

基于建筑信息模型数据驱动的铁铁路设备运维多模态知识图谱构建

林海香^{1,2}, 胡娜娜¹, 何乔³, 赵正祥¹, 白万胜¹

(1. 兰州交通大学 自动化与电气工程学院, 甘肃 兰州 730070; 2. 兰州交通大学 四电建筑信息模型工程与智能应用铁路行业重点实验室, 甘肃 兰州 730070; 3. 贝尔福-蒙贝利亚技术大学 信息学院, 贝尔福 法国 90000)

摘要: 铁路信号设备是保障行车安全、提高运输效率的核心装备, 加强信号设备智能运维是降低铁路运行风险的必要基础保障。目前, 针对我国基于建筑信息模型(BIM)的智能运维平台存在不能精准映射各设备的行为规律和相互之间互馈作用的机理, 须同时依靠经验知识进行推断等问题。首先构建了铁路设备运维文本知识图谱; 其次构建卷积神经网络(CNN)-团组图卷积神经网络(cgGCN)模型对BIM图像模态数据进行处理, 完成对20种铁路信号设备零件图信息的标注, 实验结果表明模型准确率达到95.38%, 精确率和召回率的调和平均值F1达到95.58%; 最后将BIM图像信息以视觉模态嵌入运维文本知识图谱, 利用Neo4j图数据库实现多模态知识图谱可视化展示, 从而精准映射各信号设备相互之间互馈作用的机理, 为后续现场铁路运维人员实施安全管理和运维决策提供在线服务和指导。

关键词: 铁路信号设备; 建筑信息模型(BIM); 运维; 多模态; 知识图谱

中图分类号: U284

文献标志码: A

Construction of a Multi-Modal Knowledge Graph for Railway Equipment Operation and Maintenance Based on Building Information Model Data-Driven Approach

LIN Haixiang^{1,2}, HU Nana¹, HE Qiao³, ZHAO Zhengxiang¹, BAI Wansheng¹

(1. School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China; 2. Key Laboratory of Railway Industry of Building Information Model Engineering and Intelligence for Electric Power, Traction Power Supply, Communication and Signaling, Lanzhou Jiaotong University, Lanzhou 730070, China; 3. School of Information, University of Technology of Belfort-Montbéliard, Belfort 90000, France)

收稿日期: 2023-10-24

基金项目: 国家自然科学基金重点项目(52038008); 甘肃省重点研发计划: 工业类(23YFGA0046); 四电BIM工程与智能应用铁路行业重点实验室2022年度开放课题(BIMKF-2022-02)

第一作者: 林海香, 副教授, 工学博士, 主要研究方向为交通信息数据挖掘。E-mail: linhaixiang@mail.lzjtu.cn

通信作者: 何乔, 工学硕士, 主要研究方向为交通控制信息分析。E-mail: qiao.he@utbm.fr

Abstract: Railway signal equipment is essential for ensuring traffic safety and improving transportation efficiency. Strengthening the intelligence operation and maintenance of signal equipment is essential to mitigate the risks associated with railway operations. Currently, the intelligence operation and maintenance platform based on building information model (BIM) in China is unable to accurately depict the behavior and mutual feedback mechanism of each device, thus relying on experiential knowledge for inference. To address this issue, initially, the knowledge graph was constructed using the text related to the operation and maintenance of railway equipment; Subsequently, a convolutional neural networks-clique group graph convolutional neural networks (CNN-cgGCN) model was developed to process BIM image modal data and annotate the information of 20 specific railway signal equipment part drawings. The experimental results show that the accuracy of the model reaches 95.38%, and the harmonic mean F1 of precision and recall reaches 95.58%; Finally, BIM image information is integrated into the visual knowledge graph of operation and maintenance text. This multi-modal knowledge graph is then visualized using the Neo4j graph database, so as to accurately map the mechanism of mutual feedback between signal equipment, and offer online services and guidance to on-site railway operation and maintenance personnel, facilitating safety management and operational decision-making.

Keywords: railway signal equipment; building information model (BIM); operation and maintenance; multi-modal; knowledge graph

随着我国高速铁路网的快速发展^[1], 高速铁路



论文
拓展
介绍

运营现状呈大运量、高速度、高密度等特点。铁路信号设备作为高速铁路安全可靠运营的核心设备,具有结构复杂、各设备内部构件之间的关联性强等特点,传统以现场人员密集型作业为主的运维模式逐渐难以适应现有复杂运维场景下的故障高效分析与处理。近年来,建筑信息模型(building information model, BIM)在铁路行业得到了广泛应用。在运维管理阶段, BIM主要用于支持运维工作流程的自动化,如铁路四电BIM工程设备管理系统构建了朔黄铁路全生命周期管理的基础设施智能运维技术体系^[2],有效增强了运维信息交互和模数共享能力。尽管BIM为信号设备运维信息管理提供了媒介和技术,通过建模获得的表征数据可反映设备实体之间的结构特征,但不能精准映射其行为规律和相互之间互馈作用的机理,尤其是当获得的表征数据不够完备、不够精细、不够准确、不够及时的时候,须同时依靠模型知识进行推断^[3]。因此,如何将BIM可视化三维图像数据和专家经验知识及运维文本数据有机集成融合,实现数据的多维呈现,对提高铁路信号设备运营维护效率具有十分重要的意义。

知识图谱(knowledge graph, KG)作为表示结构化知识的描述框架,能够以简洁的形式表达描述性数据和实例数据。目前已被广泛应用于电力^[4]、航空^[5]、医疗^[6]、金融^[7]等各大领域。在铁路领域,文献[8-9]以铁路信号设备相关文本数据为研究对象,研究相关知识抽取、知识融合技术,为基于知识图谱的各类应用提供数据支撑,但并未将这些知识进行高效的组织管理以服务于高速铁路故障诊断、智能运维等下游应用。文献[10]构建面向列车车载设备的故障知识图谱,为车载设备故障智能维护提供支持。文献[11]使用基于知识图谱的拓扑分析来探索铁路运营事故。文献[12]以英国铁路事故报告为研究对象,建立面向铁路安全事故的知识图谱进而应用于铁路风险评估。上述研究均为单一文本信息来源的单模态数据驱动的方法,对铁路数据中可挖掘的信息量存在一定局限性,并不能满足铁路设备运维高质量需求。多模态知识图谱(multi-modal knowledge graph, MMKG)可以在传统知识图谱的基础上嵌入多种模态的实体^[13],并构建多模态实体间的语义关系,实现数据-模型-知识的集成融合与可视化。例如Xiong等^[14]针对甲骨文信息处理构建了多模态知识图谱,该MMKG能够更好地组织和管理基础数据,为甲骨文信息处理研究服务。Fan等^[15]通过本体驱动的方式,构建了集音频、文本、图

像为一体的中国戏曲多模态知识图谱,为用户提供多维交互。上述研究说明,通过构建多模态图谱可实现多源异构数据的高效组织管理与应用,且能取得比单模态图谱更优的效果。

综上,本文从实际应用出发,提出基于BIM和数据驱动的铁路设备运维多模态知识图谱构建方法。首先利用已有的铁路信号设备运维文本数据,构建运维文本知识图谱;接着以铁路四电BIM工程设备管理系统获取的图像模态数据为研究对象,构建卷积神经网络(convolutional neural networks, CNN)-团组图卷积神经网络(clique group graph convolutional neural networks, cgGCN)模型对BIM图像数据进行信息标注;最后将图像信息以视觉模态嵌入到运维文本知识图谱中,完成BIM图像数据与运维文本数据的集成融合,构建铁路设备运维多模态知识图谱,为后续现场铁路运维人员实施安全管理和运维决策提供在线服务和指导。

1 铁路设备运维多模态知识图谱构建总体框架

本文提出的铁路信号设备运维多模态知识图谱构建总体框架如图1所示,主要包括运维文本知识图谱构建、铁路设备BIM图像模态知识抽取以及多模态知识图谱构建及可视化展示与应用。

1.1 文本知识图谱构建

由于铁路信号设备运维数据具有强专业性,故采用自顶向下与自底向上相结合的方法构建知识图谱。首先根据不同类型数据的结构化特征,自顶向下形成对应的概念节点和关联关系,然后自底向上构建相应的数据层,利用知识源之间的数据冗余性进行知识抽取,文本模态的知识抽取操作主要包括实体、关系抽取和知识融合。

首先,针对铁路信号设备非结构化运维文本数据命名实体识别任务识别精度低、识别不全问题,采用BERT-BiLSTM-CRF (bidirectional encoder representations from transformers-bi-directional long short-term memory-conditional random fields)模型进行实体抽取:由BERT将文本数据转为字符级嵌入向量,送入BiLSTM提取上下文特征信息,CRF学习标签之间的内部关系以获取最优标签序列。然后,基于BERT-CNN模型进行关系抽取,将文本数据送入BERT中进行编码生成词向量,然后送入CNN中进行训练,最后添加全连接层和softmax层

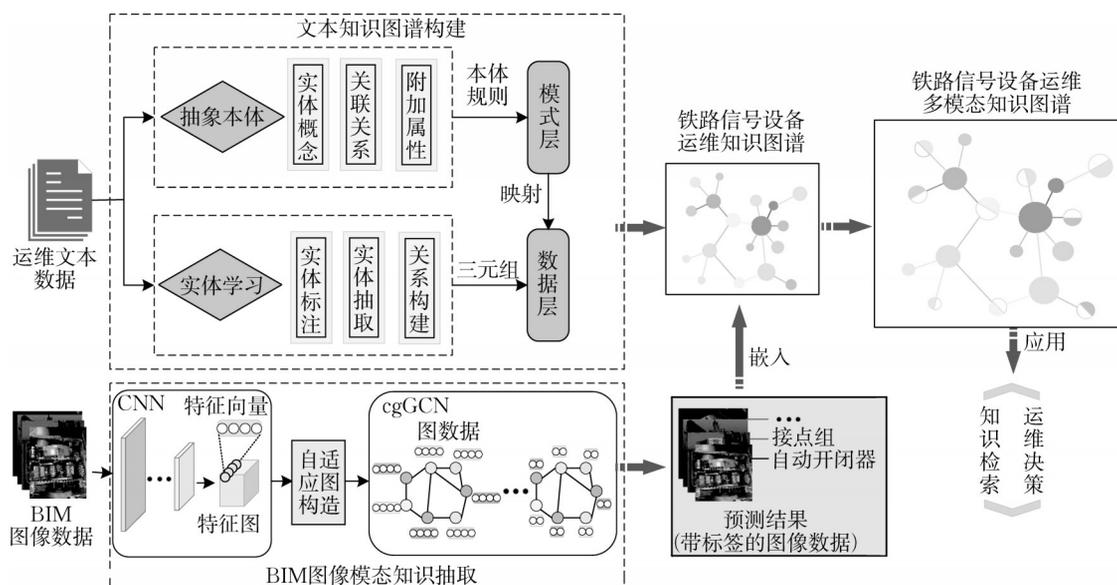


图1 铁路设备运维多模态知识图谱构建总体框架

Fig. 1 Overall framework of multimodal knowledge graph construction for railway equipment operation and maintenance

抽取出目标实体之间的关系。由于运维文本数据记录方式存在差异,对同一设备、故障现象、处理措施等内容的描述不同,导致知识重复抽取,例如“密贴检查器”和“密检器”实际为同一设备实体。为避免知识图谱中构建重复节点,本文基于铁路信号设备运维数据中的实体构建词典,针对抽取实体采用 cosine 相似度算法在词典中进行实体匹配,将抽取实体链接到对应实体上。

最后,将非结构化的源文本数据转化为结构化的知识三元组数据并存储至 Neo4j 图数据库,完成运维文本知识图谱的构建。

1.2 图像模态知识抽取

铁路领域存在诸多专有名词,在实际运维环境中,基于BIM的铁路信号设备运维平台可对信号设备进行三维几何信息的可视化。将此类BIM图像模态数据表征和关联在知识图谱中,能更直观形象地显示铁路设备内部的上下游关系,实现故障发生后的快速定位。对于BIM图像模态数据,本文借助计算机视觉处理技术进行表征,其流程如下所述。

1.2.1 自适应图的构造

为了从输入的铁路设备BIM图像中自动学习具有代表性的图数据,将原始图像送入CNN进行自适应图的构造,过程如图2所示,具体工作机理如下。

首先给定输入的铁路设备图像 $I \in R^{H \times W \times 3}$, H 和 W 分别为图像高度和宽度,采用CNN网络对其进行高维特征的学习,由卷积层和池化层生成多通

道的特征图 $I_m \in R^{H' \times W' \times C}$ ($H' \ll H$ 和 $W' \ll W$), H' 和 W' 分别为特征图的高度和宽度, C 为通道数。接着将多通道特征图中相同空间位置的像素拼接起来,形成特征向量,该向量被视为一个图节点,代表原始图像中的一个空间区域。最后将多通道特征映射调整为节点嵌入矩阵 $X \in R^{N \times C}$, 其中 N 为节点数。基于节点嵌入,由 k 近邻法 (k -nearest neighbor, k NN) 根据节点特征之间的距离生成边来构造图。

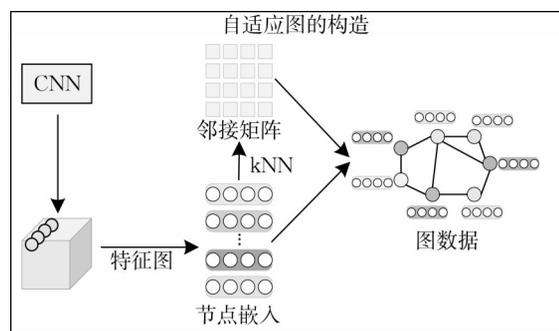


图2 自适应图的构造过程

Fig. 2 Process of adaptive graph construction

k NN 中评价两个节点之间相似度的距离函数 $D(\cdot, \cdot)$, 计算方式为

$$D(v_i, v_j) = v_i - v_j \quad (1)$$

其中: v_i 和 v_j 分别为第 i 和第 j 个节点嵌入, $i, j \in \{1, 2, \dots, N\}$ 。然后,选择离每个节点最近的 k 个节点作为相邻节点。

图的边可以计算为

$$e_{ij} = \begin{cases} 1, & \text{若 } v_j \in \text{kNN}(v_i) \\ 0, & \text{否则} \end{cases} \quad (2)$$

其中:如果图中第*i*和第*j*个节点之间有边连接,则为 $e_{ij}=1$, $e_{ij}=0$ 则为无边连接。

由于自适应图的构造嵌入到整个图像分类模型框架中,因此在训练过程中,节点特征和边缘连接都会随着整个网络的优化而更新。因此,图可以自适应更新和构造,随后图卷积网络可以学习更有效的图表示进行分类。

1.2.2 cgGCN

对构建的图数据进行图卷积,进一步学习铁路设备图像的空间特征。采用cgGCN替代传统的GCN,以捕捉节点间强大的空间相关性,学习有效的空间特征用于分类任务。

(1) cGCN

cGCN通过循环反馈网络(Clique Net)最大化信息流,进而有效重用高维图像特征,引导分层特征的细化。具体过程如图3所示。cGCN有三个图卷积层,任意两层之间都采用反馈和反向连接。特征传播过程可分为两个阶段,即阶段I和阶段II。

在阶段I中,输入层通过单向连接初始化所有层。更新后的层被串联起来以更新下一层。cGCN正向传播中第*i*层的图卷积可以表示为

$$X_i^{(1)} = \sum_{l < i} (H(X_l^{(1)}, W_{li})) + H(X_0, W_{oi}) \quad (3)$$

其中: X_0 为cGCN的输入; $X_i^{(1)}$ 为阶段I中第*i*个图卷积层的输出; W_{ij} 为连接第*i*层和第*j*层的权值矩阵。

在阶段II中,各层开始交替更新。除了顶层待更新外,所有的图卷积层都作为底层串联,对应的参数也进行串联。cGCN的反馈过程可以表示为

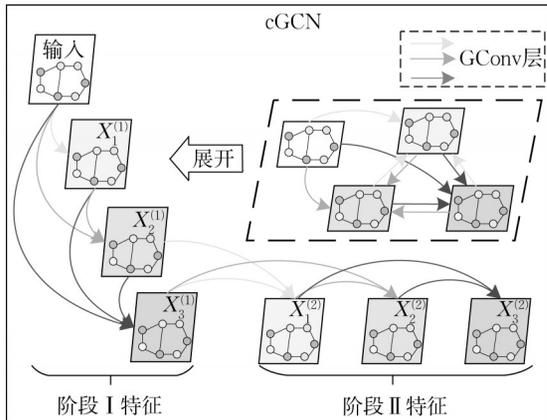


图3 cGCN结构示意图

Fig. 3 Illustration of clique graph convolution structure

$$X_i^{(2)} = \sum_{l < i} H(X_l^{(2)}, W_{li}) + \sum_{m > i} H(X_m^{(1)}, W_{mi}) \quad (4)$$

其中: $X_i^{(2)}$ 表示阶段II中第*i*个图卷积层的输出。

具有*f*维节点特征向量的cGCN输出图表示 $Y_{\text{out}} \in R^{N \times F}$ 可以表示为

$$Y_{\text{out}} = \{X_3^{(1)}, X_3^{(2)}\} \quad (5)$$

其中“{}”表示将向量进行拼接。

(2) gGCN

利用gGCN来减少冗余信息,并在相邻层之间的特征传播过程中进一步隐式地选择更优的特征表示。其结构如图4所示。cgGCN中gGCN的工作原理描述如下。

首先, G_{input} 表示gGCN的输入,将其分成两个子图 G_1 和 G_2 ,分别经过图卷积运算后产生线性图特征 G_1' 和 G_2' ,计算过程如式(6)所示:

$$\begin{cases} G_1' = G\text{Conv}(G_1) + b_1 \\ G_2' = G\text{Conv}(G_2) + b_2 \end{cases} \quad (6)$$

其中: b_1 和 b_2 分别是图卷积运算过程对应的偏置。

然后,用得到的 G_1' 和 G_2' 计算经BN+ReLU模块产生的非线性图特征 G_1'' 和 G_2'' ,计算过程如式(7)所示:

$$\begin{cases} G_1'' = \text{ReLU}(\text{BN}(G_1')) \\ G_2'' = \text{ReLU}(\text{BN}(G_2')) \end{cases} \quad (7)$$

式中:BN和ReLU分别表示批量归一化和ReLU激活函数。

接着,由GConv(Group Conv, 分组卷积)和Tanh激活函数结合成门控机制,输出 G_{1g}' 和 G_{2g}' ,计算过程如下:

$$\begin{cases} G_{1g}' = \text{Tanh}(\text{BN}(G\text{Conv}(G_1')) + b_{1g}) \\ G_{2g}' = \text{Tanh}(\text{BN}(G\text{Conv}(G_2')) + b_{2g}) \end{cases} \quad (8)$$

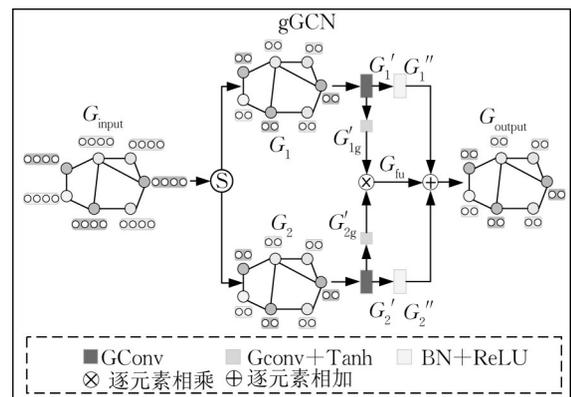


图4 gGCN结构示意图

Fig. 4 Illustration of group graph convolution structure

式(8)中: b_{1g} 、 b_{2g} 分别为图卷积运算过程对应的偏置。计算融合特征 G_{fu} 为

$$G_{fu} = G'_{1g} \cdot G'_{2g} \quad (9)$$

其中：“ \cdot ”表示元素的乘积。逐元素相乘可避免维数增加,降低后续计算的复杂性,同时可有效捕捉两个子图之间的复杂特征交互,以学习高级特征表示。

最后,gGCN的输出图表示 G_{output} 计算为

$$G_{output} = \alpha G''_1 + \beta G''_2 + \gamma G_{fu} \quad (10)$$

其中: α 、 β 和 γ 是可训练的权重变量。

gGCN可以简化学习过程,从而细化节点特征的学习,增强图表示固有的多样性,使学习特性更具鲁棒性。此外,gGCN还实现了对偶分支中不同信息的选择和融合,生成丰富的图表示。

(3) 图像标签预测

将cGCN和gGCN学习到的特征送入图池化层,通过全连接层对输出进行整合,随后采用softmax实现图像标签预测。

2 实验与评估

2.1 实验数据集及评价指标

铁路设备图像数据集来源于铁路四电工程BIM设备管理系统,共计2 135张,如图5所示为信号机BIM图像。为避免模型出现过拟合问题,对图像进行旋转、水平翻转等数据增强操作,形成包含4 270张图像的数据集,以8:1:1的比例随机划分为训练集、验证集和测试集,通过EasyData智能标注工具对训练集数据进行标注。模型性能评价指标^[16]为准确率、精确率、召回率以及精确率和召回率的调和平均值F1。



图5 BIM图像示例

Fig. 5 Example of BIM image

2.2 实验环境搭建

本文实验环境如表1所示,经多次实验确定的

部分参数如表2所示。

表1 实验环境及配置

Tab. 1 Experimental environment and configuration

实验环境	环境配置
操作系统	Linux操作系统
CPU	Intel i9 12900K
GPU	NVIDIA GeForce RTX3080Ti
编程语言	Python3.8
编程环境	pycharm2021.2
深度学习框架	TensorFlow-GPU-1.14.0

表2 实验参数设置

Tab. 2 Setting of experimental parameters

名称	参数值
学习率	0.5
权值衰减	5×10^{-4}
批处理大小	64
Epoch	200
最优邻节点数 k	8

2.3 对比实验与分析

为验证本实验所提出的模型在铁路设备图像数据集上分类的有效性,将CNN-cgGCN与ResNet50^[17]、CNN-SAGEConv^[18]、CNN-GATConv^[19]、CNN-tGCN^[20]、CNN-cGCN和CNN-gGCN进行比较。由于cgGCN输入的图是基于自适应图构造模块构建的,为便于比较,将所有基于图卷积的算法都嵌入到CNN-GCN框架中,网络层数与cgGCN相同。实验结果如图6所示。

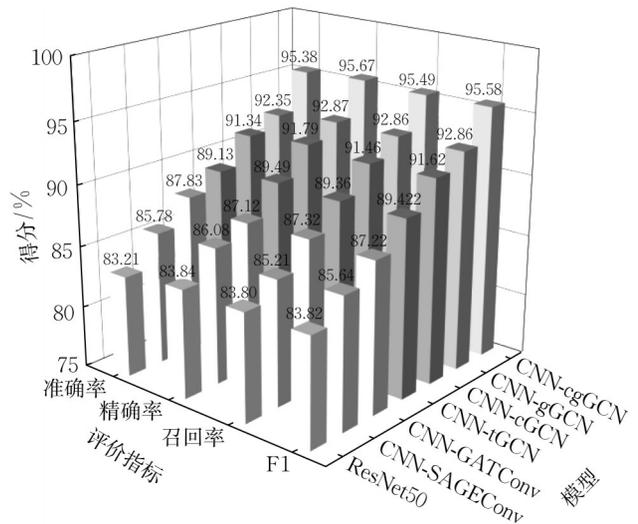


图6 不同模型图像分类结果

Fig. 6 Image classification results of different models

实验结果表明,CNN-cgGCN算法优于ResNet50、CNN-SAGEConv、CNN-GATConv、

CNN-tGCN、CNN-cGCN 和 CNN-gGCN 等所有比较算法。下面对实验结果进行具体分析:

(1) 基于 CNN-GCN 框架的所有算法在所有指标上的性能都优于基线模型 ResNet50,说明 CNN-GCN 框架的有效性,因为图卷积网络能够从相应的铁路信号设备 BIM 图像中学习各构件间的空间信息,从而提高分类性能。

(2) 与 CNN-tGCN 相比,CNN-cGCN 在分类准确率、精确率、召回率和 F1 值上分别提高了 2.34%、2.30%、2.10% 和 2.33%,表明 cGCN 中的循环反馈机制能够有效提高图卷积的性能。与此同时,CNN-gGCN 的结果优于 CNN-tGCN,因为 gGCN 将原始图划分为两个子组,然后分别提取各分支特征和融合特征,获取不同的图信息,同时还隐式地执行特征选择进而减少冗余信息,提高图卷积的分类性能。

(3) 与 CNN-cGCN 相比,CNN-cgGCN 分类准确率提高了 4.04%,表明 cgGCN 中的 gGCN 可以通过图卷积简化学习过程,从而细化节点特征,使学习的特性更具鲁棒性。此外,gGCN 还可以选择和融合各分支中的不同信息,生成丰富的图表示,提高

分类性能;与 CNN-gGCN 相比 CNN-cgGCN 模型分类准确率提高了 3.03%,表明 cgGCN 中的 cGCN 通过循环反馈机制能够最大化信息流,获得更优的特征表示,提高分类性能。综上,CNN-cgGCN 分类性能最好,证明将 cGCN 和 gGCN 结合用于铁路设备 BIM 图像处理的优越性。

3 多模态知识图谱可视化及应用

基于所述方法,首先抽取铁路设备运维文本中的知识,以三元组形式存入 Neo4j 图数据库,实现铁路设备运维文本知识图谱的可视化展示。接着通过 BIM 图像模态知识抽取获取对应信号机、自动开闭器、齿轮组、电动机、锁闭器等 20 种铁路设备零件图信息,然后令图像信息以视觉模态嵌入到运维文本知识图谱中,直观准确地反映信号设备信息之间的内在联系。另外,考虑在不同环境下各设备工作的可靠性和安全性,将转辙机、信号机等设备零部件实体节点所处的自然环境以及应用环境等相关信息作为附加属性添加。部分铁路设备运维多模态知识图谱可视化展示如图 7 所示。

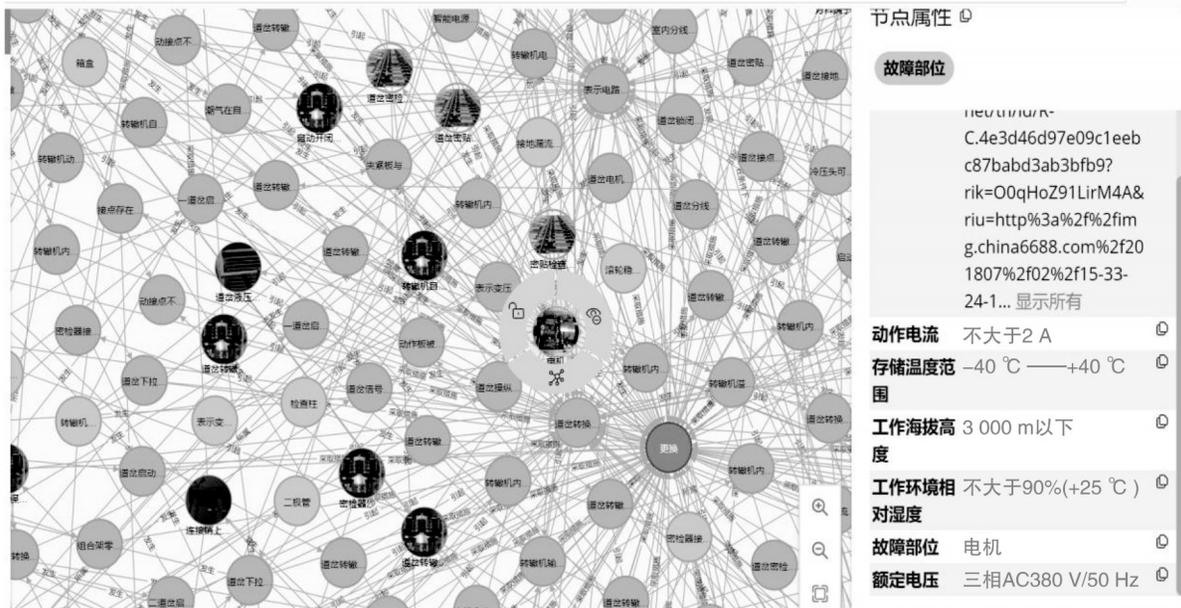


图7 铁路设备运维多模态知识图谱

Fig. 7 Multimodal knowledge graph for operation and maintenance of railway equipment

铁路设备运维多模态知识图谱的决策辅助功能包括查询分析和运维决策措施反馈。通过知识检索和图计算实现,知识检索能够在知识图谱中快速获取运维决策者所需的信息,从而形成结构化的知识结果并反馈给运维决策者。图计算利用图分析算法

对知识图谱进行分析,获得更深层次的设备运维决策规则。

为进一步对比多模态数据为铁路信号设备运维图谱带来的跨模态检索和查询能力,将本文构建的多模态知识图谱与去除 BIM 图像模态的知识图谱进

行对比。表3为不同模态知识图谱在运维过程中的故障处置表现对比。

表3 不同模态图谱对比

Tab. 3 Comparison of different modal knowledge graphs

图谱模态	构建耗时/min	查询耗时/ms	故障处置耗时/min
文本	6	345	25
文本-图像	10	351	20

由表3可看出,尽管构建多模态图谱将花费更多的时间,但在单次知识检索耗时方面并没有太大差距。在实际场景中,由于知识图谱通常是预先构建完成,并不对信号设备故障处置任务的执行造成较大影响。而从完成故障处置任务耗时方面来看,多模态知识图谱将起到较大辅助作用,其跨模态检索能力可实现故障部位的快速定位,有效缩短故障维修耗时,提升运维决策准确率。

如图8,以包含多模态数据进行知识检索和运维决策措施反馈为例,介绍运维多模态知识图谱的实际应用流程。在运营过程中信号设备发生故障后,发出包含图像和文本模态数据的报警信息,文本信息为“6号道岔尖3转辙机故障”,经知识抽取得到关键实体{故障部位:转辙机},对图像信息先定位目标对象后提取图像特征,与铁路设备运维MMKG中的图像实体进行相似度计算,确定图像对应的实体名称为“自动开闭器”,由此可快速将故障部位定位到转辙机内部的自动开闭器,同时采用Neo4j Cypher在MMKG中进行查询,返回自动开闭器对应相关实体的关联属性信息,得到其对应的历史故障记录及维修措施等运维知识。最后,运维人员进行故障维修时,选用检索推送出的历史运维决策措施作为维修依据,并结合现场状况进行维修决策,提高运维决策效率和维修效果。

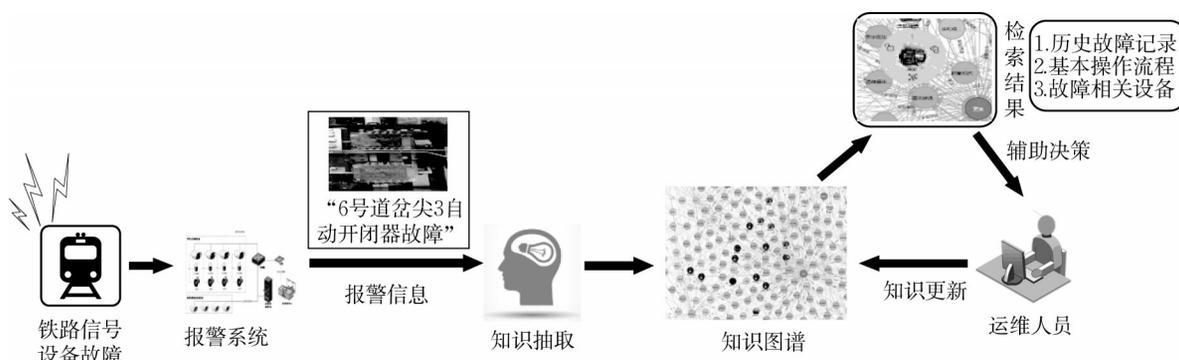


图8 铁路设备运维多模态知识图谱的应用流程

Fig. 8 Application process of multimodal knowledge graph of railway equipment operation and maintenance

4 结论

(1) 基于非结构化运维文本数据,首先构建BERT-BiLSTM-CRF模型进行实体抽取,由于抽取实体中存在“多词一义”问题,故经余弦相似度算法进行实体消歧,然后构建BERT-CNN模型抽取实体间的关系。最终将源数据转化为知识三元组数据构建运维文本知识图谱;

(2) 以铁路设备BIM图像模态的数据为对象,采用CNN-cgGCN模型对铁路设备BIM图像进行分类,得到20种铁路信号设备零件图类别,经对比实验分析,说明将该模型用于铁路领域图像模态数据知识抽取的有效性;

(3) 将BIM图像信息以视觉模态嵌入到运维文

本知识图谱中,最终构建铁路信号设备运维多模态知识图谱并进行可视化展示,完成BIM图像数据与运维文本知识的集成融合,直观地映射各设备的行为规律和相互之间互馈作用的机理,同时通过知识检索和图计算可快速获取运维决策者所需的信息,辅助运维人员进行运维决策。

本文是从自然语言处理角度出发,通过构建铁路信号设备运维多模态知识图谱实现BIM图像模态数据与运维文本数据的多维呈现,实质是知识图谱的补充。后续在本文所构建的多模态图谱的基础上会添加更多如音频、视频等模态的数据进行扩充,并且会对多模态信息融合进行研究,将尝试更多先进的方法提高图谱构建质量。在图谱应用方面,未来将结合图神经网络以探究运维图谱自动补全生成的可行性。

作者贡献声明:

林海香:数据收集,方案设计,建模思路。
 胡娜娜:知识图谱构建,初稿撰写。
 何 乔:数据分析,论文修改。
 赵正祥:实验结果可视化。
 白万胜:数据整理。

参考文献:

- [1] 叶玉玲,李文卿,张俊. 高速铁路网络复杂特性及其传播动力学研究[J]. 同济大学学报(自然科学版), 2019, 47(5): 655.
 YE Yuling, LI Wengqing, ZHANG Jun. Research on the complex characteristics and propagation dynamics of high speed railway networks[J]. Journal of Tongji University (Natural Science), 2019, 47(5): 655.
- [2] 韩自力,马战国,谷牧,等. 重载铁路线路设备智能运维关键技术研究[J]. 铁道建筑, 2022, 62(8): 23.
 HAN Zili, MA Zhanguo, GU Mu, *et al.* Research on key technologies of intelligent operation and maintenance of heavy haul railway line equipment[J]. Railway Engineering, 2022, 62(8): 23.
- [3] 朱庆,张利国,丁雨淋,等. 从实景三维建模到数字孪生建模[J]. 测绘学报, 2022, 51(6): 1040.
 ZHU Qing, ZHANG Liguo, DING Yulin, *et al.* From real 3D modeling to digital twin modeling[J]. Journal of Surveying and Mapping, 2022, 51(6): 1040.
- [4] XIAO N, PENG B, LI X, *et al.* Research on the construction and implementation of power grid fault handling knowledge graphs[J]. Energy Reports, 2023, 9: 182.
- [5] 吴闯,张亮,唐希浪,等. 航空发动机润滑系统故障知识图谱构建及应用[J/OL]. 北京航空航天大学学报, 1-14. [2023-06-06]. <https://doi.org/10.13700/j.bh.1001-5965.2022.0434>.
 WU Chuang, ZHANG Liang, TANG Xilang, *et al.* Construction and application of fault knowledge graph for aviation engine lubrication system [J/OL]. Journal of Beijing University of Aeronautics and Astronautics, 1-14. [2023-06-06]. <https://doi.org/10.13700/j.bh.1001-5965.2022.0434>.
- [6] AYOUB H, MARYEM R, MOUNIA M, *et al.* BERT based clinical knowledge extraction for biomedical knowledge graph construction and analysis[J]. Computer Methods and Programs in Biomedicine Update, 2021, 1: 100042.
- [7] MAO X, SUN H, ZhU X, *et al.* Financial fraud detection using the related-party transaction knowledge graph [J]. Procedia Computer Science, 2022, 199: 733.
- [8] 林海香,白万胜,陆人杰,等. 基于BMBC模型的高速铁路道岔故障信息实体识别[J]. 铁道科学与工程学报, 2023, 20(4): 1149.
 LIN Haixiang, BAI Wansheng, LU Renjie, *et al.* Entity recognition of high-speed railway turnout fault information based on BMBC model[J]. Journal of Railway Science and Engineering, 2023, 20(4): 1149.
- [9] LIU C, YANG S. A text mining-based approach for understanding Chinese railway incidents caused by electromagnetic interference [J]. Engineering Applications of Artificial Intelligence, 2023, 117: 105598.
- [10] 薛莲,姚新文,郑启明,等. 高铁列控车载设备故障知识图谱构建方法研究[J]. 铁道科学与工程学报, 2023, 20(1): 34.
 XUE Lian, YAO Xinwen, ZHENG Qiming, *et al.* Research on the construction method of fault knowledge graph of CTCS on-board equipment[J]. Journal of Railway Science and Engineering, 2023, 20(1): 34.
- [11] LIU J, SCHMID F, LI K, *et al.* A knowledge graph-based approach for exploring railway operational accidents[J]. Reliability Engineering & System Safety, 2021, 207: 107352.
- [12] LIU C, YANG S. Using text mining to establish knowledge graph from accident/incident reports in risk assessment [J]. Expert Systems with Applications, 2022, 207: 117991.
- [13] 陈焯,周刚,卢记仓. 多模态知识图谱构建与应用研究综述[J]. 计算机应用研究, 2021, 38(12): 3535.
 CHEN Ye, ZHOU Gang, LU Jicang. A Review on the construction and application of multimodal knowledge graph[J]. Application Research of Computers, 2021, 38(12): 3535.
- [14] XIONG J, LIU G, LIU Y, *et al.* Oracle Bone Inscriptions information processing based on multi-modal knowledge graph[J]. Computers & Electrical Engineering, 2021, 92: 107173.
- [15] FAN T, WANG H, HODEL T. Multimodal knowledge graph construction of Chinese traditional operas and sentiment and genre recognition[J]. Journal of Cultural Heritage, 2023, 62: 32.
- [16] GAO F, LI F, WANG Z F, *et al.* Research on multilevel classification of high-speed railway signal equipment fault based on text mining[J]. Journal of Electrical and Computer Engineering, 2021(2): 1.
- [17] LI C, POLLING M, CAO L, *et al.* Analysis of automatic image classification methods for Urticaceae pollen classification [J]. Neurocomputing, 2023, 522: 181.
- [18] BURKHOLZ R, QUACKENBUSH J, BOJAR D. Using graph convolutional neural networks to learn a representation for glycans [J]. Cell Reports, 2021, 35(11): 109251.
- [19] CORIZZO R, SLENN T. Distributed node classification with graph attention networks [C]// 2022 IEEE International Conference on Big Data (Big Data). Osaka, Japan: IEEE Xplore, 2022: 3720-3725.
- [20] CHEN B, GUO W, TANG R, *et al.* TGCN: Tag graph convolutional network for tag-aware recommendation [C]// Proceedings of the 29th ACM International Conference on Information & Knowledge Management. [S.l.]: ACM Digital Library, 2020: 155-164.