



耦合知识图谱和深度学习的新一代遥感影像解译范式

李彦胜¹ 张永军¹

1 武汉大学遥感信息工程学院,湖北 武汉,430079

摘要:在遥感大数据时代,遥感影像智能解译是挖掘遥感大数据价值并推动若干重大应用的关键技术,如何将知识推理和数据学习两类解译方法有机结合已成为遥感大数据智能处理的重要研究趋势。由此提出了面向遥感影像解译的遥感领域知识图谱构建与进化方法,建立了顾及遥感成像机理和地理学知识的遥感领域知识图谱。在遥感领域知识图谱支撑下,以零样本遥感影像场景分类、可解释遥感影像语义分割以及大幅面遥感影像场景图生成3个典型的遥感影像解译任务为例,研究了耦合知识图谱和深度学习的新一代遥感影像解译范式。在零样本遥感影像场景分类实验中,所提方法在不同的可见类/不可见类比例和不同的语义表示下,都明显优于其他方法;在可解释遥感影像语义分割实验中,知识推理与深度学习的联合方法取得了最好的分类结果;在大幅面遥感影像场景图生成实验中,知识图谱引导的方法精度明显高于基准的频率统计方法。遥感知识图谱推理与深度数据学习的融合可以有效提升遥感影像的解译性能。

关键词:数据学习与知识推理;遥感知识图谱;深度学习;遥感影像智能解译

中图分类号:P237

文献标志码:A

随着遥感科学、航空航天、导航通信等领域技术的飞速发展,遥感大数据时代已来临^[1-2]。海量遥感影像数据的自动精确解译是一项十分基础且重要的工作,尽管国内外研究人员提出了大量遥感影像解译方法,遥感影像自动解译技术还远远不能满足行业单位的使用需求,亟需研究更加精准、可靠、智能的遥感影像解译方法^[3]。

遥感影像解译方法随着人工智能技术的发展变化而不断演进。从发展历程来看,人工智能技术主要经历了三大阶段,即以符号主义为主要特点的第一代人工智能^[4],以联结主义为主要特点的第二代人工智能^[5]和以知识推理与数据学习联合为主要特点的第三代人工智能^[6]。第一代和第二代人工智能技术分别为单方面基于知识层面和数据层面模拟人类的智能行为,因此存在各自的局限性。在数据大爆炸背景下,如何通过知识推理与数据学习互补发展促进人工智能技术的提升,是人工智能研究与应用中亟待解决的重大数理基础问题之一。尽管第三代人工智能具有重大提升潜力,但是知识推理与数据学习耦合

难度很大,尚处在萌芽期。为了推动该技术的发展,2021年国际上专门成立了知识推理与数据学习联合国际会议。

在遥感信息领域,现有的遥感影像解译方法可分为知识驱动方法和数据驱动方法两大类。知识驱动的遥感影像解译方法是在第一代人工智能技术的基础上发展起来的,它能利用遥感解译专家在解译过程中的决策知识^[7]或者人眼视觉识别机制^[8],借助先验知识推理完成遥感影像解译。总体来说,现有的知识驱动的遥感影像解译方法只是对遥感专家决策过程和解译规则的简单模仿,通过对静态知识的应用实现遥感影像的解译工作,难以充分利用复杂的遥感成像机理、遥感影像附载的地学知识。第三代人工智能技术的兴起给遥感影像智能解译提供了方向,即将知识推理和数据学习结合起来,建立一个以数据自主学习为核心,以领域先验知识推理为引导的新一代遥感影像解译范式。

一般来说,知识推理高度依赖大规模领域知识库,然而目前还不存在面向遥感影像解译的大

收稿日期:2021-11-28

项目资助:国家自然科学基金(42030102,41971284);湖北省自然科学基金创新群体(2020CFA003)。

第一作者:李彦胜,博士,副教授,研究方向为遥感大数据智能挖掘。yansheng.li@whu.edu.cn

通讯作者:张永军,博士,教授。zhangyj@whu.edu.cn

规模知识库。在知识工程领域,知识图谱^[9]能以结构化的形式描述客观世界的概念、实体及其之间的关系,有希望将现实世界中的遥感信息表达为更接近人类认知的形式,更好地组织和管理专家先验知识、遥感成像机理、遥感影像附载的地理学知识。在遥感领域知识图谱基础上,如何通过知识图谱推理与深度数据学习互补发展促进遥感影像智能解译性能的提升,实现数据-信息-知识的智能转换,成为亟待解决的重要科学问题^[10-12]。

基于上述考虑,本文研究了面向遥感影像解译的遥感知识图谱构建与进化方法,在遥感领域知识图谱基础上构建了联合知识图谱和深度学习的新一代遥感影像智能解译范式,主要包括:(1)基于知识图谱表示学习的零样本遥感影像场景分类;(2)联合知识推理和深度学习的可解释遥感影像语义分割;(3)知识图谱引导的大幅面遥感影像场景图生成。

1 遥感领域知识图谱的协同构建与自主进化

一般来说,领域知识图谱可以广泛支撑特定领域的精准检索、智能问答、主动推荐等若干应用任务,因此领域知识图谱构建也往往是应用导向的。本文着重探讨面向遥感影像智能解译的领域知识图谱构建方法及系统。一般来说,遥感影像智能解译任务需要多学科、多领域交叉得到的高质量领域知识。结合遥感解译任务的实际需要与现有知识图谱构建方法的特点,本文采用专家协同构建与多源数据挖掘相结合的方法进行遥感领域知识图谱初始构建,通过遥感知识图谱表示学习技术^[13-14]实现知识图谱的自主迭代进化,充分考虑了遥感解译任务的逆向反馈对遥感领域知识图谱进化升级的必要性,形成了一整套面向智能遥感影像解译的遥感领域知识图谱迭代建模框架,如图 1 所示。该框架主要包括自底向上的专家交互过程和自顶向下的应用反馈过程。在系统框架的底层,本文研制了一个基于 B/S(browser/server)架构的遥感领域知识图谱协同建模与迭代进化系统,领域专家们在进行协同工作的同时可以根据一定的语义相似度模型解决可能会存在的逻辑冲突问题^[15],从而使系统在服务上层模块时可以以群智协同的方式实现人机交互,是实现遥感领域知识图谱专家交互式构建的基石。在系统架构层之上,遥感领域本体设计

依据遥感影像解译任务的需要,通过遥感领域权威专家确立遥感领域知识表达模型,构建合适的知识图谱本体模型对遥感领域知识进行存储;在本体设计的基础上,基于遥感领域知识表达模型,标注并构建遥感领域知识语料库,用于训练具体的遥感领域知识三元组抽取模型,实现从多源数据中进行智能化的领域知识抽取;完成知识三元组抽取后,通过知识图谱表示学习技术,对遥感领域知识图谱进行智能化的实体对齐与关系补全,实现遥感领域知识图谱的自主进化。这几个部分共同构成了遥感领域知识图谱构建与自主进化模式,在底层系统架构的支持下,能够比较合理地实现专家交互式遥感领域知识图谱的构建工作。在最顶层,将构建起的遥感领域知识图谱应用于智能遥感影像解译任务,并依据应用结果反馈于下层的遥感领域知识图谱构建与自主进化过程,对知识图谱进行逆向反馈优化,完善遥感领域知识图谱的内容,提高遥感领域知识图谱的完备性和实用性。

1.1 顾及遥感成像机理和地理时空知识的遥感领域本体设计

为了满足遥感影像智能解译任务的需要,首先需要参考自顶向下的知识图谱构建方式,确立遥感领域知识表达模型与对应的遥感领域本体模式,从而保证构建的遥感领域知识图谱对于领域内知识表达的权威性、正确性、体系完备性与可用性,使得遥感领域知识图谱能够服务于最终的遥感影像解译任务。

遥感本体设计是将遥感领域内错综复杂的多学科知识按照统一的体系结构进行组织的基石,是推动遥感领域完成完备性及易用性知识表达的关键。相比于通用知识图谱,遥感领域知识图谱需要包含空间对象的时空知识,即对实体空间位置、空间分布、空间形态、空间关系、空间统计、空间关联、空间对比、空间趋势、空间运动、时空变化、趋势分析等信息进行概括和凝练,形成具有时空特性的知识^[16]。因此,在进行遥感领域本体设计时,需要合理组织空间对象相关的语义信息与时空信息。此外,还需结合遥感影像解译任务,引入遥感成像探测机理相关知识,并通过地理学知识对遥感成像机理进行补充^[10],表达出空间对象相关的地理学知识对于遥感影像解译标志的作用与影响,最终形成一套将时空特征作为空间对象状态基本划分单元^[11],通过状态内部关系与外部变化对该空间对象相关的遥感领

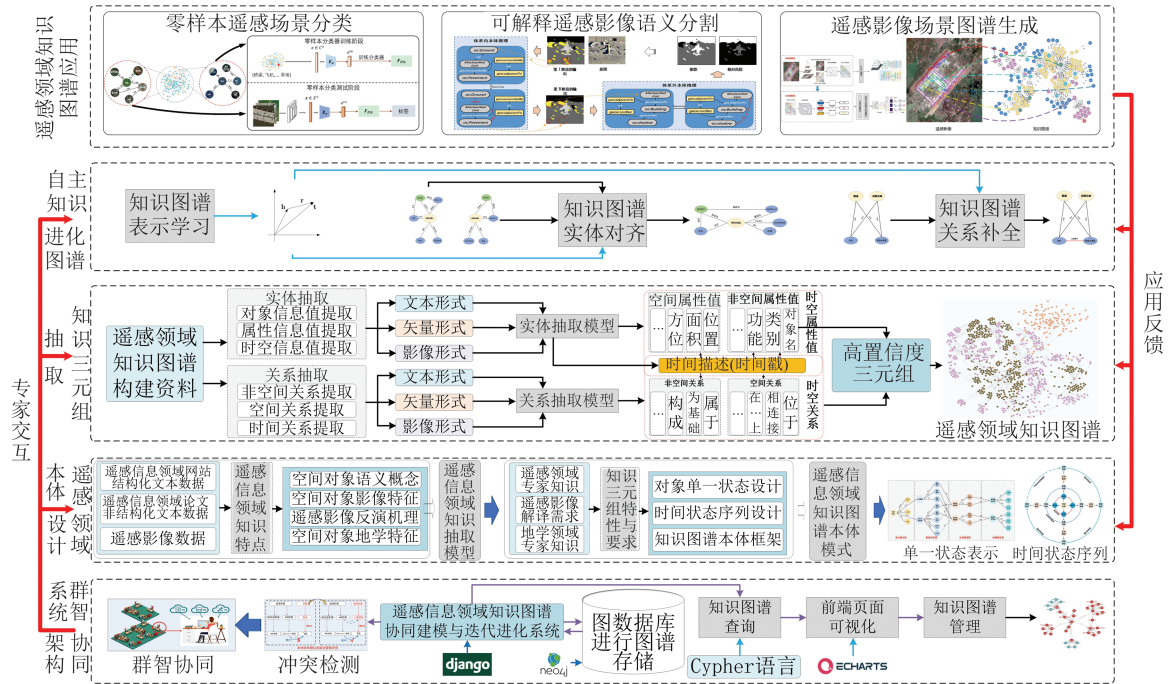


图1 面向遥感影像智能解译的遥感领域知识图谱迭代建模框架

Fig.1 Iterative Modeling Framework of Remote Sensing Knowledge Graph for Intelligent Interpretation of Remote Sensing Image

域知识的存在及发展进行表达的遥感领域知识本体模式,并将其应用到具体的知识三元组抽取模型和遥感领域知识图谱的构建任务中。

为了服务于遥感影像解译任务,首先需要构建能够将空间对象单一、特定状态的影像特征反演为实际地学特征表现的知识链路,其中空间对象具体表现为遥感影像中均质的图斑,这些图斑在地理空间中往往表示能用同一类别概括的连续的地理实体,以这类图斑为基本单元有利于遥感影像解译,可最终服务于土地覆盖分类等实际应用任务。对于空间位置等地理几何特征信息而言,这一链路的途径为遥感传感器的成像模型;而对于空间对象的地理属性特征信息而言,这一链路的途径为遥感定量反演模型。因此,为了反映这一知识链路的特性,对于知识图谱中表现空间对象单一状态的结构,设计了语义概念层—影像特征层—观测机理层—地理规律层的4层结构,如图2所示,保证了遥感领域知识图谱对于遥感影像解译任务所需知识的合理组织。

完成对于单一状态的遥感影像解译知识链路的表示后,考虑到空间对象的地理学特征往往会随着时间逐步演进,因此需要根据时间对空间对象的演进状态进行进一步组织。结合实际遥感影像解译任务的需要,以及遥感影像会按照时间序列的形式进行组织的特点,按照上述单一状

态表示结构,基于时间序列中的单景影像可以整理出空间对象在该时间节点的状态。进一步地,将时间序列中所有影像进行组织,可以获得空间对象随时间演进的状态序列,进而表现出空间对象的地理学特征随着时间演进变化的过程,如图3所示。

在完成对空间对象随时间的多状态描述之后,结合知识图谱中所包含的影像解译标志和地理规律表现等语义信息与遥感影像所具有的光谱信息与几何信息等特征,补全通过遥感影像所无法直接得到的间接特征层,从而使得遥感影像解译在知识图谱的语义辅助下能够得到更为准确或者细粒度的结果。如图4所示,以遥感影像上耕地区块的作物类别识别为例,首先,根据遥感影像所具有的坐标信息和时间信息对遥感领域知识图谱中耕地对象的子类别检索进行约束;然后,结合定量遥感反演模型,对耕地中的生物量累计进行估算,并对作物的生长周期进行拟合;最后,将作物生长周期信息与空间信息相结合,通过遥感领域知识图谱确定作物类型为双季稻,实现关联遥感领域知识图谱的遥感影像解译任务。

1.2 基于多源数据挖掘的遥感知知识三元组抽取

在自顶向下完成遥感领域本体设计和知识表达模型后,为了兼顾遥感领域知识图谱对于领

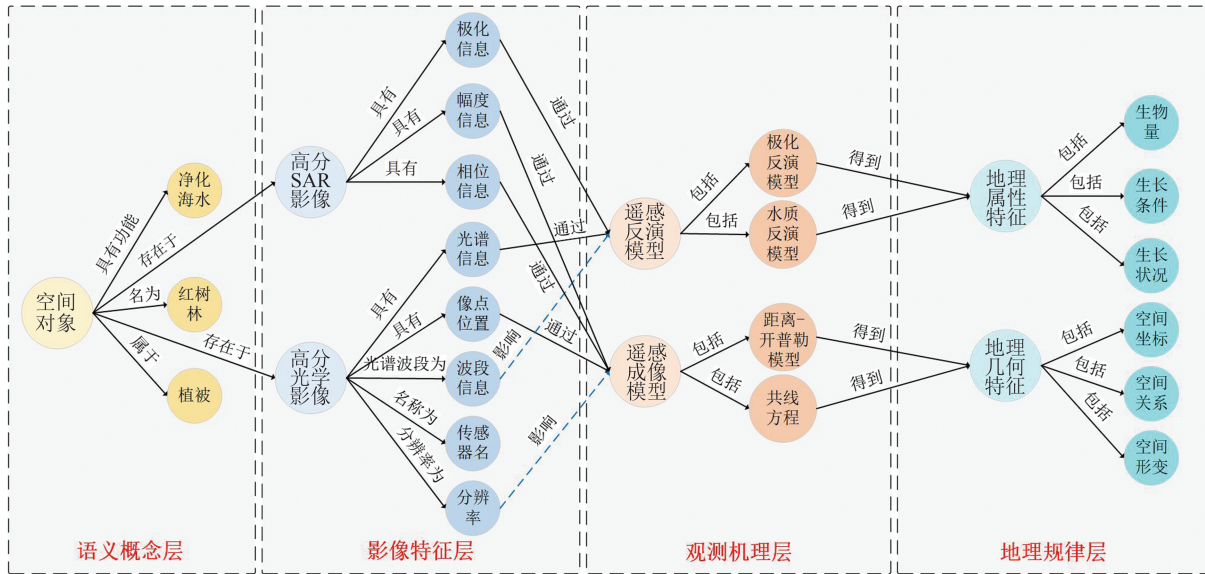


图 2 单状态遥感本体结构

Fig.2 Single-State Remote Sensing Ontology Structure

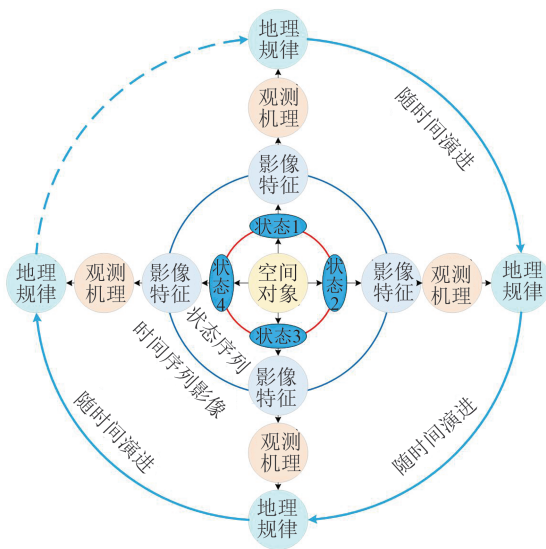


图 3 序列状态遥感本体结构

Fig.3 Sequence State Remote Sensing Ontology Structure

域内大量知识的覆盖完整度,需要结合自然语言处理、地理矢量数据解析、遥感图像目标检测等手段,从多源地学数据中进行领域知识三元组的自动化抽取。如图 5 所示,遥感知知识挖掘的数据来源主要包含结构化文本数据、非结构化文本数据、地图矢量数据和遥感影像数据 4 类地学数据。

结构化文本数据大多是领域内经过整理,采用具有明确表头信息的表格进行存储的信息。对该类数据进行知识挖掘,知识抽取主要依赖于表头信息对于表格内容在关系上的补充,即按照 <内容,表头,内容> 的形式构建三元组;对于表头信息不太明确的表格数据,文献[17]提出根据知识图谱中已有知识三元组的信息与表格内容

的对照,对表头信息的含义进行推理,从而实现从表格数据中自动化地进行知识抽取。

非结构化文本数据又可以分为具有一定结构的半结构化数据与完全非结构化数据,但核心任务都是通过自然语言处理技术从非结构化的文本中抽取领域知识。常用方法是将双向长短期记忆模型 (bi-directional long short-term memory, BiLSTM) 与卷积神经网络 (convolutional neural network, CNN) 结合^[18],基于大量标注语料库实现对于非结构化文本数据的有效建模与自主学习,能够有效捕获文本中的全局信息与局部信息,进而实现非结构化文本数据的自动知识抽取。

地图矢量数据对补充遥感知知识图谱中具体空间对象的矢量表示、空间属性等方面都具有非常大的潜力。然而,通过矢量数据进行知识三元组抽取的主要难点在于矢量数据中的地理节点与领域知识图谱中所包含的空间对象实体之间的联系不够明确,难以将二者关联起来。作为重要矢量数据来源的开源地理信息资源, OpenStreetMap (OSM)^[19] 存在大量异构的用户定义的节点表达模式,加大了数据处理的难度。为了解决这一问题,文献[20]提出的 OSM2KG (OpenStreetMap with knowledge graphs) 监督模型通过构建一种潜在的、紧凑的 OSM 节点表示方法,利用知识图谱实体与 OSM 节点间的现有关系进行训练,能够有效捕获语义节点间的相似性,发现实体与 OSM 节点间的关系,为矢量数据的知识解析提供了一个值得借鉴的思路。

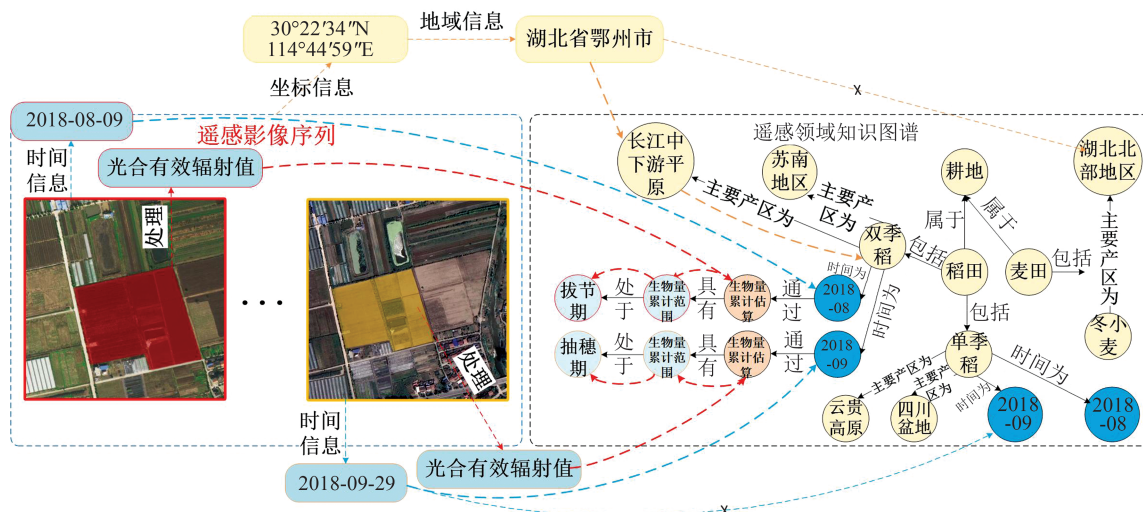


图4 时空约束下的知识图谱与遥感影像的关联解译

Fig.4 Association Interpretation of Knowledge Map and Remote Sensing Image Under the Constraints of Time and Space

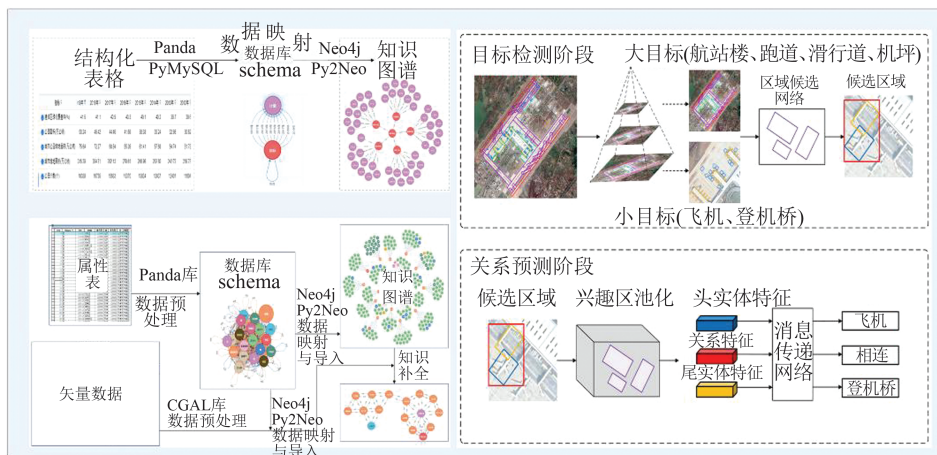


图5 基于多源地学数据的遥感领域知识三元组提取

Fig.5 Extraction of Triples of Remote Sensing Domain Knowledge Based on Multi-source Geoscience Data

由于遥感领域知识图谱最终需要服务于遥感影像解译任务,因此在进行知识抽取时需要将遥感影像数据纳入进来,进行基于遥感影像的知识三元组抽取。首先通过遥感目标检测模型(例如R3Det^[21])进行目标检测,获得贴近影像上空间对象的目标框;然后提取成对目标框的多维度特征,将多维度特征输入信息传递到神经网络中,预测空间对象之间的空间联系;最后抽取出现遥感领域知识三元组。

1.3 遥感领域知识图谱的对齐融合与关系补全

完成基于多源数据的知识三元组抽取后,考虑到经过知识抽取得到的三元组来自多源异构数据,知识图谱极大可能会存在由于知识来源不同导致的同近义实体,例如“果园”与“果木园”,或者是在一些实体间缺少内在关系等情况,需要进一步对遥感领域知识图谱的内容进行合并同

近义实体,精简图谱表达和补全缺失关系,完成知识图谱的精细化。

1.3.1 遥感领域知识图谱对齐融合

遥感知识图谱构建的初始阶段采用了基于多源数据挖掘的方法进行知识三元组抽取,因此需要对获取的多个知识图谱进行融合,对从多源数据中获取的同义实体或关系进行对齐合并,使图谱的内容更加简洁完整,进一步提高知识图谱的易用性。一般来说,对多源数据进行知识抽取得到的高质量图谱往往规模较小,能够获得的已知对齐关系的可用于训练的遥感知识图谱样本较少。因此,本文选取了在小样本条件下效果较为理想的迭代自拓展式实体对齐(bootstrapping entity alignment, BootEA)模型^[22]对遥感领域知识图谱进行实体对齐融合,该方法的基本流程如图6所示。

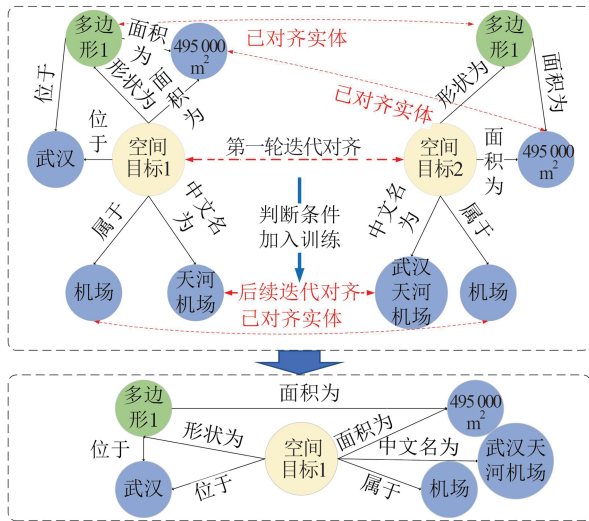


图 6 迭代自拓展式实体对齐模型基本流程
Fig.6 Flowchart of Bootstrapping Entity Alignment Model

采用迭代自拓展式实体对齐模型具有两个优势:(1)在学习知识图谱的语义映射时,对于生成的负样本,仅取正样本实体的 s 近邻实体进行负样本生成,由此生成的负样本与正样本的相似度较高,更加难以区分,通过训练能够提高模型的映射能力,解决遥感领域中对于语义相似实体的区分问题,提高对齐精度;(2)在实际的遥感领域知识图谱构建过程中,已知对齐的实体对较少,训练样本的数量匮乏会很大程度上限制模型的性能,因此在迭代训练过程中,通过一定的判别规则,将迭代过程中得到的高置信对齐实体加入训练样本中,从而提高模型在遥感领域知识图谱实体对齐任务中的表现。

迭代自拓展式实体对齐模型训练过程中的总体目标函数为:

$$O = O_e + \mu_1 O_a \quad (1)$$

$$O_e = \sum_{\tau \in T^+} [f(\tau) - \gamma_1]_+ + \mu_2 \sum_{\tau \in T^-} [\gamma_2 - f(\tau')]_+ \quad (2)$$

$$O_a = - \sum_{x \in X} \sum_{y \in Y} \phi_x(y) \log \pi(y|x; \Theta) \quad (3)$$

式中, $\mu_1, \gamma_1, \gamma_2, \mu_2$ 为超参数; O_e 为基本目标函数; O_a 为实体对齐似然度的目标函数; T^+ 为正三元组的映射向量集合; T^- 为负三元组的映射向量集合; $f(*)$ 为三元组映射向量的距离函数,对于正三元组,该函数值尽可能小,而对于负三元组,该函数值尽可能大; X, Y 为两个知识图谱中实体的映射向量集合; $\pi(y|x; \Theta)$ 为采用余弦相似度度量的两个不同知识图谱中所出现的实体的映射向量的相似度; $\phi_x(y)$ 为第一个知识图谱中实体 x

的对齐实体 y 的概率分布函数。

O_e 通过整体上对正负三元组的距离函数进行约束,加强模型对于单个知识图谱内部语义信息映射到低维稠密空间向量的能力。 O_a 将多源图谱中的同义实体尽可能映射为相同的低维稠密空间向量,从而能够通过低维向量的相似度判别实体是否为同义实体。通过以上两个约束,可在对单个图谱的内部语义信息进行合理映射的同时,将已知的对齐约束加入进来,最终得到多源子图谱间的同义实体信息。

为了验证方法有效性,本文基于已有的遥感领域知识图谱构建了两个各包含 1 595 个三元组的子图谱,两个子图谱共享一定数量的同义实体。本文将其中的 456 对实体作为训练样本,并对剩下的 458 对实体进行测试,总共训练了 50 个迭代过程,总计 500 个迭代次数。最终得到的定量结果,在链接预测中排名小于等于 1 的三元组的平均占比 (Hits@1) 为 71.83%, Hits@1 值越大越好;三元组的链接预测排名倒数的平均数 (mean reciprocal ranking, MRR) 为 75.50%, MRR 值越大越好。

1.3.2 遥感领域知识图谱关系补全

多源子图谱进行同义实体的对齐融合后,实体之间往往还会存在缺失关系的情况,因此需要对遥感领域知识图谱整体地进行关系补全操作。考虑到实际在认知知识时,遥感领域知识中会存在相当多的不同语义层级关系,例如“植被覆盖区域”和“红树林区域”之间的上下位关系,而这种语义层级关系可以有效帮助判断某些关系是否存在。为了充分利用知识图谱实体节点的语义层级约束信息,本文采用了自学习语义层次感知模型 (hierarchy-aware knowledge graph embedding, HAKE)^[23] 将知识三元组映射到极坐标,通过实体的语义层级信息判断关系存在的可能性,能够更好地适应遥感领域知识图谱的关系补全任务。自学习语义层次感知模型的流程图如图 7 所示,该模型将映射得到的低维向量分为模量和相位两个部分。模量部分旨在为属于不同语义层级的实体建模,通过模量的不同来区分不同语义层级的实体。例如,对于“植被覆盖区域”与“红树林区域”两个实体而言,“植被覆盖区域”明显位于更高的语义层级上,因此在模量部分映射时,该实体的模量会小于“红树林区域”的模量部分。相位部分则用于区分同一语义层级上的实体,同一语义层级上的实体会被映射为接近模量,但其相位部分却能够存在很大的不同,从而

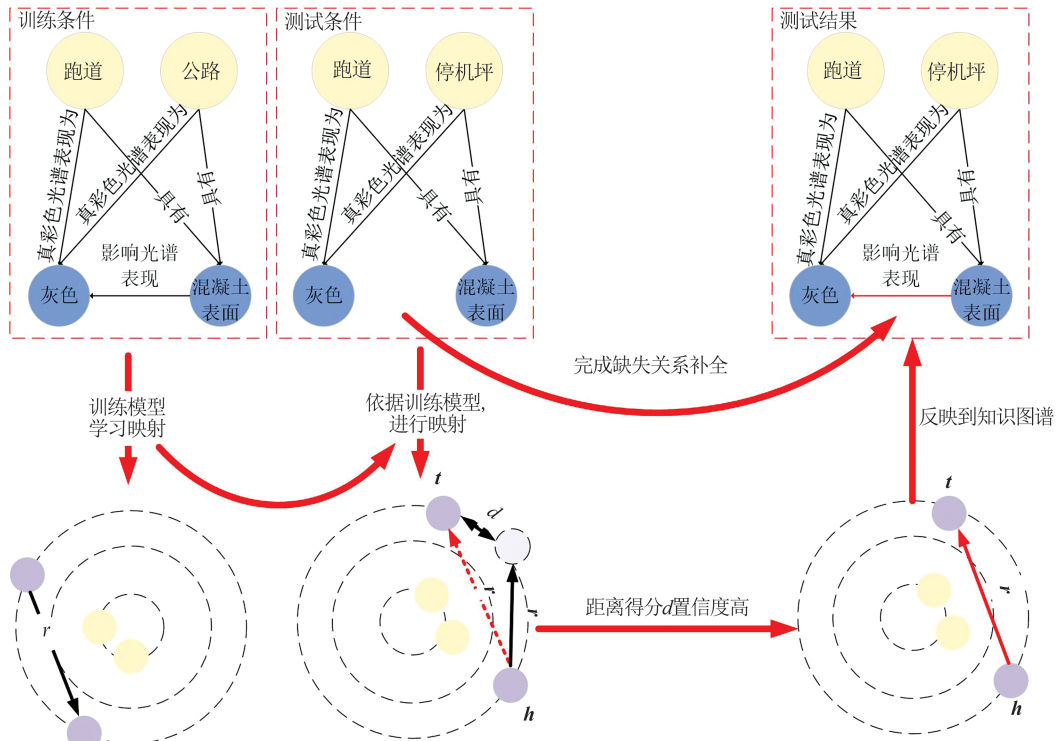


图7 自学习语义层次感知模型

Fig.7 Hierarchy-Aware Knowledge Graph Embedding Model

将其区分开来。通过模量部分与相位部分的结合,自学习语义层次感知模型将知识图谱中的实体映射为具有明显层次关系的极坐标稠密向量。

采用以上映射方式之后,由于关系在知识图谱表示学习中被映射为向量间的变换,因此对于相同语义层级实体间的关系,其模量变换部分为1,而不同语义层级实体间关系的模量变换部分不为1。因此在进行关系预测时,能够较好地利用起语义层级信息,提高关系预测的可靠性。

自学习语义层次感知模型采用的是负采样损失函数和自我对抗训练,其损失函数如下:

$$L = -\log \sigma(\gamma - d_r(h, t)) - \sum_{i=1}^n p(h'_i, r, t'_i) \log \sigma(d_r(h'_i, t'_i) - \gamma) \quad (4)$$

式中, γ 为超参数; σ 是 Sigmoid 函数; $p(h'_i, r, t'_i)$ 是采样的负三元组的概率分布函数; $d_r(h, t)$ 为三元组的距离函数。对于正确的映射向量,正样本的距离函数值应当较小,而负样本的距离函数值应当较大,因此通过上述损失函数可以对映射结果进行整体约束,训练模型的正确映射能力。

与实体对齐模型类似,本文在已构建的遥感领域知识图谱的基础上,将知识图谱以三元组为基本单位,按 8:1:1 的比例划分为训练集、验证集和测试集,关系补全的测试内容是将测试集中的

关系全部进行正确补全。实验共计训练 80 000 个迭代次数,最终在测试集上得到的定量结果 Hits@1 为 43.25%,所有三元组链接预测 MRR 为 50.00%,MRR 值越小越好。

2 耦合知识图谱和深度学习的智能遥感影像解译

在遥感领域知识图谱基础上,本文结合知识图谱的知识推理以及深度学习的大数据自主学习优势提出了耦合知识图谱和深度学习的新一代智能遥感影像解译范式。以零样本遥感影像场景分类、可解释遥感影像语义分割以及大幅面遥感影像场景图生成 3 个典型任务为例对该范式进行测试。

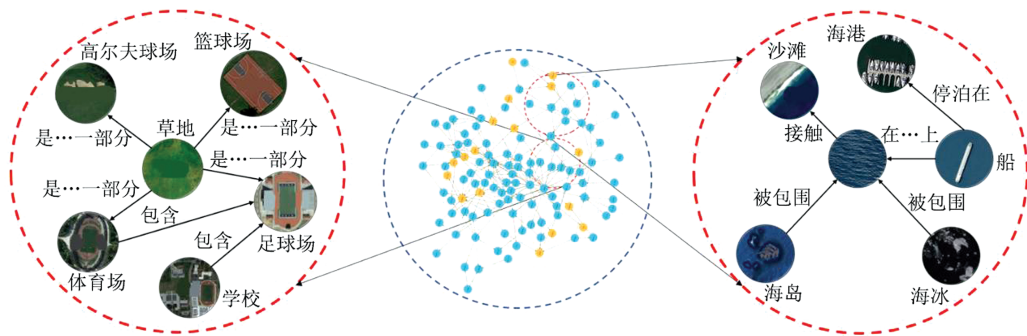
2.1 基于知识图谱表示学习的广义零样本遥感影像场景分类

遥感影像场景分类是高分辨率遥感影像实现精确解译的关键技术,具有重要的应用价值。然而在如今遥感大数据的时代背景下,遥感影像的数量以及遥感地物的类别数量都呈现出指数级的增长,仅仅依靠现有的样本库难以有效识别所有种类地物。因此,研究如何在现有样本库下实现训练阶段不可见类的识别在遥感大数据时

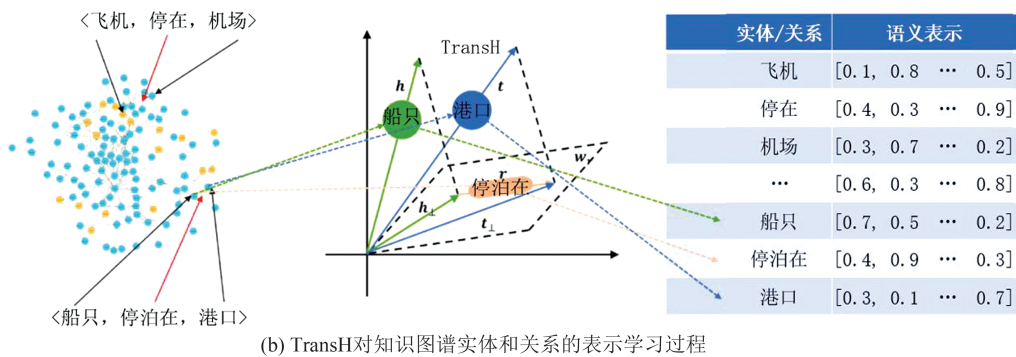
代具有重要的实际价值^[24]。

基于上述考虑,针对基于知识图谱表示学习的广义零样本遥感影像场景分类技术进行了研究^[25],探索适用于遥感领域知识图谱的构建方法,通过表示学习技术将知识图谱应用于遥感影像场景广义零样本分类,为遥感领域广义零样本分类任务提供了新视角,其流程图如图 8 所示。首先借助先验知识完成基于遥感领域专家群体的遥感领域本体建模,构建了包含五大类时空属性和非时空属性以及四大类时空关系和非时空关系的遥感知识图谱,如图 8(a)所示,其中实体和属性共有 3 870 个,关系三元组为 7 252 个。为了实现知识图谱与深度学习联合下的遥感影像分类,分别使用表示学习方法和卷积神经网络对遥感知识图谱所蕴含的语义表示信息和场景影像的视觉特征信息进行挖掘,并建立语义表示与视觉特征之间的联系,从而实现训练阶段不可见类场景的识别。在遥感知识图谱表示学习中,假设头实体的向量 h 加上关系的向量 r 应等于尾实

体的向量 t 。通过该约束,对实体和关系的向量表示进行学习,可以用于后续广义零样本遥感影像分类任务。为了解决图 8(a)中所展示的实体间一对多和多对一等复杂关系,本文采用 TransH^[26]方法将关系建模为超平面,即针对每一个关系 r ,都假设有一个对应的超平面,关系 r 位于该超平面上,其法向量为 w ,具体如图 8(b)所示。对于给定的三元组 (h, r, t) ,将实体 h 和 t 映射到关系 r 的超平面上,得到 $c_{h_{\perp}} = c_h - w^T c_h w_r, c_{t_{\perp}} = c_t - w^T c_t w_r$,其中 w_r 是超平面的法向量,对于一个正确的三元组满足 $c_{h_{\perp}} + c_r \approx c_{t_{\perp}}$ 。考虑到遥感知识图谱中的实体和标签之间存在一对一的对应关系,可获得每类遥感场景标签的语义表示。对于场景图的视觉特征,本文采用经典的残差网络进行提取,通过语义表示与视觉特征之间的匹配即可建立知识图谱与深度学习之间的联系,通过不可见类场景图的视觉特征匹配到合适的语义表示,得到最终的类别预测。



(a) 遥感知识图谱可视化结果



(b) TransH对知识图谱实体和关系的表示学习过程

图 8 遥感知识图谱表示学习

Fig.8 Remote Sensing Knowledge Graph Representation Learning

为了获得更好的不同模态间的匹配结果,本文提出了一种基于隐层空间特征对齐和多类别分布分离的深度跨模态对齐模型(deep cross-modal alignment, DCA),将视觉特征和语义特征都映射到隐层空间,在实现视觉特征和语义表示

之间匹配的同时,也增强了视觉特征和语义表示之间的耦合关系,从而有效提升广义零样本遥感场景分类的精度。模型网络结构如图 9 所示,具体参数设置可参考文献[25]。首先,利用最小化视觉特征和语义表示的重构损失,然后对齐视觉

特征和语义表示在隐层空间的分布,最后进一步分离隐层空间中不同类别特征的分布。考虑到视觉特征和语义表示在隐层空间的对齐以及多类别分布的分离,该模型的总体损失 \mathcal{L} 定义为:

$$\mathcal{L} = \mathcal{L}_{VAE} + \alpha\mathcal{L}_{CMFR} + \beta\mathcal{L}_{VSDM} + \gamma\mathcal{L}_{MCDD} \quad (5)$$

式中, \mathcal{L}_{VAE} 表示视觉特征和语义表征重构损失,VAE(variational autoencoder)表示变分自编码器; \mathcal{L}_{CMFR} 和 α 分别表示跨模态重构特征(cross-

modal feature reconstruction, CMFR)损失及其权重; \mathcal{L}_{VSDM} 和 β 分别表示视觉特征和语义分布对齐(visual and semantic distribution matching, VSDM)损失及其权重; \mathcal{L}_{MCDD} 和 γ 分别表示多类别间特征分布分离(multi-category distribution dispersion, MCDD)损失及其权重。该模型不仅适用于广义零样本遥感影像分类,同样也适用于零样本遥感影像分类。

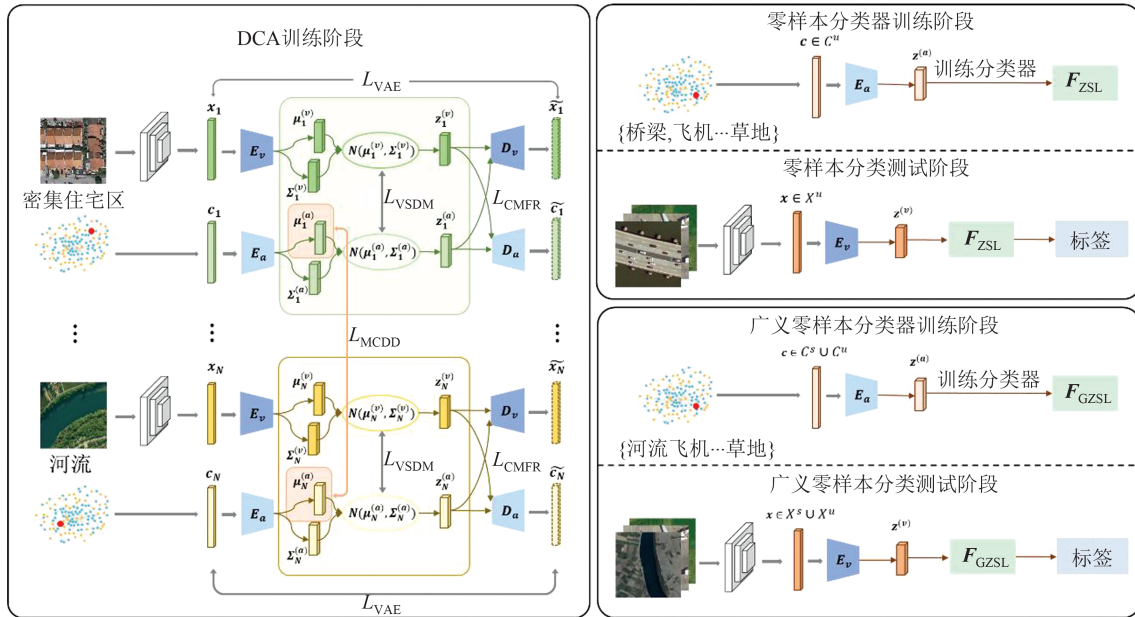


图9 深度跨模态匹配模型整体框架

Fig.9 Framework of DCA

为了定量地比较所提方法的性能,本文融合了多个现有数据集得到一个大规模遥感影像场景数据集^[25],新的数据集共包含遥感影像场景70类,每类包含800张影像,每张影像均为256×256像素,在实验中对可见类/不可见类分别采取了60/10、50/20和40/30 3种划分方式,其分类结果如表1所示,最优值加粗表示。由表1可以发现,本文方法在不同的可见类/不可见类比例和不同的语义表示下,都明显优于其他方法。值得注意的是,对比同一方法中使用的不同语义表示,遥感知识图谱在大多数情况下性能最好。这证明了基于遥感知识图谱表示学习获得的语义表示优于自然语言模型提取的语义表示,能够有效地提升分类的精度,提高了深度学习分类模型的可解释性。

2.2 联合知识推理和深度学习的可解释遥感影像语义分割

遥感影像语义分割是遥感信息处理领域的研究热点,在自然资源监测、智慧农业、生态评估

等应用中发挥着重要作用^[27-31]。然而,基于深度学习的遥感影像语义分割方法可解释性差,且模型性能严重依赖于受训练样本数量。遥感知识图谱包含了丰富的语义关系信息和强大的推理能力,能够在深度学习模型遥感影像语义分割的基础上进一步增强模型的可解释性,提高分割结果的可信度。

基于上述考虑,本文提出了一种联合知识推理和深度学习的可解释遥感影像语义分割方法,推进遥感影像的智能化解译,总体流程如图10所示。该方法采用深度学习网络从遥感影像中自主学习低层次特征,知识图谱语义推理则运用高层次的专家语义知识引导解译纠错与推理。深度学习与知识推理有机耦合在一个闭环的回路中,利用知识推理对深度学习结果进行纠正,知识推理产生的输出辅助深度学习模型的训练,二者可以彼此迭代增强,直到分类精度收敛,有助于提高遥感解译结果的精度以及可解释性。

首先,建立了遥感本体以及基于遥感本体的

表 1 广义零样本分类任务中不同划分方式下不同方法的准确率对比/%

Tab.1 Accuracy Comparison of Different Methods Under Different Partition Modes in Generalized Zero-Shot Classification Task/%

语义表示	可见类/ 不可见类	SAE ^[27]	DMaP ^[28]	CI2SL ^[29]	CADA-VAE ^[30]	DCA
Word2Vec	60/10	27.97±1.13	28.88±1.26	25.18±0.86	32.88±2.54	34.09±1.34
	50/20	20.99±1.90	20.33±1.13	15.70±0.86	30.25±3.07	31.44±1.66
	40/30	17.15±0.55	16.78±1.10	9.10±1.32	26.06±0.79	25.63±0.26
BERT	60/10	28.57±0.94	26.57±0.65	25.00±1.25	36.34±2.03	37.96±1.65
	50/20	21.52±1.38	19.52±1.42	14.95±1.51	31.51±2.27	31.45±1.85
	40/30	16.65±0.40	16.31±1.24	8.57±0.57	27.05±0.79	28.15±1.16
Attribute	60/10	28.58±0.93	30.71±0.78	23.88±0.87	36.00±2.19	37.60±1.24
	50/20	20.52±1.75	23.55±0.87	14.27±1.05	32.17±2.41	32.66±0.80
	40/30	16.73±1.06	16.12±0.82	8.11±0.98	26.13±0.79	28.79±0.92
知识图谱	60/10	28.86±0.60	30.11±1.39	23.65±0.61	38.10±1.89	40.25±0.84
	50/20	23.66±1.06	23.41±1.21	13.93±1.01	32.94±1.42	34.11±0.45
	40/30	16.94±1.03	16.20±1.62	8.14±0.87	28.11±0.79	29.61±0.82

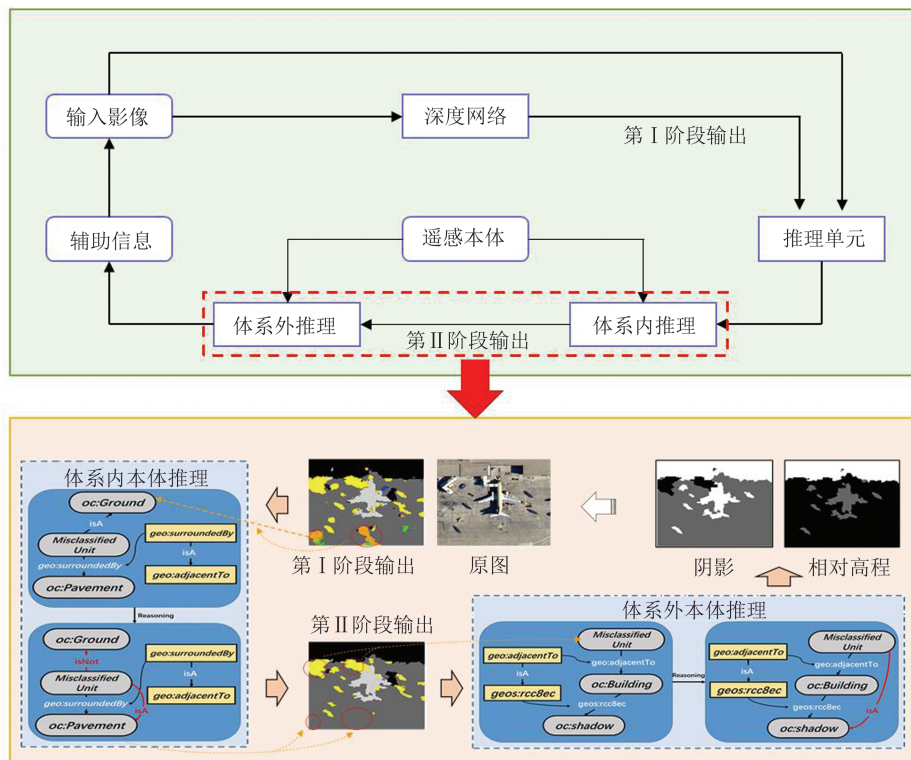


图 10 可解释性遥感影像语义分割方法的总体流程图

Fig.10 Overall Flowchart of Explainable Remote Sensing Image Semantic Segmentation

推理规则,并将所有规则都加入到推理规则库中,其中遥感知识图谱语义推理包括体系内和体系外推理,完整的推理规则可参考文献[32]。体系内图谱推理根据推理规则直接纠正深度学习网络输出的错误分割结果,体系外图谱推理利用改正后的结果提取阴影和相对高程信息,并将其作为辅助信息与原始影像一同输入到深度学习

网络中进行下一轮语义分割,从而增强网络的泛化能力。在每一次迭代中,首先单独训练用于遥感影像语义分割的深度学习模型,使用体系内图谱推理规则直接对模型的输出结果中的错分结果进行纠正,得到第I阶段的输出。然后,使用超像素分割方法进行超像素分割,通过超像素聚类得到推理单元,再根据知识图谱体系内推理规则

纠正分类错误的推理单元从而得到第Ⅱ阶段的输出,即改正后的语义分割结果。最后,根据知识图谱体系外推理规则对改正后的结果进行推理,提取阴影和相对高程信息,将其附加到原始遥感影像中输入深度学习模型中进行下一次迭代,直到模型收敛为止。

为了验证该方法的有效性,本文将Potsdam数据集重新划分为水体、建筑物、车辆等6类常见土地覆盖类型,并从该数据集所包含的38张尺寸为6 000×6 000像素的遥感影像中裁剪出多张512×512像素的图像,将这些裁剪得到的图像按6:2:2的比例随机划分为训练集、验证集和测试集,分类结果如表2所示。由表2可以发现,知识推理与深度学习的联合方法取得了最好的分类结果,并且相比于第Ⅰ阶段,经过了体系内推理改正分类错误的第Ⅱ阶段的分类精度提高明显。实验结果充分证明了将知识推理融入深度卷积神经网络的分割方法对地物分类的有效性,该方法能有效提升分类精度和合理性。

表2 在Potsdam数据集上的分类结果/%

Tab.2 Classification Results on the Potsdam Dataset /%

方法	总体精度	平均交并比
U-Net ^[33]	81.29	64.44
Semantic Referee ^[12]	82.76	66.69
知识推理与深度学习的联合方法(第Ⅰ阶段)	84.58	67.64
知识推理与深度学习的联合方法(第Ⅱ阶段)	85.51	68.93

2.3 知识图谱引导的大幅面遥感影像场景图生成

随着深度学习技术的推广应用,大量遥感影

像解译任务都得到了发展^[34]。然而,传统解译任务缺乏对影像地物间的关系的认知^[35]。场景图生成源于计算机视觉领域,是更高层次的图像理解任务。基于场景图可以进一步提升图像智能理解水平,直接作用于图像检索、图像自然语言描述以及图像视觉问答等下游任务。场景图数据结构中的节点和边分别表示图像中的目标和关系,根据输入图像自动生成目标检测结果以及一系列描述图像场景的目标关系三元组,形成场景图。

完整的遥感场景通常是大幅面的,遥感目标通常具有方向性、尺度变化大、分布密集等数据特点。因此,区别于计算机视觉领域的场景图,定义遥感场景图需要以大幅面完整的遥感场景影像为底图,通过旋转目标框描述场景目标。结合这些特性,亟需研究大幅面遥感影像场景图生成方法。

一般来说,场景图生成任务首先进行目标检测定位分类目标,再对成对的目标进行关系预测。然而,大幅面遥感影像中包含的目标数量众多,如果不加以处理,直接预测所有目标对之间的关系将造成大量的计算资源损耗。遥感领域知识图谱是丰富的知识库,包含了地物目标及其关系的一般规律事实,而遥感场景图可以看作遥感领域知识图谱的实例。因此可以通过遥感领域知识图谱的先验知识优化遥感场景图的关系生成,知识图谱引导的大幅面遥感影像场景图生成总体流程如图11所示。首先对大幅面遥感影像进行旋转目标检测;然后在关系预测阶段,利用知识图谱引导进行目标关系搜索优选,以提高计算效率;最后基于知识图谱先验知识对关系预测置信度进行修正,以提高预测精度。

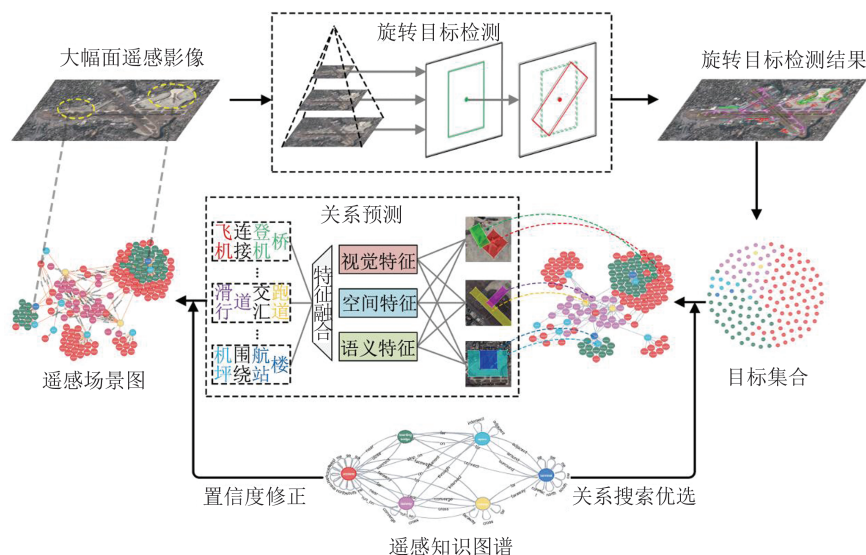


图11 知识图谱引导的大幅面遥感影像场景图生成流程

Fig.11 Flowchart of Knowledge Graph-Guided Scene Map Generation for Large-Size Remote Sensing Images

对大幅面遥感影像进行旋转目标检测,可利用旋转目标检测网络如 R3Det^[21]得到大幅面遥感影像的目标检测结果。记旋转目标类别集合为 $O=\{O_1, O_2 \cdots O_n\}$, 目标框集合为 $B=\{B_1, B_2 \cdots B_n\}$ 。对于目标集合通过知识图谱先验知识,对目标进行优选,搜索出存在潜在关系的目标对。用于目标对搜索的规则集合可表示为:

$$\{((O_i, B_i), (O_j, B_j)) | \text{hasRelation}(O_i, O_j, \text{KG}) \wedge (\text{Dist}(B_i, B_j) < \tau) \vee (\text{Intersect}(B_i, B_j) > 0)\} \quad (6)$$

式中, $((O_i, B_i), (O_j, B_j))$ 为搜索出的目标对; KG 表示遥感领域知识图谱 (knowledge graph); $\text{hasRelation}(O_i, O_j, \text{KG})$ 表示判断类别 O_i 和 O_j 在 KG 中是否存在关系; $\text{Dist}()$ 表示计算目标之间的中心距离; τ 为距离阈值; $\text{Intersect}()$ 表示判断目标之间是否相交。

对于搜索出的目标对 $((O_i, B_i), (O_j, B_j))$, 根据目标信息以及对应图像内容, 提取出目标对联合的目标视觉特征 F_{visual} 、目标框空间特征 F_{spatial} 、目标类别语义特征 F_{semantic} , 用以表征目标关系。其中, F_{visual} 为经过卷积神经网络后的影像深度特征图在目标对联合目标框范围的区域池化结果;

F_{spatial} 为目标对之间的目标框交并比、目标中心距离以及目标中心连线方向角的归一化值; F_{semantic} 为经过语言模型 (如 Word2vec) 提取的目标类别名词向量。通过融合上述 3 种特征, 将 $P=\sigma([F_{\text{visual}}, F_{\text{spatial}}, F_{\text{semantic}}], \Gamma)$ 作为目标间关系预测结果, 其中, Γ 为模型参数, $\sigma()$ 为激活过程。在遥感知识图谱引导下, 遥感影像目标间关系预测可以表示为:

$$R = \Phi(P, F) \quad (7)$$

式中, F 为知识图谱中先验目标关系统计概率; $\Phi()$ 为关系预测置信度修正过程, 可通过加权求和实现。

利用大幅面的机场场景图数据集进行遥感场景图生成实验, 实验数据集共包括 30 张大幅面影像, 共包含目标实例 6 137 个、关系实例 15 332 个, 将数据集按照 6:2:2 的比例随机划分为训练集、验证集和测试集, 实验结果如表 3 所示, 最优值加粗表示, 其中 $R@K$ 表示计算了置信度最高的前 K 个三元组召回率, 是场景图生成的常用指标, $K=500, 1\ 000, 1\ 500$ 。由表 3 可知, 知识图谱引导的方法精度明显高于基准的频率统计方法, 并且在知识图谱引导的置信度修正下, 方法精度比基本的多特征方法高。

表 3 遥感场景图生成方法的精度对比结果/%

Tab.3 Accuracy Comparison Results of Remote Sensing Scene Map Generation Methods /%

方法	关系分类			场景图分类			场景图生成		
	R@500	R@1 000	R@1 500	R@500	R@1 000	R@1 500	R@500	R@1 000	R@1 500
频率统计 ^[36]	32.46	47.53	55.44	22.81	34.09	40.75	10.84	12.71	13.03
多特征方法 ^[37]	41.41	54.83	59.77	24.09	38.05	43.72	13.39	15.77	16.03
知识图谱引导方法	44.66	56.41	60.37	25.66	38.80	44.78	13.48	15.80	16.06

3 结 语

本文首先设计了顾及遥感成像机理和地理学知识的遥感本体模式, 并在此基础上, 讨论了基于数据挖掘的多源遥感信息领域相关知识自动化提取方法。为了突破遥感领域知识集成难这一难题, 研制了遥感领域知识图谱的协同构建与自主进化系统。此外, 在遥感领域知识图谱基础上, 本文从广义零样本遥感影像场景分类、遥感影像语义分割以及大幅面遥感影像场景图生成等 3 个典型案例出发, 对本文提出的耦合知识图谱和深度学习的新一代智能遥感影像解译范式进行了初步探索。大量实验结果表明, 遥感知识图谱的引入能够有效提高深度学习模型的可

解释性、计算效率、准确率等。值得说明的是, 无论是遥感领域知识图谱建模方法, 还是知识图谱与深度学习的耦合理论方法都尚处于萌芽阶段, 需要大量研究工作来推进。

参 考 文 献

- [1] Li Deren, Zhang Liangpei, Xia Guisong. Automatic Analysis and Mining of Remote Sensing Big Data [J]. *Acta Geodaetica et Cartographica Sinica*, 2014, 43(12): 1211-1216 (李德仁, 张良培, 夏桂松. 遥感大数据自动分析与数据挖掘[J]. 测绘学报, 2014, 43(12): 1211-1216)
- [2] Fu Kun, Sun Xian, Qiu Xiaolan, et al. Multi-Satellite Integrated Processing and Analysis Method Under Remote Sensing Big Data [J]. *National Remote*

- Sensing Bulletin*, 2021, 25(3): 691-707 (付琨, 孙显, 仇晓兰, 等. 遥感大数据条件下多星一体化处理与分析[J]. 遥感学报, 2021, 25(3): 691-707)
- [3] Pan Zhuokun, Hu Yueming, Wang Guangxing, et al. Cognitions and Perspectives in Remote Sensing of Urban Renewal Monitoring [J]. *Remote Sensing Technology and Application*, 2020, 35(4): 911-923 (潘灼坤, 胡月明, 王广兴, 等. 对遥感在城市更新监测应用中的认知和思考[J]. 遥感技术与应用, 2020, 35(4): 911-923)
- [4] McCarthy J, Minsky M, Rochester N, et al. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955 [J]. *AI Mag*, 2006, 27: 12-14
- [5] Rosenblatt F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain [J]. *Psychological Review*, 1958, 65(6): 386-408
- [6] Zhang Bo, Zhu Jun, Su Hang. Toward the Third Generation of Artificial Intelligence [J]. *Scientia Sinica (Informationis)*, 2020, 50(9): 1281-1302 (张钹, 朱军, 苏航. 迈向第三代人工智能[J]. 中国科学: 信息科学, 2020, 50(9): 1281-1302)
- [7] Goodenough D G, Goldberg M, Plunkett G, et al. An Expert System for Remote Sensing [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1987, GE-25(3): 349-359
- [8] Matsuyama T. Knowledge-Based Aerial Image Understanding Systems and Expert Systems for Image Processing [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1987, GE-25(3): 305-316
- [9] Xu Jun, Pei Tao, Yao Yonghui. Conceptual Framework and Representation of Geographic Knowledge Map [J]. *Journal of Geo - Information Science*, 2010, 12(4): 496-502 (许珺, 裴韬, 姚永慧. 地学知识图谱的定义、内涵和表达方式的探讨[J]. 地球信息科学学报, 2010, 12(4): 496-502)
- [10] Wang Zhihua, Yang Xiaomei, Zhou Chenghu. Geographic Knowledge Graph for Remote Sensing Big Data [J]. *Journal of Geo - Information Science*, 2021, 23(1): 16-28 (王志华, 杨晓梅, 周成虎. 面向遥感大数据的地学知识图谱构想[J]. 地球信息科学学报, 2021, 23(1): 16-28)
- [11] Zhang Xueying, Zhang Chunju, Wu Mingguang, et al. Spatiotemporal Features Based Geographical Knowledge Graph Construction [J]. *Scientia Sinica (Informationis)*, 2020, 50(7): 1019-1032 (张雪英, 张春菊, 吴明光, 等. 顾及时空特征的地理知识图谱构建方法[J]. 中国科学: 信息科学, 2020, 50(7): 1019-1032)
- [12] Alirezaie M, Långkvist M, Sioutis M, et al. Semantic Referee: A Neural-Symbolic Framework for Enhancing Geospatial Semantic Segmentation [J]. *Semantic Web*, 2019, 10(5): 863-880
- [13] Mikolov T, Sutskever I, Chen K, et al. Distributed Representations of Words and Phrases and Their Compositionality [J]. *CoRR*, 2013, DOI: 1310.4546
- [14] Bordes A, Usunier N, Garcia-Duran A, et al. Translating Embeddings for Modeling Multi-relational Data [C]// Neural Information Processing Systems (NIPS), Carson City, USA, 2013
- [15] Chen Y W, Zhang S L, Peng X, et al. A Collaborative Ontology Construction Tool with Conflicts Detection [C]//The 4th International Conference on Semantics, Knowledge and Grid, Beijing, China, 2008
- [16] Du Qingyun, Ren Fu. Representation Model of Spatial Information in Natural Language [J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(6): 682-688 (杜清运, 任福. 空间信息的自然语言表达模型[J]. 武汉大学学报·信息科学版, 2014, 39(6): 682-688)
- [17] Chen J Y, Jimenez-Ruiz E, Horrocks I, et al. Learning Semantic Annotations for Tabular Data [C]//The 28th International Joint Conference on Artificial Intelligence, Macao, China, 2019
- [18] Chen T, Xu R F, He Y L, et al. Improving Sentiment Analysis via Sentence Type Classification Using BiLSTM-CRF and CNN [J]. *Expert Systems with Applications*, 2017, 72: 221-230
- [19] Haklay M, Weber P. OpenStreetMap: User-Generated Street Maps [J]. *IEEE Pervasive Computing*, 2008, 7(4): 12-18
- [20] Tempelmeier N, Demidova E. Linking OpenStreetMap with Knowledge Graphs—Link Discovery for Schema-Agnostic Volunteered Geographic Information [J]. *Future Generation Computer Systems*, 2021, 116: 349-364
- [21] Yang X, Liu Q Q, Yan J C, et al. R3Det: Refined Single-Stage Detector with Feature Refinement for Rotating Object [J]. *arXiv*, 2019, DOI: 1908.05612
- [22] Sun Z Q, Hu W, Zhang Q H, et al. Bootstrapping Entity Alignment with Knowledge Graph Embedding [C]//The 27th International Joint Conference on Artificial Intelligence, Stockholm, Sweden, 2018
- [23] Zhang Z Q, Cai J Y, Zhang Y D, et al. Learning Hierarchy-Aware Knowledge Graph Embeddings for Link Prediction [C]//The AAAI Conference on Artificial Intelligence, New York, USA, 2020
- [24] Li Yansheng, Kong Deyu, Zhang Yongjun, et al.

- Zero-Shot Remote Sensing Image Scene Classification Based on Robust Cross-Domain Mapping and Gradual Refinement of Semantic Space [J]. *Acta Geodaetica et Cartographica Sinica*, 2020, 49(12): 1564-1574 (李彦胜, 孔德宇, 张永军, 等. 联合稳健跨域映射和渐进语义基准修正的零样本遥感影像场景分类[J]. 测绘学报, 2020, 49(12): 1564-1574)
- [25] Li Y S, Kong D Y, Zhang Y J, et al. Robust Deep Alignment Network with Remote Sensing Knowledge Graph for Zero-Shot and Generalized Zero-Shot Remote Sensing Image Scene Classification[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 179: 145-158
- [26] Wang Z, Zhang J, Feng J, et al. Knowledge Graph Embedding by Translating on Hyperplanes [C]// AAAI Conference on Artificial Intelligence, Québec City, Canada, 2014
- [27] Kodirov E, Xiang T, Gong S G. Semantic Autoencoder for Zero-Shot Learning[C]//IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017
- [28] Li Y N, Wang D H, Hu H H, et al. Zero-Shot Recognition Using Dual Visual-Semantic Mapping Paths [C]// IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017
- [29] Elhoseiny M, Elfeki M. Creativity Inspired Zero-Shot Learning[C]// IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019
- [30] Schönfeld E, Ebrahimi S, Sinha S, et al. Generalized Zero and Few-Shot Learning via Aligned Variational Autoencoders [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019
- [31] Tong X Y, Xia G S, Lu Q K, et al. Land-Cover Classification with High-Resolution Remote Sensing Images Using Transferable Deep Models [J]. *Remote Sensing of Environment*, 2020, 237: 111322
- [32] Li Y S, Ouyang S, Zhang Y J. Collaboratively Boosting Data-Driven Deep Learning and Knowledge-Guided Ontological Reasoning for Semantic Segmentation of Remote Sensing Imagery [J]. *arXiv*, 2020, DOI: 2010.02451
- [33] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[M]//Cham: Springer International Publishing, 2015
- [34] Ma L, Liu Y, Zhang X L, et al. Deep Learning in Remote Sensing Applications: A Meta-Analysis and Review[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2019, 152: 166-177
- [35] Li Deren. Brain Cognition and Spatial Cognition: On Integration of Geo-spatial Big Data and Artificial Intelligence [J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(12): 1761-1767 (李德仁. 脑认知与空间认知: 论空间大数据与人工智能的集成[J]. 武汉大学学报·信息科学版, 2018, 43(12): 1761-1767)
- [36] Zellers R, Yatskar M, Thomson S, et al. Neural Motifs: Scene Graph Parsing with Global Context [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018
- [37] Zhu Y H, Jiang S Q, Li X Y. Visual Relationship Detection with Object Spatial Distribution [C]// IEEE International Conference on Multimedia and Expo, Hong Kong, China, 2017

A New Paradigm of Remote Sensing Image Interpretation by Coupling Knowledge Graph and Deep Learning

LI Yansheng¹ ZHANG Yongjun¹

¹ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

Abstract: Objectives: In the remote sensing (RS) big data era, intelligent interpretation of remote sensing images (RSI) is the key technology to mine the value of big RS data and promote several important applications. Traditional knowledge-driven RS interpretation methods, represented by expert systems, are highly interpretable, but generally show poor performance due to the interpretation knowledge being difficult to be completely and accurately expressed. With the development of deep learning in computer vision and other

fields, it has gradually become the mainstream technology of RSI interpretation. However, the deep learning technique still has some fatal flaws in the RS field, such as poor interpretability and weak generalization ability. In order to overcome these problems, how to effectively combine knowledge inference and data learning has become an important research trend in the field of RS big data intelligent processing. Generally, knowledge inference relies on a strong domain knowledge base, but the research on RS knowledge graph (RS-KG) is very scarce and there is no available large-scale KG database for RSI interpretation now. **Methods:** To overcome the above considerations, this paper focuses on the construction and evolution of the RS-KG for RSI interpretation and establishes the RS-KG takes into account the RS imaging mechanism and geographic knowledge. Supported by KG in the RS field, this paper takes three typical RSI interpretation tasks, namely, zero-shot RSI scene classification, interpretable RSI semantic segmentation, and large-scale RSI scene graph generation, as examples, to discuss the performance of the novel generation RSI interpretation paradigm which couples KG and deep learning. **Results and Conclusions:** A large number of experimental results show that the combination of RS-KG inference and deep learning can effectively improve the performance of RSI interpretation. The introduction of RS-KG can effectively improve the interpretation accuracy, generalization ability, anti-interference ability, and interpretability of deep learning models. These advantages make RS-KG promising in the novel generation RSI interpretation paradigm.

Key words: data learning and knowledge inference; remote sensing knowledge graph; deep learning; remote sensing images intelligent interpretation

First author: LI Yansheng, PhD, associator professor, majors in intelligent mining of remote sensing big data. E-mail: yansheng.li@whu.edu.cn

Corresponding author: ZHANG Yongjun, PhD, professor. E-mail: zhangyj@whu.edu.cn

Foundation support: The National Natural Science Foundation of China(42030102, 41971284); Innovative Research Groups of the Natural Science Foundation of Hubei Province (2020CFA003).

引文格式: LI Yansheng, ZHANG Yongjun. A New Paradigm of Remote Sensing Image Interpretation by Coupling Knowledge Graph and Deep Learning[J]. Geomatics and Information Science of Wuhan University, 2022, 47(8):1176-1190. DOI:10.13203/j.whugis20210652(李彦胜, 张永军. 耦合知识图谱和深度学习的新一代遥感影像解译范式[J]. 武汉大学学报·信息科学版, 2022, 47(8):1176-1190. DOI:10.13203/j.whugis20210652)