数据特点,对知识图谱中的类别属性和类别关系进行针对性设计。

为探索知识图谱在高光谱图像零样本分类中的应用,本文提出一种基于知识图谱语义表示的高光谱遥感图像零样本分类网络(Knowledge Graph semantic representations based Zero-Shot Network for hyperspectral remote sensing image classification, KG-ZSNet),其在生成式零样本学习框架下,通过变分自编码器建立高光谱视觉特征与知识图谱语义特征之间的跨模态关联,并在此基础上合成未知类别的视觉特征表示,从而将零样本识别问题转化为可监督的分类任务。总体网络结构包括视觉特征提取模块、语义特征提取模块、视觉-语义跨模态对齐模块和未知类分类模块。训练过程包含三个步骤:首先训练二维卷积神经网络进行高光谱数据的视觉特征提取,构建高光谱城市场景的知识图谱,并采用 DistMult 方法进行各类地物的语义特征提取;然后训练两个 VAE 网络将已知类的视觉和语义特征进行跨模块重构和对齐,获得视觉和语义的对齐隐层特征;最后利用未知类的隐层语义特征训练未知类语义分类器。在测试阶段,对未知类样本,通过视觉特征提取模块和 VAE 对齐模块,提取其隐层视觉特征,由于两种模态在共享的隐层空间中具有一致的分布结构,因此利用未知类语义分类器,能够实

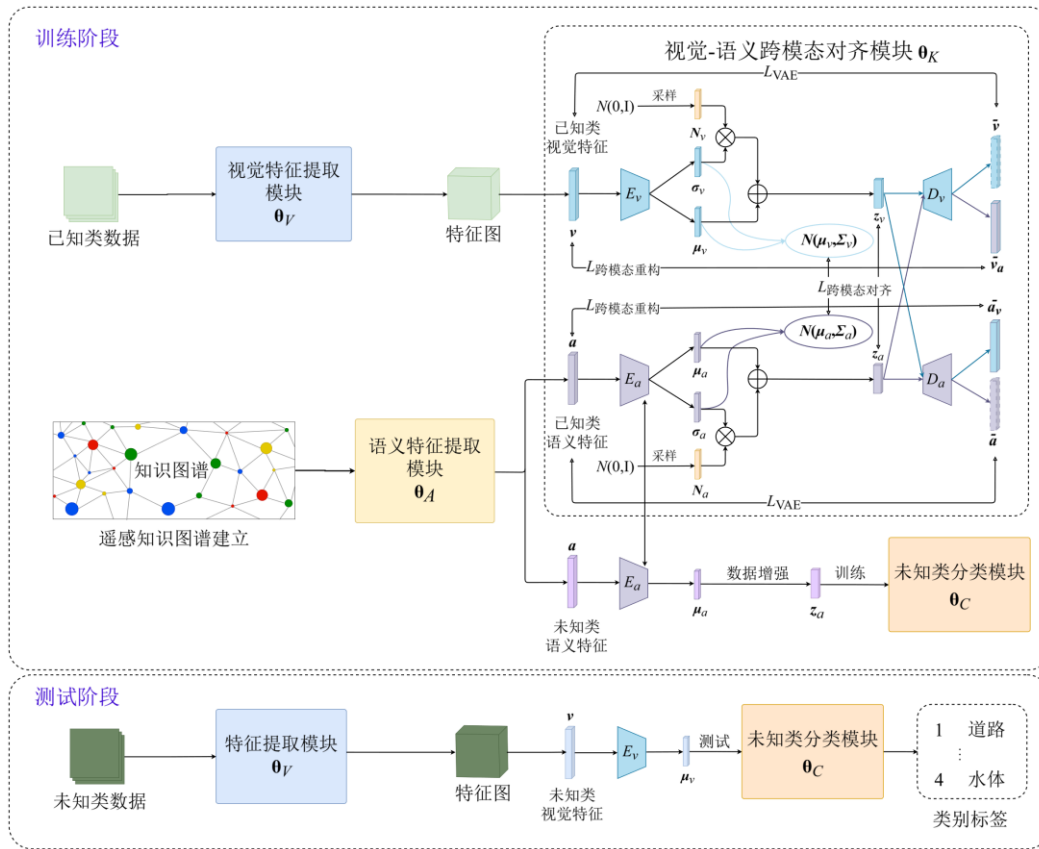


图 1 KG-ZSNet 模型结构
Fig.1 Architecture of KG-ZSNet

现未知类隐层视觉特征的零样本分类。

本文的主要创新点概括如下：

(1) 现有高光谱遥感图像像素级零样本分类研究主要集中于嵌入式方法，使用词向量进行语义表达，对生成式零样本学习框架和知识图谱的探索仍较为有限。本文首次在高光谱零样本分类任务中引入生成式零样本学习框架，利用知识图谱对地物类别的属性信息及类别间关系进行结构化表达，在仅存在已知类别标注样本条件下，提取基于知识图谱的地物语义信息，实现对未知地物类别的有效识别。

(2) 现有基于知识图谱的零样本分类研究主要面向自然图像分类或遥感场景分类，由于高光谱遥感图像分类是像素级分类任务，因此，对高光谱图像分类，有必要对知识图谱进行针对性设计。本文面向城市高光谱数据集中的地物类别构建知识图谱，全面描述地物类别属性及类别间的语义关系、层级关系和空间关系，并着重设计了在遥感图像分类中具有关键作用的光谱反射率属性。通过将知识图谱中蕴含的类别属性信息及类别间关系信息引入零样本学习过程，为模型提供了相比人工属性向量和通用词向量更加丰富且结构化的语义信息，从而提升高光谱零样本分类性能。

1 方法

本文研究的是高光谱遥感图像在零样本学习设定下的像素级分类问题。训练阶段仅使用已知地物类别的标注样本，测试阶段的数据为训练阶段未出现的未知地物类别。在该设定下，传统监督分类方法由于缺乏未知类别的标注样本，难以对未知类样本进行有效识别。本文在训练阶段利用已知类别的高光谱视觉特征及其对应的语义信息进行学习，并借助未知类别的语义表示，在测试阶段实现对未知地物类别视觉特征的识别。

1.1 KG-ZSNet 模型

KG-ZSNet 整体模型结构如图 1 所示，分为训练阶段和测试阶段，训练阶段的网络由以下四个功能模块构成：(1) 视觉特征提取模块 (θ_V)：使用二维卷积神经网络 (2D-CNN) 对高光谱图像进行特征提取，获得类别对应的视觉特征表示。需要注意的是，视觉特征提取器仅使用已知类别样本进行训练，未知类别样本在训练阶段不参与任何参数更新。(2) 语义特征提取模块 (θ_A)：基于构建的知识图谱，采用 DistMult 知识表示学习方法对类别节点进行嵌入建模，获得包含类别属性与关系信息的低维稠密语义特征表示。(3) 视觉-语义跨模态对齐模块 (θ_K)：

通过双重变分自编码器分别对视觉特征与语义特征进行编码与解码，并在共享的隐层空间中实现二者的分布一致性与特征关联。该模块在保持各模态特征重构一致性的同时，使已知类别的视觉特征与语义特征在隐层空间中对齐，从而为未知类别相关的跨模态视觉和语义特征对齐提供支持。（4）未知分类模块（ θ_c ）：利用线性层与对数 Softmax 组成的分类器执行零样本分类任务。该模块在训练阶段以未知类别的隐层语义向量作为类别语义原型，通过在隐层空间中重复并引入随机扰动进行数据特征增强，生成对应的未知类隐层语义表示用于未知类分类器训练。在测试阶段，分类器对未知类的隐层视觉向量进行识别，实现零样本分类。

KG-ZSNet 的总体优化目标函数定义为:

$$L = L_{\text{CF}} + L_{\text{KG}} + L_{\text{JUNK}}, \quad (1)$$

其中, L_{CE} 表示交叉熵损失, 用于在已知类标签监督下学习判别性视觉特征; L_{KG} 表示视觉-语义跨模态对齐模块损失, 用于实现语义特征与视觉特征的对齐学习; L_{UNK} 表示未知类零样本分类器损失, 采用负对数似然损失进行建模, 用于约束数据增强后的未知类隐层语义向量对未知类分类器的训练。在模型测试阶段, 仅使用经过特征提取和 VAE 编码器的未知类视觉特征进行预测, 从而实现高光谱遥感数据中未知地物类别的零样本识别。

1.2 高光谱遥感知知识图谱的建立

为刻画高光谱遥感数据集中地物类别之间的语义属性及其相互关系,本文首先构建了与具体城市数据集对应的高光谱遥感知识图谱,为后续语义特征提取与跨模态对齐提供结构化语义信息基础。知识图谱是一种以图结构形式组织和表示知识的语义模型,通常表示为 $G=(E,R)$,其中 E 表示所有实体的集合, R 表示所有关系的集合。知识图谱以事实三元组作为基本表达单元,其形式为 $T=\{<h,r,t>|h,t\in E,r\in R\}$,其中 h 和 t 分别代表头实体和尾实体, r 代表头实体和尾实体之间的关系。通过大量共享实体的三元组相互连接,知识图谱构建起结构化的语义网络,从而实现对客观事实及其内在关系的系统化表达。

我们选择根据现有的高光谱图像数据，有针对性地构建小型城市知识图谱，类似的也可以建立农业知识图谱等其他小型知识图谱。在实体选择方面，本文选取高光谱遥感数据中涉及的地物类别及其属性值作为实体，以提高知识图谱对高光谱遥感场景

的针对性。其中,地物类别包括道路、裸土、树木等,在事实三元组中既可以作为头实体,也可以作为尾实体;属性值包括颜色值、光谱反射率范围值等,主要作为事实三元组中的尾实体。在关系构造方面,我们根据高光谱地物知识中光谱信息、语义信息和空间信息的不同表达需求,将知识图谱中的关系划分为三类:I类关系为光谱相关关系,主要描述地物局部光谱特征块中和光谱相关的属性信息;II类关系为语义相关关系,用于刻画地物类别之间的层级结构、语义相似性等全局先验属性;III类关系为空间相关关系,用于描述地物在空间分布上的特性。这种知识组织方式既能够刻画未知类别自身的属性特征,也能够反映其与其他类别之间的关联关系及其在整体语义结构中的位置,从而为生成式零样本学习框架中视觉-语义跨模态对齐提供更具判别性的语义支撑。基于上述构建原则,我们选取了43个地物类别和27个属性值作为实体,并设计了12种关系用于描述这些实体之间的联系,形成适用于零样本分类任务的高光谱遥感小型知识图谱。此知识图谱总共构建了380条事实三元组,部分事实三元组示例如表1所示,展示了构建的知识图谱中地物类别属性、类别间语义关系以及空间关系的具体三元组表达形式。构建的部分知识图谱如图2所示,其中圆形表示实体,箭头表示关系。

表 1 部分事实三元组示例
Table 1 Example of partial triples

三元组类型	遥感地物实体	遥感知识图谱三元组
光谱相关关系	裸土	<裸土, 颜色为, 棕色>
	树木	<树木, 可见光谱反射率, 偏低>
语义相关关系	植被	<植被, 拥有二级类, 草地>
	木桥	<木桥, 相似于, 石砖桥>
空间相关关系	体育场座位	<体育场座位, 围绕, 跑道>
	阴影	<阴影, 相接于, 建筑>
	桥梁	<桥梁, 穿过, 水体>
	车辆	<车辆, 停在, 停车场>

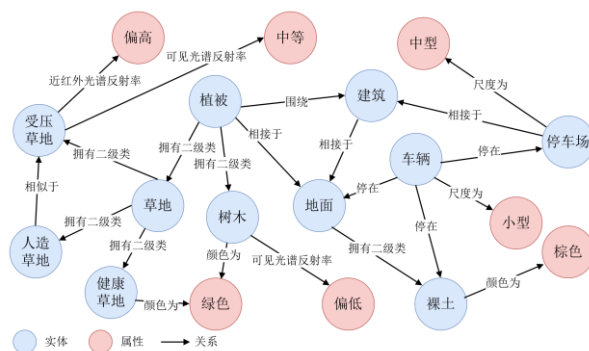


图 2 部分知识图谱
Fig.2 Part of a knowledge graph

1.3 知识图谱的语义特征提取

为使知识图谱中的类别语义信息能够被深度学习模型有效利用，本文引入语义特征提取模块，将知识图谱中的实体及其关系转化为连续的语义向量表示，为后续与视觉特征的对齐学习提供输入。知识图谱表示学习方法的基本思想是根据知识图谱结构，将实体与关系映射到连续向量空间，从而通过语义向量的形式便于下游任务模型对知识图谱内所保存的知识进行处理与运算。

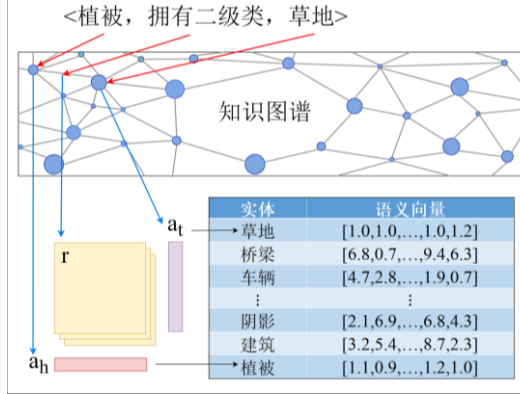


图3 基于 DistMult 方法的实体语义表示
Fig.3 Semantic representation of entities using DistMult method.

考虑到本文构建的知识图谱规模相对较小，语义匹配模型 DistMult 具有结构简洁、参数规模较低、实现相对直接等特点，能够刻画实体间语义关联性以及建模知识图谱中的多对多关系（李彦胜等，2024），满足本文对类别语义向量构建的需求。因此，本文选择 DistMult 模型来从知识图谱中获得实体的知识图谱语义向量。如图3所示，对于知识图谱中的三元组 $\langle h, r, t \rangle$ ，DistMult 模型将实体映射为语义向量 $a_h \in \mathbb{R}^d$ ， $a_t \in \mathbb{R}^d$ ，并将关系映射为双线性变换矩阵 $a_r \in \mathbb{R}^{d \times d}$ 。其对应的三元组评分函数表示如下：

$$f(h, r, t) = a_h^T \text{diag}(a_r) a_t, \quad (2)$$

其中， $\text{diag}(a_r)$ 代表关系 r 对应的变换矩阵。三元组得分的取值范围为 0 到 1。为训练 DistMult 模型，需要构造负三元组集合 $T = \{\langle h', r, t' \rangle | \langle h', r, t' \rangle \notin G\}$ ，其方法通常是通过随机替换正三元组集合 $T = \{\langle h, r, t' \rangle | \langle h, r, t' \rangle \in G\}$ 中的头实体或尾实体来生成。知识图谱中存在的正三元组应当具有较高的得分，而不在知识图谱中的负三元组则应当具有较低的得分。用于训练 DistMult 模型的损失函数表示如下：

$$L = \sum_{\langle h, r, t \rangle \in T} \sum_{\langle h', r, t' \rangle \in T'} \max(0, f(h, r, t) + \tau f(h', r, t')), \quad (3)$$

其中， τ 为设定的边界阈值。

1.4 视觉-语义的跨模态对齐

在获得高光谱视觉特征与类别语义特征之后，本文设计了视觉-语义跨模态对齐模块，用于在共享隐层空间中学习视觉特征与语义特征之间的对齐表示，从而为未知类别的语义-视觉关联提供支持。如图1所示，视觉-语义跨模态对齐模块由两个变分自编码器（VAE）构成，用于同时执行特征重构与视觉特征和语义特征的跨模态对齐。具体而言，两个 VAE 分别通过编码器 E_v 和 E_a 对输入的视觉特征 v 和语义特征 a 进行重构。在隐层空间中，视觉与语义特征可表示为 μ_v, σ_v 与 μ_a, σ_a ，其中 μ 表示均值向量， Σ 表示标准差向量。通过缩小视觉与语义特征在隐层空间中的距离实现跨模态对齐，其损失函数定义为：

$$L_{\text{跨模态对齐}} = \sum_{i=1}^N \sqrt{\|\mu_v^{(i)} - \mu_a^{(i)}\|_2^2 + \|\sigma_v^{(i)} - \sigma_a^{(i)}\|_2^2}, \quad (4)$$

其中， N 表示训练样本数， $\|\cdot\|_2$ 表示 L2 范数。通过这种跨模态特征对齐机制，知识图谱中蕴含的结构化语义信息得以有效迁移到视觉特征学习过程中。

在隐层向量生成过程中，标准差 σ_v 与从多维标准正态分布 $N(0, I)$ 中采样的随机变量相乘，并加至均值 μ_v 上，以获得隐层向量 z_v 。语义特征亦采用相同方式得到隐层向量 z_a 。该隐层向量捕获了输入数据的隐层特征，表示其低维编码。随后，隐层向量 z_v 和 z_a 分别输入解码器 D_v 和 D_a ，得到重构结果 $\bar{a}$ 和 $\bar{v}$ 。两个 VAE 的重构损失定义为：

$$L_{\text{VAE}} = \text{MAE}(v, \bar{v}) + \gamma D_{\text{KL}}(N(\mu_v, \Sigma_v) \| N(0, I)) + \text{MAE}(a, \bar{a}) + \gamma D_{\text{KL}}(N(\mu_a, \Sigma_a) \| N(0, I)), \quad (5)$$

其中， D_{KL} 表示 KL 散度计算函数， $N(0, I)$ 为隐层向量 z_v 与 z_a 的先验分布， $N(\mu_v, \Sigma_v)$ 和 $N(\mu_a, \Sigma_a)$ 分别为 z_v 与 z_a 的后验分布， I 表示与隐层向量维度一致的单位矩阵， Σ_v 和 Σ_a 表示协方差矩阵，采用对角形式 $\Sigma_v = \text{diag}(\sigma_v^2)$ 和 $\Sigma_a = \text{diag}(\sigma_a^2)$ ， γ 为平衡系数。通过 KL 散度项，后验分布被约束为接近标准正态分布。该损失可以约束两个 VAE 学习对应特征向量的重构模型，并获得具有较强表达力的特征隐层空间向量。

为进一步提升对齐性能，更好地重构出视觉特征与对应语义向量在隐层空间的相似分布，引入跨模态特征重构损失，定义如下：

$$L_{\text{跨模态重构}} = \sum_{i=1}^N |v^{(i)} - \bar{v}_a^{(i)}| + |a^{(i)} - \bar{a}_v^{(i)}|, \quad (6)$$

其中， $\bar{v}_a = D_v(z_a)$ 并且 $\bar{a}_v = D_a(z_v)$ 。若 z_v 和 z_a 对齐良好，则 $\bar{v}_a$ 将接近 $\bar{v} = D_v(z_v)$ ，可视为对视觉特征 v 的重构。

在上述损失函数的约束下，视觉-语义跨模态对齐模块对视觉模态与语义模态分别进行概率编码，并在隐层空间中实现跨模态对齐，使得同一类别的视觉特征与语义特征在分布层面保持一致性，有效缓解了跨模态语义鸿沟问题。该机制不仅有助于提升已知类别的特征表达能力，同时也为模型在测试阶段对未知类别的泛化与识别提供了稳定的语义先验支撑。最终，视觉-语义跨模态对齐模块的整体损失函数定义为：

$$L_{KG} = L_{VAE} + \alpha L_{\text{跨模态重构}} + \beta L_{\text{跨模态对齐}}, \quad (7)$$

其中， α 和 β 为平衡系数。

2 实验与分析

2.1 数据集描述

为验证所提出的基于知识图谱语义表示的高光谱遥感图像零样本分类方法的有效性，本文选取了三个具有代表性的公开高光谱数据集进行实验评估，包括 University of Pavia (PU)、Houston2012 (HU12) 以及 Houston2018 (HU18) 数据集。PU 数据集由反射光学光谱成像系统 (ROSIS) 采集，成像区域位于意大利北部帕维亚市。PU 数据集的原始图像尺寸为 610×340 像素，初始包含 115 个光谱波段。通过去除 12 个噪声波段后，最终保留 103 个有效波段。剔除无效区域后，PU 数据集的有效尺寸为 610×315 像素，其光谱覆盖范围为 430-860 nm，空间分辨率为 1.3 m。PU 数据集原始共包含 15 类地物，为降低类别间语义冗余并避免极少样本类别对模型训练的干扰，对其中的建筑类地物进行了合并，同时剔除了样本数量极少的地物类别，最终保留 12 类地物用于实验分析。HU12 数据集采集于 2012 年，由 ITRES 紧凑型机载光谱成像仪 1500 (CASI-1500) 传感器采集，成像区域覆盖美国休斯顿大学校园及其周边区域。该数据集的图像尺寸为 349×1905 像素，包含 144 个光谱波段，光谱范围为 380-1050 nm，空间分辨率为 2.5 m。该数据集中共标注了 15 类地物，涵盖自然地物与人工地物类型。HU18 数据集采集于 2018 年，由休斯顿大学高光谱图像分析组联合 NSF 资助的机载激光测绘中心 (NCALM) 获取，成像区域覆盖美国休斯顿大学校园及其周边区域。该数据集的图像尺寸为 1202×4768 像素，包含 50 个光谱波段，光谱范围为 380-1050 nm，空间分辨率为 1 m。该数据集标注了 20 个类别，适合开展复杂城市场景下的地物识别研究。图 4 展示了三个数据集

的三波段假彩色合成图像及对应的地物类别标注信息。PU 数据集和 HU18 数据集中各类别的名称及样本数量如表 2、4 所示，HU12 数据集中各类别的名称、训练样本数、测试样本数以及总样本数如表 3 所示。在实际实验中，鉴于 PU 和 HU18 数据集像素数量较大，分别随机选取 30% 和 1% 的样本用于实验。

在未知类别的设置上，本文尽量选择与已知类别之间具有一定语义关联的地物类别，以满足零样本学习中由已知类向未知类进行知识迁移的基本前提。对于 PU 数据集，选择第 4 类树木、第 7 类沥青铺面及第 11 类停车场作为未知类别。对于 HU12 数据集，选择第 2 类受压草地、第 8 类非住宅建筑、第 11 类铁路以及第 13 类未铺面停车场作为未知类别。对于 HU18 数据集，选择选取第 2 类受压草地、第 4 类常青树、第 8 类住宅建筑以及第 15 类铁路作为未知类别。

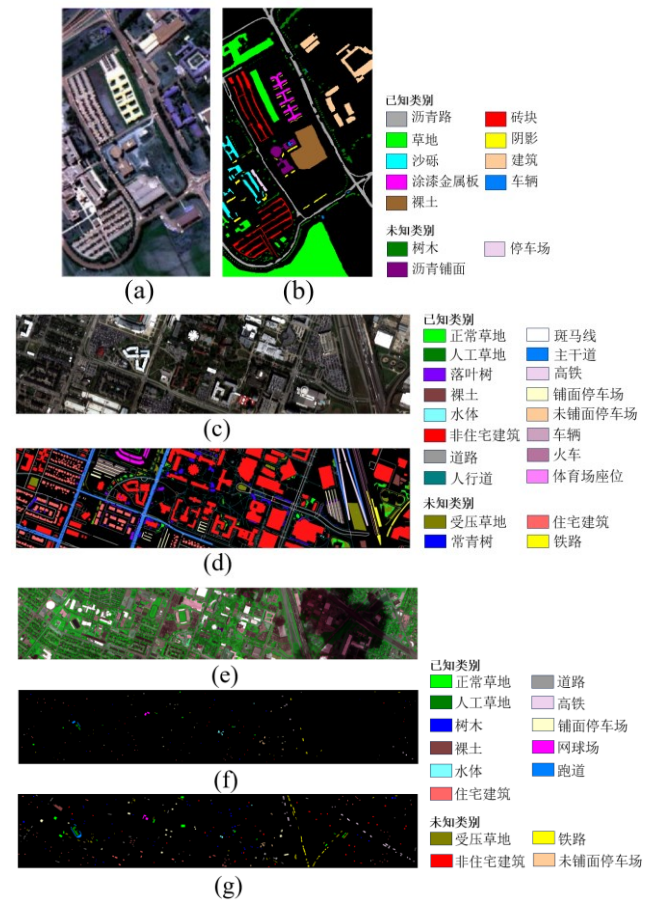


图 4 数据集示例图像。(a) PU 数据集的假彩色合成图；(b) PU 数据集的地物真实标注图；(c) HU18 数据集的假彩色合成图；(d) HU18 数据集的地物真实标注图；(e) HU12 数据集的假彩色合成图；(f) HU12 训练集的地物真实标注图；(g) HU12 测试集的地物真实标注图。

Fig.4 dataset images. (a) False-color image of PU. (b) Ground-truth map of PU. (c) False-color image of PC. (d) Ground-truth map of HU18. (e) False-color image of HU18. (f) Ground-truth map of training data in HU12. (g) Ground-truth map of testing data in HU12.

表 2 PU 数据集的类别信息
Table 2 Class information of the PU datasets

标签	类别名称	数目
已知类	1 沥青路	6631
	2 草地	18649
	3 沙砾	2099
	5 涂漆金属板	1345
	6 裸土	5029
	8 砖块	3682
	9 阴影	947
	10 建筑	4663
	12 车辆	136
未知类	4 树木	3064
	7 沥青铺面	1330
	11 停车场	340

表 3 HU12 数据集的类别信息
Table 3 Class information of the HU12 dataset

标签	类别名称	数目	训练集数目	测试集数目
已知类	1 正常草地	1251	198	1053
	3 人工草地	697	192	505
	4 树木	1244	188	1056
	5 裸土	1242	186	1056
	6 水体	325	182	143
	7 住宅建筑	1268	196	1072
	9 道路	1252	193	1059
	10 高铁	1227	191	1036
	12 铺面停车场	1233	192	1041
	14 网球场	428	181	247
	15 跑道	660	187	473
未知类	2 受压草地	1254	190	1064
	8 非住宅建筑	1244	191	1053
	11 铁路	1235	181	1054
	13 未铺面停车场	469	184	285

表 4 HU18 数据集的类别信息
Table 4 Class information of the HU18 dataset

标签	类别名称	数目
已知类	1 正常草地	39196
	3 人工草地	2736
	5 落叶树	20172
	6 裸土	18064
	7 水体	1064
	9 非住宅建筑	894769
	10 道路	183283
	11 人行道	136035
	12 斑马线	6059
	13 主干道	185438
	14 高铁	39438
	16 铺面停车场	45932
	17 未铺面停车场	587
	18 车辆	26289
	19 火车	21479
	20 体育场座位	27296
未知类	2 受压草地	130008
	4 常青树	54322
	8 住宅建筑	158995
	15 铁路	27748

2.2 实验设置

KG-ZSNet 模型实验设置如下。2D-CNN 网络仅利用已知类别的原始视觉特征进行训练，训练轮数为 200 个 epoch，并在第 100 个 epoch 时进行一次学习率衰减，所有数据集的 batch size 均设为 10；双重 VAE 网络利用已知类别的视觉特征与语义特征进行联合训练，训练轮数为 300 个 epoch，所有数据集的

batch size 统一设为 50；未知类分类器基于对齐后的未知类别语义特征进行训练，训练轮数为 50 个 epoch，所有数据集的 batch size 设为 32。所有网络均在 4 张 NVIDIA GTX 2080 Ti GPU（单卡显存 11 GB）上完成训练。为保证实验结果的稳定性与可重复性，所有实验均在 5 个随机选取的种子下重复进行，模型使用总体分类精度（Overall Accuracy, OA）和类别平均分类精度（Average Accuracy, AA）作为评价指标，相关结果以 5 次独立实验的均值及标准差（Mean $\pm$ Std）报告。本文提出的模型涉及三个关键超参数，分别为：①视觉-语义跨模态对齐模块损失 L_{KG} 中 KL 散度损失 D_{KL} 的权重 γ ；②视觉-语义跨模态对齐模块损失 L_{KG} 中跨模态重构损失 $L_{跨模态重构}$ 的权重 α ；③视觉-语义跨模态对齐模块损失 L_{KG} 中跨模态对齐损失 $L_{跨模态对齐}$ 的权重 β 。针对本文所采用的数据集，通过实验验证后确定最优参数配置为 $\alpha = 10$ 、 $\beta = 10$ 、 $\gamma = 0.001$ ，所有实验均在上述最优参数设置下完成。

2.3 知识图谱语义表示可视化

本文采用 t-SNE 方法对通过 DistMult 知识表示学习获得的类别语义向量进行可视化。t-SNE 的目标是在尽可能保持数据点局部结构关系的前提下，将高维特征映射到二维空间中。图中每个点表示一个类别，不同点之间的距离反映了类别之间语义关系的接近程度，距离越近表示语义相似性越高。从图 5 可以观察到，诸如草地和树木等语义属性相近的类别在嵌入空间中分布较为接近，而由于其属性存在差异，这些类别之间仍保持一定的区分度。相比之下，桥梁与建筑等在语义属性上差异显著的类别，其在嵌入空间中的距离明显更远。

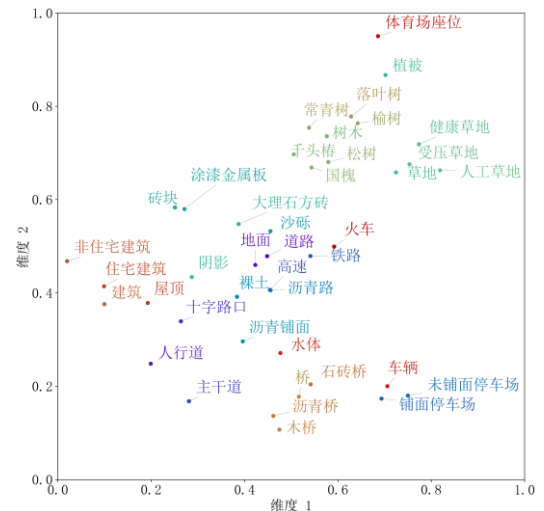


图 5 知识图谱语义特征的 t-SNE 可视化
Fig.5 t-SNE visualization of knowledge graph semantic features

表 5 不同零样本方法在 PU 数据集上的分类性能对比
Table 5 Comparison of classification performance of different ZSL methods on the PU dataset

类别名称	SJE	DeViSE	ALE	ESZSL	DEM	SAE	RN	GAZSL	CVAE	KG-ZSNet
树木	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	42.434 ± 32.755	6.695 ± 0.137	0.651 ± 0.202	53.326 ± 37.701	94.460 ± 4.078	70.955 ± 30.591	84.764 ± 1.476
沥青铺面	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	98.613 ± 1.595	0.000 ± 0.000	99.627 ± 0.446	98.880 ± 0.662	12.640 ± 7.849	83.680 ± 27.786	98.827 ± 0.320
停车场	100.000 ± 0.000	100.000 ± 0.000	100.000 ± 0.000	54.821 ± 45.749	62.500 ± 4.685	0.000 ± 0.000	66.786 ± 20.356	25.893 ± 26.270	39.821 ± 35.118	82.679 ± 12.289
AA(%)	33.333 ± 0.000	33.333 ± 0.000	33.333 ± 0.000	65.290 ± 17.877	23.065 ± 1.531	33.426 ± 0.165	72.997 ± 13.764	44.331 ± 9.978	64.819 ± 9.024	88.756 ± 4.039
OA(%)	7.778 ± 0.000	7.778 ± 0.000	7.778 ± 0.000	58.028 ± 21.490	9.292 ± 0.312	26.375 ± 0.180	66.236 ± 24.814	67.820 ± 1.434	71.847 ± 18.035	88.264 ± 1.301

表 6 不同零样本方法在 HU12 数据集上的分类性能对比
Table 6 Comparison of classification performance of different ZSL methods on the HU12 dataset

类别名称	SJE	DeViSE	ALE	ESZSL	DEM	SAE	RN	GAZSL	CVAE	KG-ZSNet
受压草地	99.825 ± 0.392	96.252 ± 1.769	97.528 ± 1.658	79.968 ± 6.832	0.000 ± 0.000	57.416 ± 7.700	83.987 ± 4.722	93.222 ± 4.227	79.697 ± 15.480	93.397 ± 8.919
非住宅建筑	41.367 ± 15.735	76.576 ± 2.147	78.762 ± 0.760	9.518 ± 4.857	35.941 ± 1.012	0.788 ± 0.223	0.322 ± 0.390	3.826 ± 5.669	15.964 ± 2.320	16.174 ± 4.029
铁路	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	69.376 ± 2.562	46.607 ± 2.720	88.259 ± 4.900	84.227 ± 4.582	60.194 ± 7.939	46.462 ± 12.120	72.923 ± 5.840
未铺面停车场	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	79.318 ± 0.604	70.021 ± 5.056	0.853 ± 1.086	14.883 ± 11.427	10.576 ± 1.830	85.075 ± 6.720	69.296 ± 9.852
AA(%)	35.298 ± 3.955	43.207 ± 0.779	44.072 ± 0.543	59.546 ± 1.318	38.142 ± 0.808	36.829 ± 0.496	45.854 ± 2.973	41.955 ± 0.898	56.799 ± 5.190	62.948 ± 4.282
OA(%)	42.037 ± 4.683	51.400 ± 0.970	52.423 ± 0.648	55.926 ± 1.612	32.121 ± 0.203	43.403 ± 0.778	51.575 ± 1.601	47.825 ± 0.780	51.661 ± 5.950	61.828 ± 3.714

表 7 不同零样本方法在 HU18 数据集上的分类性能对比
Table 7 Comparison of classification performance of different ZSL methods on the HU18 dataset

类别名称	SJE	DeViSE	ALE	ESZSL	DEM	SAE	RN	GAZSL	CVAE	KG-ZSNet
受压草地	100.000 ± 0.000	99.969 ± 0.070	96.885 ± 0.560	64.938 ± 2.291	97.788 ± 0.903	0.000 ± 0.000	97.196 ± 3.684	33.256 ± 16.630	58.816 ± 7.930	74.704 ± 18.200
常青树	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	2.774 ± 1.100	0.000 ± 0.000	0.262 ± 0.587	2.996 ± 2.221	0.000 ± 0.000	42.434 ± 12.378	28.052 ± 13.283
住宅建筑	0.000 ± 0.000	5.827 ± 3.342	97.170 ± 0.312	73.769 ± 2.682	28.502 ± 3.908	98.579 ± 2.450	40.165 ± 14.377	26.586 ± 19.040	76.168 ± 0.492	73.972 ± 6.712
铁路	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	89.728 ± 0.306	1.365 ± 0.540	30.171 ± 9.577	99.522 ± 0.389	99.864 ± 0.187	40.887 ± 10.933	64.216 ± 8.804
AA(%)	25.000 ± 0.000	26.448 ± 0.844	48.514 ± 0.111	57.802 ± 6.028	31.914 ± 1.060	32.253 ± 1.955	59.970 ± 3.576	39.926 ± 8.880	54.576 ± 5.985	60.236 ± 0.307
OA(%)	34.825 ± 0.000	37.301 ± 1.487	75.275 ± 0.154	62.457 ± 0.556	46.347 ± 1.778	44.573 ± 0.359	59.360 ± 6.124	30.881 ± 13.910	62.436 ± 4.205	65.007 ± 7.347

2.4 对比实验结果及分析

2.4.1 零样本高光谱图像分类方法对比实验

为全面评估所提出方法的性能，本文选取多种具有代表性的零样本方法作为对比模型。为了追求公平的比较，对比模型的参数是根据其论文中推荐的参数进行设置的，三个数据集相关实验结果如表 5-7 所示。具体而言，SJE(Akata et al., 2015)、DeViSE(Frome et al., 2013)、ALE(Akata et al., 2013)、ESZSL(Romera-Paredes and Torr, 2015)、DEM(Zhang et al., 2017)、SAE(Kodirov et al., 2017)以及 RN(Sung et al., 2018)均属于基于嵌入学习的零样本方法，而 GAZSL(Zhu et al., 2018)和 CVAE(Mishra et al., 2018)则属于基于生成模型的零样

本方法。本文提出的 KG-ZSNet 模型使用双重变分自编码器结构，属于基于生成模型的零样本方法。

由表 5 可知，在 PU 数据集上，不同零样本方法在未知类别上的分类性能差异较为明显。传统嵌入式方法在部分类别上能够取得较高精度，但在其余类别上表现较弱，导致整体分类性能受限。生成式方法整体表现更为稳健，但在不同类别上的分类精度仍存在较大方差。相比之下，KG-ZSNet 在 PU 数据集上在所有未知类别上均取得较高的分类精度，其 OA 和 AA 分别达到 88.264%和 88.756%的平均分类精度，且标准差较小，表现出更为一致和稳定的分类性能。由表 6 可知，在 HU12 数据集上，由于城市场景更加复杂、地物类别差异较大，各类

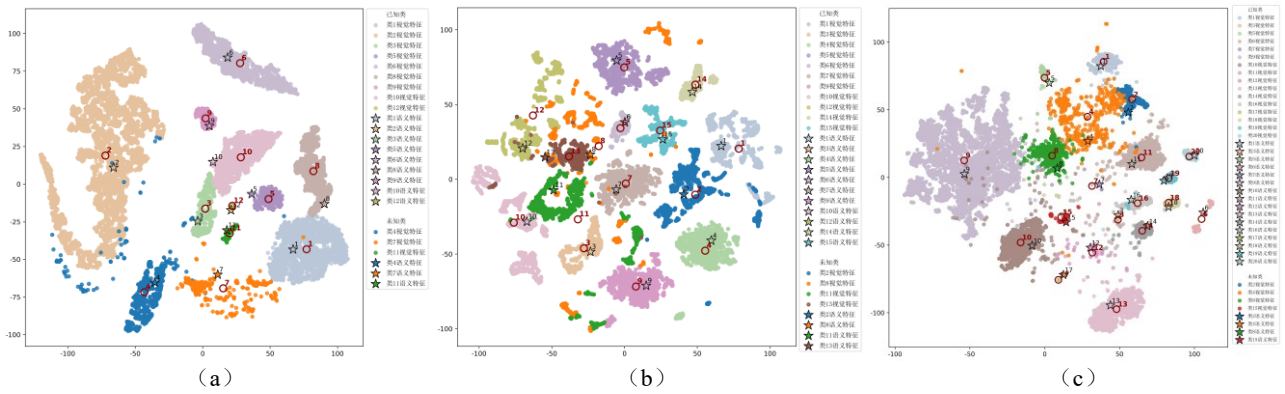


图 6 三个数据集已知类与未知类的隐层对齐特征散点图。(a) PU 数据集；(b) HU12 数据集；(c) HU18 数据集。
Fig.6 Latent alignment feature scatter plots of seen and unseen classes on three datasets. (a) PU dataset; (b) HU12 dataset; (c) HU18 dataset.

零样本方法的整体性能均有所下降。部分嵌入式方法在“铁路”和“未铺面停车场”等类别上仍存在明显性能退化，限制了其整体分类效果。生成式方法虽在个别类别上取得较好结果，但在不同类别之间的性能差异较大，整体均衡性不足。KG-ZSNet 在 HU12 数据集上 OA 和 AA 分别达到 61.828% 和 62.948% 的平均分类精度，且标准差较小，在多数未知类别上保持了相对稳定的分类表现。由表 7 可知，在更具挑战性的 HU18 数据集上，不同方法之间的性能差距进一步拉大。嵌入式方法在多个未知类别上依然难以有效识别目标类别，而生成式方法虽然在特定类别上表现较好，但整体稳定性有限。KG-ZSNet 在 HU18 数据集上仍保持最优的整体性能，其 OA 和 AA 分别达到 65.007% 和 60.236% 的平均分类精度。综合三组实验结果可以看出，嵌入式零样本方法对语义表示质量和嵌入空间结构较为敏感，在高光谱复杂场景下易出现性能退化；生成式方法通过生成未知类别的隐层表示，在一定程度上缓解了样本缺失问题。KG-ZSNet 在使用 VAE 生成式零样本模型的同时，进一步引入知识图谱提供的结构化语义信息，增强了类别语义的判别性，使模型在不同数据集和场景下均取得了稳定且具有竞争力的分类性能，验证了所提出方法在零样本高光谱遥感图像分类任务中的有效性与稳定性。

2.4.2 隐层对齐特征可视化分析

为更直观地展示本文方法在不同数据集上的视觉-语义跨模态对齐效果，本文采用 t-SNE 对模型学习得到的对齐隐层特征进行二维可视化，结果如图 6 所示。图中分别给出了 PU、HU12 和 HU18 数据集上已知类与未知类样本在隐层空间中的特征分布情况。为便于辨别视觉特征簇的类别归属，我们计算了各类别视觉特征簇的中心位置，并采用红色空

心圆和红色类别编号进行标记。从图 6 整体来看，三个数据集的已知类视觉特征与语义特征间均表现出较好的对齐效果，说明所设计的对齐模块能够有效实现已知类视觉特征与语义特征的一致性对齐。

对于零样本学习，未知类视觉-语义对齐是未知类分类的关键，由图 6 (a) 可见，在 PU 数据集中，第 4 类“树木”、第 7 类“沥青铺面”和第 11 类“停车场”这三个未知类的语义特征与对应视觉特征簇之间具有较好的对齐关系，利用其语义特征训练得到的未知类分类器，能够较好地地区分对应的未知类视觉特征，其分类精度也应相对较高。由图 6 (b) 可见，在 HU12 数据集中，第 8 类“非住宅建筑”的视觉特征分布较为分散。尽管其语义特征整体上在相应视觉特征区域附近，但视觉特征簇本身聚集性较弱，并且与邻近的第 11 类“铁路”和第 13 类“未铺面停车场”这两类未知类存在一定重叠，因此该类在分类时较易与相近类别发生混淆，其分类精度相对较低。而第 2 类“受压草地”、第 11 类“铁路”和第 13 类“未铺面停车场”的语义特征与视觉特征对齐关系较清晰，视觉特征类内聚集性也相对更好，与其他未知类别的分离程度较高，因此这些类别分类精度较高。由图 6 (c) 可见，在 HU18 数据集中，第 4 类“常青树”虽然在一定程度上实现了语义特征与视觉特征的对齐，但该类的视觉特征簇在隐层空间中仍存在较大的类内离散性以及一定的类间混叠，因此其分类精度可能受到明显影响。而第 2 类“受压草地”、第 8 类“住宅建筑”和第 15 类“铁路”的语义特征与视觉特征簇对齐关系较为明确，类内分布相对更为紧凑，与其他未知类别之间也具有更好的分离性，因此这些未知类更可能获得较高的分类精度。

综合来看，三个数据集上的隐层对齐特征散点图分析结果与定量实验结论保持一致，能够较好地反映各未知类的实际分类表现。

2.4.3 语义知识向量对比实验

为进一步验证知识图谱在模型中的优势，本文还使用了三种其他知识表示学习方法进行对比实验，包括对比式语言-图像预训练模型 CLIP 以及两种自然语言处理模型 Word2Vec 和 BERT。在对比实验中，分别使用上述三种先验知识获得的语义向量替代知识图谱语义向量，对应的方法分别记为 Word2Vec-ZSNet、Bert-ZSNet 和 CLIP-ZSNet。对于 Word2Vec 方法，采用预训练的 Word2Vec 模型将类别标签映射为 300 维的语义向量。对于 BERT 方法，首先为 43 个地物类别分别整理对应的领域描述文本，例如对人工草地的描述为“人工制造的一片草地”，随后利用预训练的 BERT 模型将上述领域描述映射为 1024 维的语义向量。对于 CLIP 方法，同样利用预训练的 CLIP-ViT-B-32 模型将领域描述文本映射为 512 维的语义向量。

三个数据集采用不同语义向量的对比实验结果如表 8-10 所示。与其他几种语义表示相比，基于知识图谱语义表示的方法能够更有效地引导深度网络进行零样本分类。相较于表现相对较好的 Bert-ZSNet 模型，KG-ZSNet 在 PU 数据集上的总体分类精度（OA）平均提升了 1.681%，在 HU12 数据集上平均提升了 7.306%，在 HU18 数据集上平均提升了 6.032%，充分验证了知识图谱语义向量在提升零样本分类精度方面的有效性。

表 8 不同语义知识在 PU 数据集上的分类性能对比
Table 8 Comparison of classification performance of different semantic knowledge on the PU dataset

类别名称	Word2Vec -ZSNet	CLIP -ZSNet	Bert -ZSNet	KG-ZSNet
树木	91.438 ± 1.985	85.876 ± 3.704	84.596 ± 8.954	84.764 ± 1.476
	27.573 ± 19.085	98.667 ± 0.675	95.467 ± 4.069	98.827 ± 0.320
停车场	0.000 ± 0.000	48.929 ± 36.049	73.750 ± 30.778	82.679 ± 12.289
	39.670 ± 6.170	77.824 ± 12.437	84.604 ± 12.744	88.756 ± 4.039
AA(%)	67.694 ± 4.693	86.333 ± 4.349	86.583 ± 7.838	88.264 ± 1.301

表 9 不同语义知识在 HU12 数据集上的分类性能对比
Table 9 Comparison of classification performance of different semantic knowledge on the HU12 dataset

类别名称	Word2Vec -ZSNet	CLIP -ZSNet	Bert -ZSNet	KG-ZSNet
受压草地	70.510 ± 15.418	96.890 ± 1.833	98.341 ± 0.831	93.397 ± 8.919
	17.428 ± 6.874	10.627 ± 5.785	15.740 ± 3.448	16.174 ± 4.029
非住宅建筑	63.320 ± 5.590	41.895 ± 7.872	42.607 ± 1.933	72.923 ± 5.840
	35.991 ± 6.822	68.998 ± 6.968	71.599 ± 10.248	69.296 ± 9.852
AA(%)	46.812 ± 2.474	54.602 ± 2.669	57.072 ± 1.832	62.948 ± 4.282
OA(%)	48.829 ± 3.115	52.075 ± 2.218	54.522 ± 0.531	61.828 ± 3.714

表 10 不同语义知识在 HU18 数据集上的分类性能对比
Table 10 Comparison of classification performance of different semantic knowledge on the HU18 dataset

类别名称	Word2Vec -ZSNet	CLIP -ZSNet	Bert -ZSNet	KG-ZSNet
受压草地	11.402 ± 11.835	62.991 ± 9.218	74.907 ± 11.243	74.704 ± 18.200
	97.753 ± 1.475	11.835 ± 10.798	39.438 ± 19.459	28.052 ± 13.283
住宅建筑	51.383 ± 14.824	47.437 ± 23.172	49.467 ± 8.035	73.972 ± 6.712
	70.853 ± 11.649	96.587 ± 2.679	75.904 ± 13.838	64.216 ± 8.804
AA(%)	57.848 ± 2.642	54.712 ± 4.553	59.929 ± 2.823	60.236 ± 0.307
	45.723 ± 3.485	51.603 ± 7.434	58.975 ± 3.833	65.007 ± 7.347

2.5 消融实验

2.5.1 知识图谱关系消融实验

为进一步验证所构建知识图谱中不同类别属性与关系结构对零样本分类性能的贡献，本文设计了针对知识图谱关系类型的消融实验。在 1.2 节我们将知识图谱中的关系划分为三类：I 类关系为光谱相关关系，II 类关系为语义相关关系，III 类关系为空间相关关系。在此基础上我们构建单一关系类型或不同关系组合的多种知识图谱变体，通过 DistMult 模型映射为语义向量，用于零样本分类过程，从而系统分析各类关系及其组合对分类性能的影响。

表 11-13 展示了三个数据集在不同关系组合的知识图谱变体下的消融实验结果，知识图谱关系变体后的数字代表该知识图谱包含的关系数量。实验结果表明，仅引入单一类型的关系难以充分刻画复杂城市场景中地物类别的语义差异，而不同类型关系在零样本分类任务中具有明显的互补性。当同时引入 I、II 和 III 三类关系时，KG-ZSNet 模型在平均分类精度（AA）和总体分类精度（OA）上均取得最优表现，充分验证了所构建知识图谱在结构设计上的合理性，以及结构化语义信息对提升高光谱零样本分类性能的有效作用。

表 11 不同知识图谱关系类型在 PU 数据集上的消融实验结果

Table 11 Ablation results of different knowledge graph relation types on the PU dataset

知识图谱 关系变体	I 类	II 类	III 类	AA(%)	OA(%)
KG-106	√			38.208 ± 17.457	48.153 ± 25.785
KG-102		√		37.025 ± 13.568	29.986 ± 18.308
KG-172			√	56.129 ± 18.459	64.069 ± 21.563
KG-208	√	√		67.960 ± 7.931	64.542 ± 12.135
KG-278	√		√	73.274 ± 11.108	62.694 ± 21.345
KG-274		√	√	83.571 ± 5.231	82.917 ± 6.435
KG-380 (KG-ZSNet)	√	√	√	88.756 ± 4.039	88.264 ± 1.301

表 12 不同知识图谱关系类型在 HU12 数据集上的消融实验结果

Table 12 Ablation results of different knowledge graph relation types on the HU12 dataset

知识图谱 关系变体	I 类	II 类	III 类	AA(%)	OA(%)
KG-106	√			29.154 ± 9.652	21.085 ± 9.341
KG-102		√		38.661 ± 7.940	36.492 ± 11.963
KG-172			√	47.040 ± 11.843	51.985 ± 11.325
KG-208	√	√		51.265 ± 7.335	52.023 ± 7.568
KG-278	√		√	56.416 ± 3.195	52.137 ± 5.528
KG-274		√	√	48.110 ± 5.517	52.504 ± 4.694
KG-380 (KG-ZSNet)	√	√	√	62.948 ± 4.282	61.828 ± 3.714

表 13 不同知识图谱关系类型在 HU18 数据集上的消融实验结果

Table 13 Ablation results of different knowledge graph relation types on the HU18 dataset

知识图谱 关系变体	I 类	II 类	III 类	AA(%)	OA(%)
KG-106	√			23.075 ± 6.834	15.411 ± 5.524
KG-102		√		51.164 ± 8.762	55.530 ± 8.330
KG-172			√	36.620 ± 4.399	44.421 ± 5.161
KG-208	√	√		44.788 ± 2.301	56.734 ± 3.827
KG-278	√		√	52.171 ± 8.801	58.069 ± 11.359
KG-274		√	√	52.232 ± 3.877	59.821 ± 4.597
KG-380 (KG-ZSNet)	√	√	√	60.236 ± 0.307	65.007 ± 7.347

2.5.2 损失函数消融实验

为分析各损失项在模型训练与性能提升中的作用，本文对 KG-ZSNet 模型中视觉-语义跨模态对齐模块的关键损失函数进行了系统性的消融实验。整体视觉-语义跨模态对齐模块损失函数由三部分组成，分别为基础变分自编码器损失 L_{VAE} 、跨模态重构损失 $L_{跨模态重构}$ 以及跨模态对齐损失 $L_{跨模态对齐}$ ，其对应权重参数设置为最佳参数。通过分别移除或组合不同损失项，构建多种损失函数变体，并在相同实验设置下对其分类性能进行对比分析。

表 14-16 展示了三个数据集在不同损失函数变体下的消融实验结果。实验结果表明， L_{VAE} 提供了稳定且具有良好泛化能力的隐层表示空间， $L_{跨模态重构}$ 增强了视觉特征与语义特征之间的双向一致性， $L_{跨模态对齐}$ 进一步强化了类别层面的判别结构。仅引入单一损失函数时，模型整体分类性能较低，说明单独的隐层分布建模或跨模态约束难以支撑零样本分类任务。引入两种损失函数后，模型性能均得到显著提升，验证了不同损失项之间的互补性。当三种损失项同时引入时，模型能够更有效地建立视觉空间与语义空间之间的映射关系，在 AA 和 OA 上均取得最优性能，验证了完整损失函数设

计的合理性。

表 14 不同损失函数组合在 PU 数据集上的消融实验结果
Table 14 Ablation results of different loss function combinations on the PU dataset

L_{VAE}	$L_{跨模态重构}$	$L_{跨模态对齐}$	AA(%)	OA(%)
√			47.290 ± 9.138	48.458 ± 13.163
	√		44.182 ± 12.465	49.583 ± 28.355
		√	33.333 ± 0.000	26.042 ± 0.000
√	√		79.473 ± 21.404	77.764 ± 23.622
√		√	46.474 ± 15.963	54.708 ± 18.810
	√	√	45.924 ± 16.560	51.583 ± 30.051
√	√	√	88.756 ± 4.039	88.264 ± 1.301

表 15 不同损失函数组合在 HU12 数据集上的消融实验结果
Table 15 Ablation results of different loss function combinations on the HU12 dataset

L_{VAE}	$L_{跨模态重构}$	$L_{跨模态对齐}$	AA(%)	OA(%)
√			26.535 ± 5.206	27.101 ± 5.069
	√		25.323 ± 14.390	26.678 ± 15.636
		√	26.283 ± 2.295	31.223 ± 2.798
√	√		59.408 ± 4.347	57.744 ± 3.396
√		√	60.185 ± 3.640	61.766 ± 1.408
	√	√	61.197 ± 4.380	60.590 ± 2.781
√	√	√	62.948 ± 4.282	61.828 ± 3.714

表 16 不同损失函数组合在 HU18 数据集上的消融实验结果
Table 16 Ablation results of different loss function combinations on the HU18 dataset

L_{VAE}	$L_{跨模态重构}$	$L_{跨模态对齐}$	AA(%)	OA(%)
√			22.930 ± 4.896	23.792 ± 6.551
	√		25.226 ± 4.675	28.131 ± 8.151
		√	24.997 ± 0.006	39.571 ± 3.875
√	√		42.815 ± 7.578	40.423 ± 11.725
√		√	44.213 ± 5.391	60.651 ± 7.149
	√	√	54.104 ± 4.417	59.040 ± 11.195
√	√	√	60.236 ± 0.307	65.007 ± 7.347

2.6 模型敏感性分析

2.6.1 未知类别设置敏感性分析

算法对未知类别设置具有一定敏感性，这是零样本学习任务中的普遍现象。未知类零样本分类精度受到多种因素影响，包括未知类与已知类之间的语义关联性、未知类自身的分布聚集性，以及未知类与其他未知类之间的特征判别性。为验证本文方法在不同未知类别设置下的表现，以 HU12 数据集为例，我们在前文选择 4 个未知类的基础上，又设置了 2 个未知类、3 个未知类两种情况，并随机选择不同未知类组合进行了未知类分类实验。表 17-18 进一步给出了不同零样本方法在不同未知类划分条

表 17 不同零样本方法在 HU12 数据集上对 2 类未知类的分类性能对比

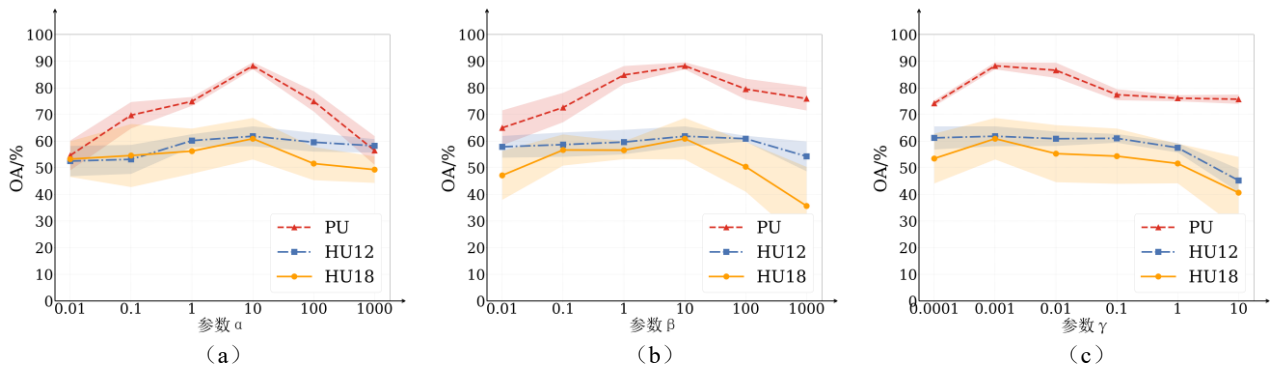
Table 17 Comparison of classification performance of different ZSL methods on two unseen classes of the HU12 dataset

类别名称	SJE	DeViSE	ALE	ESZSL	DEM	SAE	RN	GAZSL	CVAE	KG-ZSNet
受压草地	100.000 ± 0.000	100.000 ± 0.000	100.000 ± 0.000	93.334 ± 2.953	38.517 ± 8.308	63.894 ± 1.595	93.334 ± 2.953	51.060 ± 17.666	73.113 ± 3.268	99.330 ± 0.396
铁路	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	86.154 ± 1.672	98.121 ± 4.041	86.397 ± 0.736	86.154 ± 1.672	99.514 ± 0.079	99.676 ± 0.198	93.279 ± 3.907
AA(%)	50.000 ± 0.000	50.000 ± 0.000	50.000 ± 0.000	89.744 ± 1.615	68.119 ± 3.498	75.345 ± 0.897	89.744 ± 1.615	75.187 ± 9.133	86.333 ± 1.677	96.305 ± 1.797
OA(%)	50.382 ± 0.000	50.382 ± 0.000	50.382 ± 0.000	89.771 ± 1.758	68.892 ± 3.235	74.857 ± 0.985	89.771 ± 1.758	75.203 ± 10.517	83.630 ± 1.748	96.328 ± 1.781

表 18 不同零样本方法在 HU12 数据集上对 3 类未知类的分类性能对比

Table 18 Comparison of classification performance of different ZSL methods on three unseen classes of the HU12 dataset

类别名称	SJE	DeViSE	ALE	ESZSL	DEM	SAE	RN	GAZSL	CVAE	KG-ZSNet
正常草地	100.000 ± 0.000	100.000 ± 0.000	100.000 ± 0.000	99.345 ± 0.608	0.592 ± 0.386	85.731 ± 0.354	99.904 ± 0.178	81.327 ± 0.932	99.361 ± 0.000	99.920 ± 0.179
铁路	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	86.769 ± 8.268	86.332 ± 0.253	100.000 ± 0.000	74.170 ± 5.627	96.264 ± 3.744	97.189 ± 1.177	78.889 ± 7.198
未铺面停车场	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	48.972 ± 34.260	0.000 ± 0.000	0.000 ± 0.000	97.477 ± 1.407	10.374 ± 10.956	3.578 ± 1.736	96.749 ± 1.940
AA(%)	33.333 ± 0.000	33.333 ± 0.000	33.333 ± 0.000	78.362 ± 8.682	28.974 ± 0.304	61.984 ± 0.109	90.517 ± 1.685	62.122 ± 2.625	66.009 ± 0.163	91.762 ± 2.201
OA(%)	42.931 ± 0.000	42.931 ± 0.000	42.931 ± 0.000	86.616 ± 1.571	36.443 ± 0.131	79.901 ± 0.143	88.641 ± 2.260	77.777 ± 0.343	84.350 ± 0.305	90.503 ± 2.865

图 7 KG-ZSNet 在 PU、HU12 与 HU18 数据集上的参数敏感性分析。(a) 参数 α ; (b) 参数 β ; (c) 参数 γ 。Fig.7 Sensitivity analysis of parameters in KG-ZSNet on the PU, HU12 and HU18 datasets. (a) Parameter α . (b) Parameter β . (c) Parameter γ .

件下的分类性能。实验结果表明，不同未知类别划分下的分类精度确实存在明显差异，这说明未知类别设置会直接影响零样本分类任务的难度。同时，在不同未知类别划分条件下，本文方法整体上均取得了优于其他对比方法的分类结果，说明尽管未知类别组合的变化会引起性能波动，但本文方法在不同未知类设置下仍保持了较好的稳定性和适应能力。

2.6.2 参数敏感性分析

论文存在三个自由参数： α 与 β 分别控制跨模态重构损失和跨模态对齐损失的权重， γ 用于控制 VAE 中的 KL 散度损失权重。为分析模型对关键超参数的敏感性及其在训练过程中的作用，在其余参数固定为最优取值的条件下，对损失函数中的参数 α 、 β 和 γ 进行了参数敏感性分析，相关结果如图 7 所示，模型性能以总体分类精度 (OA) 来评价。图 7 (a) 表明，PU 数据该参数较敏感，在 $\alpha=10$ 时取

得最优分类精度，HU12 和 HU18 数据对不同 α 取值不太敏感，在 1 到 10 范围内都能取得较好的分类精度。因此，我们建议 α 取值为 10，在三个数据集都具有最好的表现。图 7 (b) 表明，对于参数 β ，三个数据集在 1 到 10 范围内都能取得较好的分类效果，在 $\beta=10$ 时达到最优分类精度。图 7 (c) 表明，对于参数 γ ，在较小取值区间 (0.001~0.01) 能够获得较好分类效果，在 $\gamma=0.001$ 取值下，三组数据都得到最优分类精度，随着 γ 增大，分类性能迅速下降。

结合上述参数敏感性分析结果可以看出， α 和 β 在较大取值条件下 (1~10) 能够取得更优分类性能，表明重构损失和特征对齐损失在零样本分类中的作用，这两个损失用于实现视觉-语义跨模态对齐，适当提高其权重，有助于增强跨模态对齐效果，从而提升模型的零样本分类能力。 γ 在较小取值条件下

表 20 不同零样本方法在三个数据集上的计算复杂性与运行效率
Table 20 Computational complexity and running efficiency of different zero-shot methods on three datasets

数据集		SJE	DeViSE	ALE	ESZSL	DEM	SAE	RN	GAZSL	CVAE	KG-ZSNet
PU 数据集	总训练时间 (s)	305.647	132.054	125.510	658.627	651.142	724.045	1076.795	453.803	511.648	852.746
	总测试时间 (s)	0.124	0.106	0.116	13.541	0.150	0.284	0.082	0.109	0.056	0.018
	参数量	3.07M	3.07M	3.07M	3.07M	8.24M	6.14M	9.06M	21.13M	8.37M	10.19M
HU12 数据集	总训练时间 (s)	520.510	167.610	192.981	635.289	636.696	728.713	1068.282	506.495	392.626	654.377
	总测试时间 (s)	0.286	0.302	0.298	14.227	0.479	0.485	0.165	0.364	0.069	0.021
	参数量	3.07M	3.07M	3.07M	3.07M	8.24M	6.14M	9.06M	21.14M	8.37M	10.19M
HU18 数据集	总训练时间 (s)	723.069	680.758	561.949	652.064	707.885	732.922	1097.033	476.351	616.929	1028.215
	总测试时间 (s)	0.263	0.263	0.289	15.014	0.431	0.483	0.143	0.438	0.083	0.020
	参数量	3.07M	3.07M	3.07M	3.07M	8.24M	6.14M	9.06M	21.17M	8.37M	10.19M

(0.001~0.01) 时模型性能更优, 表明 VAE 损失中的特征重构损失更为关键, KL 散度损失在本文模型中主要起到对隐层空间分布进行约束和正则化的作用, 若其权重过大, 可能会对特征表达能力形成较强限制, 不利于模型学习更具判别性的隐层表示。

2.7 计算复杂性与运行效率分析

为进一步分析本文方法的计算复杂性与运行效率, 我们针对 KG-ZSNet 模型内部模块, 在不同数据集上分析模型各组成部分的主要时间开销来源。表 19 展示了 KG-ZSNet 中不同模块在三个数据集上的训练时间效率, 以每轮训练时间 (s/epoch) 进行统计。可以看出, 模型的主要时间开销集中在视觉特征提取模块和视觉-语义跨模态对齐模块, 其中视觉-语义跨模态对齐模块占据主导地位, 是整体训练过程中计算负担最重的部分。语义特征提取模块和未知类分类模块的时间开销均较小。这表明本文方法的复杂度主要来源于跨模态对齐学习过程。

表 19 KG-ZSNet 不同模块的训练时间效率 (s/epoch)
Table 19 Training time efficiency of different models in KG-ZSNet (s/epoch)

模块	PU 数据集	HU12 数据集	HU18 数据集
视觉特征提取模块	1.436	1.123	0.836
语义特征提取模块	0.034	0.034	0.034
视觉-语义跨模态对齐模块	2.173	2.835	3.421
未知类分类模块	0.049	0.043	0.052

我们进一步对不同零样本方法的参数量、总训练时间和总测试时间进行比较, 以评估本文方法在模型规模与运行效率方面的表现。表 20 展示了不同零样本方法在三个数据集上的计算复杂性与运行效率。其中, 参数量用于衡量模型规模, 总训练时间和总测试时间分别用于反映模型在训练阶段和测试阶段的整体时间开销。从表 18 可以看出, 不同类型零样本方法在计算复杂性与运行效率方面表现出明显差异。嵌入式方法中, SJE、DeViSE、ALE 和 ESZSL 整体参数量较小, 训练开销相对较低, 体现

出较好的轻量性; DEM、SAE 和 RN 由于引入了更复杂的映射或关系建模机制, 参数量和训练时间进一步增加。生成式方法 GAZSL 和 CVAE 通过生成伪视觉特征来增强未知类表征能力, 模型规模和训练开销整体高于一般嵌入式方法。本文提出的 KG-ZSNet 在参数量和训练时间上处于较高水平, 但测试时间在三个数据集上始终保持较低水平, 说明其主要计算负担集中在训练阶段, 而推理过程仍具有较高效率。总体来看, 尽管 KG-ZSNet 的结构相对复杂, 但其测试分类精度最高, 同时测试阶段时间开销较低, 表现出较好的综合性能。

3 结论

针对高光谱遥感数据中标注类别不完整, 未标注地物类别无法被传统监督学习方法有效处理这一实际应用问题, 本文提出了一种基于知识图谱语义表示的高光谱遥感图像零样本分类网络 KG-ZSNet。该方法通过构建契合高光谱城市场景特性的遥感领域知识图谱, 引入结构化语义信息, 并在生成式零样本学习框架下, 利用变分自编码器对高光谱视觉特征与语义特征进行跨模态对齐, 从而在共享隐层空间中实现对未知类别样本的有效识别。

在多个公开高光谱遥感数据集上的实验结果表明, 与现有基于嵌入学习和生成模型的零样本方法相比, KG-ZSNet 在整体分类精度上表现更优。隐层对齐特征可视化结果也表明了已知类视觉特征与语义特征在共享隐层空间中具有较好的对应关系, 不同数据集上的特征分布与定量实验结果保持一致, 能够较好反映各未知类别的实际分类表现。进一步的对比实验表明, 相较于常见的语义表示方式 (如 Word2Vec、BERT 和 CLIP), 知识图谱语义表示能够更有效地引导深度网络进行零样本分类, 显著提升未知类别的识别精度。通过对知识图谱关系结构

的消融实验, 本文验证了不同类别属性与关系类型在零样本分类任务中的互补作用, 完整关系结构的引入在平均分类精度和总体分类精度上均取得最优表现, 体现了知识图谱结构设计的合理性。损失函数消融实验与模型敏感性分析结果进一步表明, 所提出模型在关键损失项与参数配置下具有良好的稳定性与鲁棒性。计算复杂性与运行效率分析结果表明, KG-ZSNet 的计算开销主要集中在训练阶段, 但测试阶段时间开销较低, 模型在保持较高未知类别分类精度的同时兼顾了测试效率。

在现有研究的基础上, 未来研究可进一步从知识表示优化与任务场景拓展两个方面展开。基于本文实验结果, 后续研究仍将继续关注知识图谱构建与语义表示学习, 通过更加精细的属性设计和关系设计, 增强类别语义表示对高光谱地物特征及类别间关联的表达能力, 从而进一步提升基于知识图谱的零样本分类精度。在此基础上, 可进一步拓展至广义零样本学习 (Generalized Zero-Shot Learning, GZSL) 场景, 即在测试阶段同时对已知类别与未知类别进行识别, 从而更全面评估模型在实际遥感应应用条件下的泛化能力与稳定性。

References

- Akata,Z., Perronnin,F., Harchaoui,Z., et al., 2013. Label-Embedding for Attribute-Based Classification. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Portland, OR, USA: IEEE, 819-826. <https://doi.org/10.1109/CVPR.2013.111>.
- Akata,Z., Reed,S., Walter,D., et al., 2015. Evaluation of Output Embeddings for Fine-Grained Image Classification. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, MA, USA: IEEE, 2927-2936. <https://doi.org/10.1109/CVPR.2015.7298911>.
- Bollacker,K., Evans,C., Paritosh,P., et al., 2008. Freebase: A Collaboratively Created Graph Database for Structuring Human Knowledge. *International Conference on Management of Data (SIGMOD)*. Vancouver Canada: ACM, 1247-1250. <https://doi.org/10.1145/1376616.1376746>.
- Bordes,A., Usunier,N., Garcia-Durón,A., et al., 2013. Translating Embeddings for Modeling Multi-Relational Data. *International Conference on Neural Information Processing Systems (NeurIPS)*. Red Hook, NY, USA: Curran Associates Inc., 2787-2795.
- Chen,J.D., Wang,A., Chen,J.J., et al., 2019. CN-Probase: A Data-Driven Approach for Large-Scale Chinese Taxonomy Construction. *IEEE International Conference on Data Engineering (ICDE)*. Macao, China: IEEE, 1706-1709. <https://doi.org/10.1109/ICDE.2019.00178>.
- Chen,J.Y., Geng,Y.X., Chen,Z., et al., 2023. Zero-Shot and Few-Shot Learning with Knowledge Graphs: A Comprehensive Survey. *Proceedings of the IEEE*, 111(6): 653-685. <https://doi.org/10.1109/JPROC.2023.3279374>.
- Chen,W.T., Xu,J.H., Wang R., et al., 2025. Hyperspectral Remote Sensing Inversion of Black Soil Organic Matter Content Based on Multi-Scale Feature Enhancement. *Earth Science*, 50(12): 4909-4918(in Chinese with English abstract).
- Chen,Z.X., Wang,Z.H., Bu,H.J., et al., 2026. A Geological Information and Enhanced Graph Convolutional Network Method for Extracting Information on Lithium and Beryllium-Rich Pegmatites from Hyperspectral Imagery. *Earth Science*, 51(3): 1057-1064(in Chinese with English abstract).
- Fensel,D., Şimşek,U., Angele,K., et al., 2020. Introduction: What Is a Knowledge Graph?. *Knowledge Graphs*. Cham: Springer International Publishing, 1-10. https://doi.org/10.1007/978-3-030-37439-6_1.
- Freitas,S., Silva,H., Silva,E., 2022. Hyperspectral Imaging Zero-Shot Learning for Remote Marine Litter Detection and Classification. *Remote Sensing*, 14(21): Art.no.5516. <https://doi.org/10.3390/rs14215516>.
- Frome,A., Corrado,G.S., Shlens,J., et al., 2013. DeViSE: A Deep Visual-Semantic Embedding Model. *Advances in Neural Information Processing Systems (NIPS)*. Lake Tahoe, Nevada: Curran Associates Inc., 2121-2129.
- Goodfellow,I.J., Pouget-Abadie,J., Mirza,M., et al., 2014. Generative Adversarial Networks. *arXiv:1406.2661*. <https://doi.org/10.48550/arXiv.1406.2661>.
- Jia,C., Yang,Y.F., Xia,Y., et al., 2021. Scaling Up Visual and Vision-Language Representation Learning with Noisy Text Supervision. *38th International Conference on Machine Learning (ICML)*. PMLR, 4904-4916.
- Jiang,B.C., Wan,G., Xu,J., et al., 2018. Geographic Knowledge Graph Building Extracted from Multi-Sourced Heterogeneous Data. *Acta Geodaetica et Cartographica Sinica*, 47(8): 1051-1061(in Chinese with English abstract).
- Kingma,D.P., Welling,M., 2022. Auto-Encoding Variational Bayes. *arXiv:1312.6114*. <https://doi.org/10.48550/arXiv.1312.6114>.
- Kodirov,E., Xiang,T., Gong,S.G., 2017. Semantic Autoencoder for Zero-Shot Learning. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI: IEEE, 4447-4456. <https://doi.org/10.1109/CVPR.2017.473>.
- Kumar,V., Singh,R.S., Rambabu,M., et al., 2024. Deep Learning for Hyperspectral Image Classification: A survey. *Computer Science Review*, 53: Art.no.100658. <https://doi.org/10.1016/j.cosrev.2024.100658>.
- Lehmann,J., Isele,R., Jakob,M., et al., 2015. Dbpedia-A Large-Scale, Multilingual Knowledge Base Extracted from Wikipedia. *Semantic Web*, 6(2): 167-195. <https://doi.org/10.3233/SW-140134>.
- Li,Y.S., Kong,D.Y., Zhang,Y.J., et al., 2021. Robust Deep Alignment Network with Remote Sensing Knowledge

- Graph for Zero-shot and Generalized Zero-Shot Remote Sensing Image Scene Classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 179: 145-158. <https://doi.org/10.1016/j.isprsjprs.2021.08.001>.
- Li,Y.S., Wu,M.L., Zhang,Y.J., 2024. Knowledge Graph-Guided Deep Network for High-Resolution Remote Sensing Image Scene Classification. *Acta Geodaetica et Cartographica Sinica*, 53(4): 677-688(in Chinese with English abstract).
- Li,Y.S., Wang,Y., Yu,L., et al., 2025. Learning to Reason over Multi-Granularity Knowledge Graph for Zero-Shot Urban Land-Use Mapping. *Remote Sensing of Environment*, 330: Art.no.114961. <https://doi.org/10.1016/j.rse.2025.114961>.
- Liu,C., Ma,S.Q., Li,Z., et al., 2023. Multi-Level Cross-Modal Feature Alignment via Contrastive Learning towards Zero-shot Classification of Remote Sensing Image Scenes. arXiv:2306.06066.<https://doi.org/10.48550/arXiv.2306.06066>.
- Mishra,A., Reddy,S.K., Mittal,A., et al., 2018. A Generative Model For Zero Shot Learning Using Conditional Variational Autoencoders. arXiv:1709.00663. <https://doi.org/10.48550/arXiv.1709.00663>.
- Pan,E., Ma,Y., Fan,F., et al., 2021. Hyperspectral Image Classification across Different Datasets: A Generalization to Unseen Categories. *Remote Sensing*, 13(9): Art.no.1672. <https://doi.org/10.3390/rs13091672>.
- Pourpanah,F., Abdar,M., Luo,Y.X., et al., 2023. A Review of Generalized Zero-Shot Learning Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4): 4051-4070. <https://doi.org/10.1109/TPAMI.2022.3191696>.
- Romera-Paredes,B., Torr,P., 2015. An Embarrassingly Simple Approach to Zero-Shot Learning. 32nd International Conference on Machine Learning (ICML). Lille, France: PMLR, 2152-2161. https://doi.org/10.1007/978-3-319-50077-5_2.
- Sun,Z.Q., Deng,Z.H., Nie,J.Y., et al., 2019. RotatE: Knowledge Graph Embedding by Relational Rotation in Complex Space. arXiv:1902.10197. <http://arxiv.org/abs/1902.10197>.
- Sung,F., Yang,Y.X., Zhang,L., et al., 2018. Learning to Compare: Relation Network for Few-Shot Learning. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, UT, USA: IEEE, 1199-1208. <https://doi.org/10.1109/CVPR.2018.00131>.
- Tan,X.M., Xi,B.B., Li,J.J., et al., 2024. Review of Zero-Shot Remote Sensing Image Scene Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 11274-11289. <https://doi.org/10.1109/JSTARS.2024.3410995>.
- Trouillon,T., Welbl,J., Riedel,S., et al., 2016. Complex Embeddings for Simple Link Prediction. 33rd International Conference on International Conference on Machine Learning (ICML). New York, NY, USA: PMLR, 2071-2080.
- Wang,W., Zheng,V.W., Yu,H., et al., 2019. A Survey of Zero-Shot Learning: Settings, Methods, and Applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2): 13:1-13:37. <https://doi.org/10.1145/3293318>.
- Xie,G.S., Zhang,Z., Shao,L., et al., 2023. Towards Zero-Shot Learning: A Brief Review and an Attention-Based Embedding Network. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(3). <https://doi.org/10.1109/TCSVT.2022.3208071>.
- Xu,B., Xu,Y., Liang,J.Q., et al., 2017. CN-DBpedia: A Never-Ending Chinese Knowledge Extraction System. International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (IEA/AIE). Arras, France: Springer, 428-438. https://doi.org/10.1007/978-3-319-60045-1_44.
- Yang,B.S., Yih,W.T., He,X.D., et al., 2014. Embedding Entities and Relations for Learning and Inference in Knowledge Bases. International Conference on Learning Representations (ICLR). Banff, AB, Canada: OpenReview, 1-6.
- Zhang,L., Xiang,T., Gong,S.G., 2017. Learning a Deep Embedding Model for Zero-Shot Learning. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 3010-3019. <https://doi.org/10.1109/CVPR.2017.321>.
- Zhang,X.Y., Zhang,C.J., Wu,M.G., et al., 2020. Spatio-Temporal Features Based Geographical Knowledge Graph Construction. *Scientia Sinica Informationis*, 50(7): 1019-1032(in Chinese with English abstract).
- Zhu,Y.Z., Elhoseiny,M., Liu,B.C., et al., 2018. A Generative Adversarial Approach for Zero-Shot Learning from Noisy Texts. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, UT, USA: IEEE, 1004-1013. <https://doi.org/10.1109/CVPR.2018.00111>.

附中文参考文献

- 陈伟涛, 徐佳辉, 王锐, 等, 2025. 基于多尺度特征增强的黑土有机质含量高光谱卫星遥感反演. *地球科学*, 50(12): 4909-4918.
- 陈志行, 王正海, 卜浩坚, 等, 2025. 基于地质信息的改进 GCN 高光谱富锂铍伟晶岩信息提取方法. *地球科学*, 51(3): 1057-1064.
- 蒋秉川, 万刚, 许剑, 等, 2018. 多源异构数据的大规模地理知识图谱构建. *测绘学报*, 47(8): 1051-1061.
- 李彦胜, 吴敏郎, 张永军, 2024. 知识图谱约束深度网络的高分辨率遥感影像场景分类. *测绘学报*, 53(4): 677-688.
- 张雪英, 张春菊, 吴明光, 等, 2020. 顾及时空特征的地理知识图谱构建方法. *中国科学:信息科学*, 50(7): 1019-1032.