



# 计算机科学

COMPUTER SCIENCE

## 基于深度学习的羽毛球知识图谱补全模型构建

陈玉珏, 胡赫, 李强

引用本文

陈玉珏, 胡赫, 李强. [基于深度学习的羽毛球知识图谱补全模型构建](#)[J]. 计算机科学, 2023, 50(11A): 220900205-6.

CHEN Yujue, HU He, LI Qiang. [Construction of Badminton Knowledge Graph Completion Model Based on Deep Learning](#) [J]. Computer Science, 2023, 50(11A): 220900205-6.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

### [基于边缘引导的多尺度医学影像分割方法](#)

Medical Image Segmentation Based on Multi-scale Edge Guidance

计算机科学, 2023, 50(11A): 220900059-7. <https://doi.org/10.11896/jsjx.220900059>

### [一种融合CNN和Swin Transformer的医学显微图像分割模型](#)

Medical Microscopic Image Segmentation Model Based on CNN Structure and Swin Transformer

计算机科学, 2023, 50(11A): 230200119-8. <https://doi.org/10.11896/jsjx.230200119>

### [基于语义注意力的医学图像超分辨率方法](#)

Medical Image Super-resolution Method Based on Semantic Attention

计算机科学, 2023, 50(11A): 221200107-6. <https://doi.org/10.11896/jsjx.221200107>

### [一种基于因果推理的垃圾分类方法](#)

Novel Method for Trash Classification Based on Causal Inference

计算机科学, 2023, 50(11A): 220800218-6. <https://doi.org/10.11896/jsjx.220800218>

### [接诉即办智能派单业务调度算法研究](#)

Study on Scheduling Algorithm of Intelligent Order Dispatching

计算机科学, 2023, 50(11A): 230300029-7. <https://doi.org/10.11896/jsjx.230300029>

# 基于深度学习的羽毛球知识图谱补全模型构建

陈玉珏<sup>1</sup> 胡赫<sup>2</sup> 李强<sup>3</sup>

1 湖南师范大学体育学院 长沙 410081

2 西安科技大学计算机学院 西安 710000

3 青海师范大学体育学院 西宁 810009

(cyj295297@163.com)

**摘要** 为提升知识图谱在羽毛球领域的应用价值,首先对补全模型的研究现状进行了分析,其次结合深度学习技术和注意力机制,基于图卷积神经网络构建子图结构解耦的知识图谱补全模型,最后对模型的改进性能进行评估。结果表明,所提模型在所有子数据集都取得了良好的结果,与最佳基线模型相当;在实验中选择3个数据集上,两个测试指标都有不同程度的降低,这表明了实体特征解耦的有效性;只需3个或8个基底就足以表达模型中不同关系的特征。本研究得到了改进效果良好的知识图谱补全模型,为知识图谱在羽毛球领域的推广奠定了基础。

**关键词** 知识图谱补全;深度学习;羽毛球;图卷积神经网络;注意力机制

**中图分类号** TP181

## Construction of Badminton Knowledge Graph Completion Model Based on Deep Learning

CHEN Yujue<sup>1</sup>, HU He<sup>2</sup> and LI Qiang<sup>3</sup>

1 School of Physical Education, Hunan Normal University, Changsha 410081, China

2 School of Computer Science, Xi'an University of Science and Technology, Xi'an 710000, China

3 School of Physical Education, Qinghai Normal University, Xining 810009, China

**Abstract** To enhance the application value of knowledge graph in badminton field, this research first analyzes the research status of the completion model, then combines the deep learning technology and attention mechanism to build a knowledge graph completion model based on graph convolution neural network with subgraph structure decoupling, and finally evaluates the improved performance of the model. The results show that the proposed model has achieved good results in all sub datasets, which is equivalent to the best baseline model. On the three data sets selected in the experiment, the two test indicators are reduced to varying degrees, which indicates the effectiveness of entity feature decoupling. Only 3 or 8 bases are sufficient to express the characteristics of different relationships in the model. In this paper, a knowledge graph completion model with good improvement effect is obtained. This study lays a foundation for the popularization of knowledge atlas in badminton.

**Keywords** Knowledge graph completion, Deep learning, Badminton, Graph convolution neural network, Attention mechanism

## 1 引言

2012年5月,Google提出了知识图谱这一概念,并在其搜索引擎中应用了Google知识图谱,以达到增强搜索结果的目的。在随后的一段时间里,知识图谱得到了工业界与学术界研究者的广泛关注<sup>[1-2]</sup>。用数据结构对知识进行表示是知识图谱的本质目的,这一思想来源于1950s万维网之父Berners-Lee,该学者提出了“语义网络”,随后Berners-Lee还提出了“链接数据”,知识图谱便是基于这些技术发展起来的<sup>[3]</sup>。我们可以将知识图谱看作一张规模庞大的图,图中的节点代表实体或概念,边代表实体或概念之间的关系,通过这种方法对真实世界中的实体或概念进行描述。要对特定知识的相关概念及知识之间的关系进行识别与挖掘,达到知识互连的目的,那么对自动化构建知识覆盖面广、质量高的知识

图谱进行研究就变得十分必要<sup>[4-5]</sup>。

特定领域的知识图谱和通用领域的知识图谱是知识图谱的两个大的分类。特定领域知识图谱(Domain Specific Knowledge Graph, DSKG)的语义网络构建以特定实体或关系为基础,这些实体及实体之间的关系来源于特定领域的数据库资源<sup>[6]</sup>。该类型图谱中所包含的知识针对性和专业性较强。由于特定领域知识图谱广泛应用于对某个专业领域进行深入研究,因此对DSKG的准确性也提出了较高要求。与特定领域知识图谱相比,通用领域知识图谱(Universal Knowledge Graph, UKG)的知识是具有通用性质的,没有划分领域<sup>[7-8]</sup>。Cai等<sup>[9]</sup>利用知识图谱对国外体育法学近十年的研究热点及研究现状进行了分析,从研究机构、地域分布等方面进行了总结,并对近年来国外体育法学的发展趋向和发展脉络进行了较为清晰的描述,指出国际体育法、反兴奋剂、运动

基金项目:国家自然科学基金(11551003)

This work was supported by the National Natural Science Foundation of China(11551003).

通信作者:胡赫(huhepingzi@163.com)

员权益、预防、反垄断等是国外体育法目前研究的重点问题。如今,大多数知识图谱由半自动形式或人工形式构造,容易出现实体信息不完整现象<sup>[10]</sup>。实体信息的缺乏会导致知识图谱稀疏,使得实体之间的许多语义链接没有得到反映,这将影响基于知识图谱的下游任务的效果。

随着科研体系的不断成熟和完善,而羽毛球运动与科技结合的基础理论的研究还不够丰富,极易导致科研发展的滞后,从宏观角度对羽毛球运动的发展进行整体把握,同时对羽毛球运动发展过程中面临的挑战进行分析,是目前研究的一个重点方向。本研究的创新点在于在知识图谱补全模型中引入基底表示不同语义,并采用语义相关注意力机制进行特征更新。本文首先对羽毛球知识图谱构建的相关代表性研究进行了分析,对目前应用较为广泛的两类基于深度学习的羽毛球知识抽取技术进行了阐述;其次介绍图神经网络,并在此基础上提出了知识图谱补全模型;最后对改进模型进行了验证。本研究最终得到了可行的研究方法,研究结果为知识图谱赋能的深度学习技术发展提供了一定思考。

## 2 相关工作

知识图谱通过各种渠道对信息进行收集,达到增强搜索引擎的目的,是谷歌及其服务使用的知识库。三元组是知识图谱的基本组成形式,通过实体间的相互关系对其进行联结,从而构成复杂的网状知识结构。图1给出了一个简单的知识图谱网络模型。

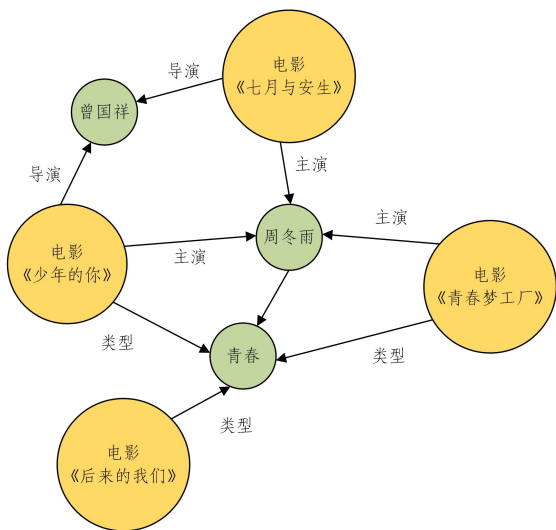


图1 知识图谱网络模型实例

Fig. 1 Example of a knowledge mapping network model

根据知识覆盖领域的差异,可以将知识图谱细分为领域知识图谱和通用知识图谱。Yang等<sup>[11]</sup>从知识组织、表示及应用这3个方面对特定领域的知识图谱和开放领域的知识图谱进行了对比,如图2所示。

知识图谱的补全方法大致可分为基于符号推理、路径发现、强化学习、元学习和表示学习的方法。其中,基于符号推理的方法对规则本身进行建模,并通过知识图集中实体和关系的符号直接完成推理。基于路径搜索的知识图谱完成方法通过随机游走对实体之间的路径进行建模,构造路径特征并训练分类器,从而预测实体之间的潜在关系<sup>[12]</sup>。基于强化学习的知识图谱完成方法改进了路径发现方法,将路径查找过

程表示为序列决策,并在向量空间中编码;通过设置各种奖励函数,路径发现过程比随机行走更灵活、可控。基于元学习的知识映射完成方法不再直接学习数据到标签的映射,而是学习数据到函数的映射,从而使模型获得“学习”的能力<sup>[13]</sup>。基于路径查找和符号推理的方法存在准确率和补全效率低的问题,往往无法对结构复杂或规模较大的知识图谱进行处理。基于元学习和强化学习的方法目前在复杂任务处理中的研究较少。但已有研究表明,基于表示学习的知识图谱补全方法具有良好的特征表达能力和计算效率<sup>[14]</sup>。

维度		领域知识图谱	通用知识图谱
知识表示	广度	窄	宽
	深度	深	浅
	粒度	细	粗
知识组织	数据模式	动态	固定
	质量要求	苛刻	高
	自动化程度	低	高
	复杂程度	高	低
知识应用	推理关系链	长	短
	应用复杂度	高	低

图2 领域知识图谱与通用知识图谱对比

Fig. 2 Comparison of domain knowledge mapping and generic knowledge mapping

基于表示学习的知识图谱完成方法使用向量来表示实体和关系,在连续向量空间中学习它们的最终表示,然后根据评分函数计算和排序候选三元组的评分结果,从而实现知识图完成。根据评分函数的不同,这些方法可具体分为基于平移距离的方法、基于语义匹配的方法和基于神经网络的方法。研究人员对距离平移方法的研究发现,该方法可以将实体空间中的实体映射到不同关系空间中进行不同关系对实体属性的关注偏好的学习<sup>[15]</sup>。基于语义匹配的方法中,目前的研究已引入负值嵌入,以简化任务的复杂度<sup>[16]</sup>。基于神经网络的方法通过循环神经网络建模以得到不同的关系路径特征。表示学习方法通常需要在训练阶段获得所有实体的特征表示,但在“测试集包含训练集中没有的新实体(Out of Knowledge Base, OOKB)”问题中,新实体不会出现在训练过程中,因此通用表示学习方法不适用于 OOKB 问题的知识图完成任务<sup>[17]</sup>。要解决这个问题,可以从知识地图的路径信息中提取语义规则,这些语义规则独立于实体,因此可以用于改进 OOKB 问题下的知识映射完成任务。

通过对知识地图完成研究现状进行分析,可以看出基于表示学习的方法仍然占主导地位。在众多表示学习方法中,基于图神经网络的方法因其强大的图特征处理能力而备受关注<sup>[18]</sup>。对于 OOKB 问题,与在训练阶段直接学习所有实体的特征表示的一般表示学习方法不同,它需要学习独立于实体的语义规则或为实体设置一般初始特征,以便未经训练阶段的新实体也可以获得高质量的特征表示<sup>[19]</sup>。

## 3 基于深度学习的羽毛球知识图谱相关技术

### 3.1 图神经网络

图神经网络(Graph Neural Networks, GNN)是对保持图

对称性的图的所有属性的可优化转换。对称信息指的是把这个顶点进行另外一种排序后,整个结果不会变。GNN 被用于非欧几里得数据图结构的特征提取<sup>[20]</sup>。图卷积神经网络(Graph Convolution Networks,GCN)是图神经网络的重要分支。GCN 根据卷积方式的不同可分为基于空域和基于谱域两种。基于谱域的图卷积方法借鉴了数字信号处理的思想:通过傅里叶变换将时域信号变换为频域信号进行分析,然后再完成一些在时域信号中不方便完成的操作后将其逆变换回来。基于谱域的图卷积方法不支持图结构的动态变化,因此受到的限制较大<sup>[21]</sup>。基于空域的图卷积方法与卷积神经网络从卷积核中提取图像局部特征的思路类似。基于空域的图卷积方法可以分为消息传递和特征更新两个步骤<sup>[22]</sup>。

图注意力网络属于基于空域的图卷积方法。神经网络中的注意力机制(Attention Mechanism)是在计算能力有限的情况下,将计算资源分配给更重要的任务,同时解决信息超载问题的一种资源分配方案。在神经网络学习中,一般而言,模型的参数越多则模型的表达能力越强,模型所存储的信息量也越大,但这会带来信息过载的问题。那么通过引入注意力机制,在众多的输入信息中聚焦于对当前任务更为关键的信息,降低对其他信息的关注度,甚至过滤掉无关信息,就可以解决信息过载问题,并提高任务处理的效率和准确性<sup>[23]</sup>。图注意力网络(Graph Attention Networks,GAT)的关键即是图注意力层,图注意力层的处理对象即为整张图的每个结点,如对于  $N$  个结点,每个结点特征维度为  $F$  的一张图,通过图注意力层得到的输出即为  $N * F$ 。如图 3 所示。

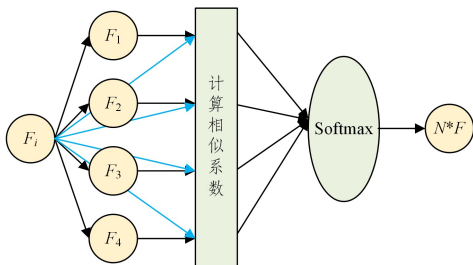


图 3 GAT 中的注意力机制  
Fig. 3 Attention mechanism in GAT

图神经网络中编码解码结构的基本思想是将输入数据转换为图结构数据,这也被广泛应用于知识图谱的各种任务中。根据学习方法的不同,图编解码结构可分为具有自监督学习的图自编码器和具有监督学习的图形编解码模型<sup>[24]</sup>。图自动编码器(Graph Auto-Encoder,GAE)是一种自编码模型,由编码器和解码器组成。编码器负责学习输入图的邻接矩阵中的结构信息,并将其转换为中间向量  $Z$ ;解码器负责将中间向量  $Z$  恢复到输出图的邻接矩阵,这称为解码。该模型的训练目标是使输入邻接矩阵和输出邻接矩阵尽可能相似。整个训练过程属于自监督学习。

在现有的基于图神经网络的知识映射完成方法中,更多的是使用有监督学习的编码-解码模型。其中,编码器部分通常通过图形神经网络提取图形数据中的结构特征,并获得三元组的嵌入特征<sup>[25]</sup>。解码器使用翻译距离或语义匹配对三元组的嵌入特征进行解码,并输出三元组分数。有监督学习的编解码模型通常需要使正样本分数大于负样本分数,并相应地训练编码器和解码器。

### 3.2 知识图谱补全

知识图谱补全是在知识图中完成不完整的三元组,在现有实体集和关系集的基础上向知识图中添加三元组并确保添加的三元组不与现有事实三元组冲突,从而解决知识图中普遍存在的数据稀疏问题<sup>[26]</sup>。基于表示学习的知识图完成方法通过表示学习模型获得相似的潜在知识或语义规则,并判断候选三元组属于正确三元组的可能性,从而完成知识图。根据补全对象的不同,知识图谱补全可分为实体预测任务和关系预测任务。

基于表示学习的知识图谱补全方法使用向量表示所有实体和关系,并不断优化连续向量空间中实体和关系的表示。最后,通过评分函数计算所有候选三元向量表示的得分,并选择得分最高的候选三元中的元素作为完成结果。根据评分函数的不同,现有的基于表示学习的知识图完成方法可分为基于翻译距离的方法、基于语义匹配的方法和基于神经网络的方法。基于表示学习的知识图谱补全方法整体流程为:首先根据正样本进行负采样;然后通过模型计算正负样本三元组得分;最后排序正负三元组得分,计算损失并进行训练<sup>[27]</sup>。

前文提到的 OOKB 问题可以通过重新训练来解决,但重新训练整个知识图的成本太高。实体拥有的其他信息,如图像描述和文本描述,也可用于生成实体特征,但这种情况需要额外的存储空间和数据来提供描述支持。因此,目前对 OOKB 的研究主要集中在如何为实体生成高质量的嵌入特征,而不需要完全的再训练和附加信息。对于一般知识图谱补全任务,当测试集中的实体已经包含在训练集中时,它实际上正在进行传导学习,即直接将训练集中学习的知识扩展到测试集中。在 OOKB 问题下完成知识图完成的过程本质上是归纳学习。由于训练集不再包含测试集中的实体,因此不可能将知识直接扩展到测试集。它只能总结训练集中的知识模式,然后将它们应用于新的测试集。

## 4 基于子图结构的知识图谱补全模型构建

### 4.1 基于子图结构的知识图谱补全模型框架

本研究提出的基于子图结构解耦的知识图谱补全模型共包含 4 个部分,分别是子图抽取、特征初始化、多维度特征更新及特征融合,如图 4 所示。

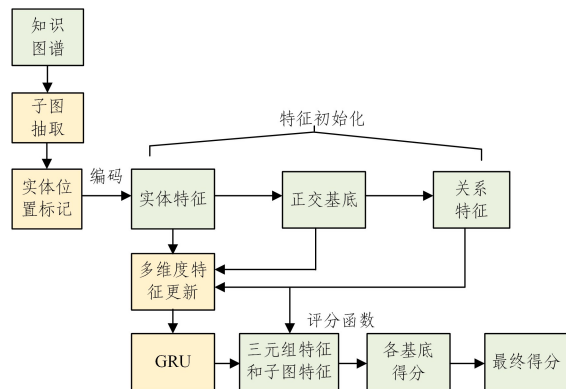


图 4 补全模型整体架构

Fig. 4 Overall architecture of the complementary model

对于子图提取,需要选择头部和尾部实体的  $k$  跳邻域中的实体以形成子图,作为模型后续操作的基础。特征初始化通过编码等步骤为关系和实体设置通用初始特征,以满足归

纳学习的要求。多维特征更新是指在解耦后将实体特征映射到不同的语义空间,并结合语义相关性分别更新特征。最后,考虑到不同语义的待预测关系的偏好程度不同,我们通过组合选通循环单元,聚合多语义空间中的结果,并提取实体排列之间的序列特征,最后将特征输出到评分函数以获得预测结果。

在预测两个实体之间的关系时,并非所有实体连接都有助于关系预测,并且,彼此远离的实体很少携带有助于进行关系预测的语义信息。因此,该模型需要处理原始知识图集中的数据集,并且只考虑两个实体之间的局部子图。考虑到隐藏在知识图结构中的规则信息,可以从子图结构的角度构造实体的初始特征<sup>[28]</sup>。

为了获取子图中的潜在信息,本研究对子结构进行了解耦,具体如式(1)一式(3)所示:

$$A = x_u \cdot \mathbf{M}_{\text{base}}^{-1} \quad (1)$$

$$h_{u,i} = a_i \mathbf{M}_{\text{base},i} \quad (2)$$

$$X_u = [h_{u,1}, h_{u,2}, \dots, h_{u,m_1}] \quad (3)$$

其中, $A$ 为实体, $x$ 为初始特征, $\mathbf{M}_{\text{base},i}$ 为基地矩阵的伪逆, $h$ 为实体特征, $i$ 为语义空间序号, $a$ 为元素值, $X$ 为输出特征, $m$ 为实体更新次数。

在得到实体特征在不同语义空间的映射后,使用图神经网络对子图进行实体特征更新,并对各语义空间注意力系数进行计算。计算过程如式(4)和式(5)所示:

$$s_{(u,v,j)}^k = f_1(h_{(u,j)}^k, h_{(v,j)}^k, M_{(rel,u,v,j)}, M_{(rel,label,j)}) \quad (4)$$

$$\alpha_{(u,v,j)}^k = \text{softmax}(s_{(u,v,j)}^k) = \frac{\exp(s_{(u,v,j)}^k)}{\sum_{(v') \in N(u)} \exp(s_{(u,v',j)}^k)} \quad (5)$$

其中, $k$ 表示图神经网络层数, $f$ 为全连接层。 $h$ 为语义空间的特征, $u$ 和 $v$ 表示实体, $j$ 表示语义空间, $M_{(rel,u,v,j)}$ 为连接两个实体的特征, $\alpha_{(u,v,j)}^k$ 为注意力系数。不同基底的实体特征更新如式(6)所示:

$$h_{(u,j)}^k = \sum_{v \in N(u)} \alpha_{(u,v,j)}^{k-1} W_h h_{(v,j)}^{k-1} + h_{(u,j)}^{k-1} \quad (6)$$

最终得到的不同基底下的特征如式(7)所示:

$$X_u^k = [\hat{h}_{(u,1)}^k, \hat{h}_{(u,2)}^k, \dots, \hat{h}_{(u,m_1)}^k] \quad (7)$$

则三元组得分为式(8):

$$\text{score}_i = f_2((x_i^{\text{head}} + x_{r,i}^{\text{relation}} - x_i^{\text{tail}}) \parallel x_i^{\text{avg}}) \quad (8)$$

其中, $x_i^{\text{head}}$ 为头实体, $x_i^{\text{tail}}$ 为尾实体, $i$ 为基底, $x_{r,i}^{\text{relation}}$ 为关系 $r$ 在基底 $i$ 上的特征, $x_i^{\text{avg}}$ 为平均特征。不同基底得分在该类型关系预测中的权值 $\text{att}$ 如式(9)所示:

$$\text{att} = \text{abs}(\mathbf{W}_{\text{rel}}(r)) \quad (9)$$

其中, $\mathbf{W}_{\text{rel}}$ 为系数矩阵, $r$ 表示关系; $\text{abs}$ 为取绝对值操作。不同基底上的权值代表的偏好程度不同,最终的得分结果如式(10)所示:

$$\text{score} = \text{att} * [\text{score}_1, \text{score}_2, \dots, \text{score}_{m_1}] \quad (10)$$

## 4.2 模型训练与实验设置

由于需要保证训练出的模型在正样本上的评分高于负样本,因此本研究采用边际损失函数来训练模型。结合伯努利负采样随机替换三元组尾部或头部,使用随机梯度下降优化公式进行分析<sup>[29]</sup>,如式(11)所示:

$$\mathcal{L} = \sum_{i=1}^{|\epsilon|} \max(0, \text{score}(n_i) - \text{score}(p_i) + \gamma) \quad (11)$$

其中, $\epsilon$ 为三元组样本集合, $p_i$ 为正三元组, $n_i$ 为负三元组, $\gamma$ 为边距超参数。

本研究中使用的数据集为3个常用知识图谱数据集

wn18rr,fb15k-237和nell-995。将每个数据集划分为4个子数据集,每个子数据集都含有测试图和训练图。

本研究使用AUC-PR和Hits@ $n$ 指标对模型的补全能力进行评估。Hits@ $n$ 指标为命中率,表示正确三元组在排名结果中小于 $n$ 的占比, $n$ 的值一般取1,3或10。该指标越大模型性能越好。计算公式如式(12)所示:

$$\text{Hits}@n = \frac{1}{|S|} \sum_{i=1}^{|S|} \mathbb{I}(\text{rank}_i \leq n) \quad (12)$$

AUC-PR表示精度和召回率曲线图。计算过程如式(13)所示:

$$AP = \sum_n (R_n - R_{n-1}) P_n \quad (13)$$

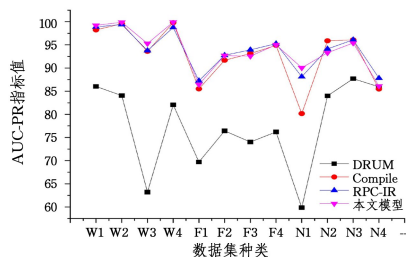
其中, $R$ 表示召回率, $P$ 表示精度, $n$ 表示阈值序号。

本研究使用的优化器为Adam,batch size设置为16。数据集wn18rr,fb15k-237和nell-995的损失函数边距分别设置为8,16,10,基底数为3,32,16,学习率均为0.01。所有模型最大训练轮数为10,实体特征和关系特征的嵌入维度为32。在测试阶段和训练阶段生成子图时, $K$ 值设置为3<sup>[30]</sup>。对于不同的指标,Hits@ $n$ 的测试在每个三元组随机采样50个负样本的前提下进行,计算完成后,查看测试样本排名是否在前10内。AUC-PR指标通过对比每个测试三元组随机匹配的一个负样本大小来实现。每个子数据集下均对模型进行了3次训练,选取测试结果的平均值为最终性能。结果分析部分将具体分为3个实验,分别是本文模型与基线模型的对比分析、模型各组件消融实验分析以及基底数目对比分析。

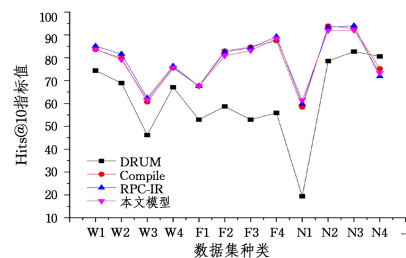
## 5 知识图谱补全模型结果分析

### 5.1 本文模型与基线模型的对比分析

本研究选择了3个常见基线模型与本文构建模型进行不同数据集下的对比分析,这3个模型分别是DRUM(Differentiable Rule Mining)、Compile和RPC-IR(Remote Procedure Call-Information Retrieval)<sup>[31]</sup>。如图5所示。



(a) AUC-PR



(b) Hits@ $n$

图5 本文模型与基线模型的对比

Fig. 5 Comparison between the model in this paper and the baseline model

图5中,W,F,N分别表示wn18rr,fb15k-237和nell-995,数字1-4为它们的子集。本文模型在所有子数据集上

都取得了良好的结果,与最佳基线模型相当。特别是在 wn18rr 数据集上,本文方法的 AUC-PR 指数与最佳基线相比平均增加了 0.5%。Hits@10 指标在大多数子数据集上达到了第二。另一方面,fb15k-237 和 nell-995 的性能不一致,在某些子数据集上仍然可以达到最佳效果,但在某些数据集上性能较差。这可能是因为在由每个子数据集的训练集生成的子图中存在断开连接的结构。

## 5.2 模型各组件消融实验分析

为验证本文模型的改进有效性,在不同数据集上进行了消融实验,结果如图 6 所示。

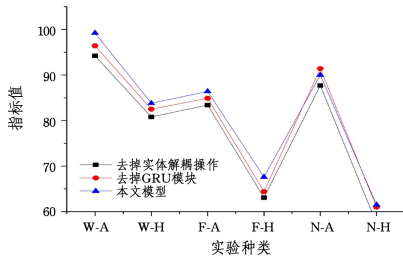


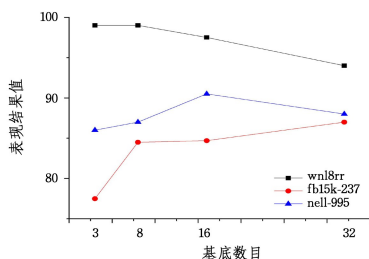
图 6 消融实验结果

Fig. 6 Results of ablation experiment

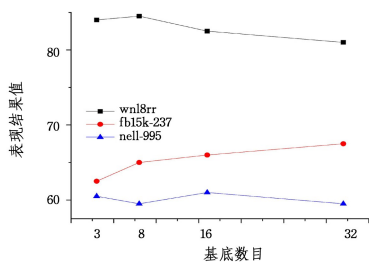
从图 6 中可以得到,当实体解耦操作从模型中移除时,在实验中选择 3 个数据集上,两个测试指标都不同程度地下降,这表明实体特征解耦的有效性。当 GRU 模块从模型中移除时,Hits@10 指标在 3 个子数据集中仍呈下降趋势,而在 AUC-PR 指标 nell-995 中,第一个小数据集分类的指标值反而增加。出现这种现象的原因可能是 nell-995 的子数据集的知识图相对稀疏,导致生成的子图上的连接较少,使得 GRU 无法从子图中提取有效序列特征,破坏了原始特征分布。

## 5.3 基底数目对比分析

分别选取 wn18rr,fb15k-237 和 nell-995 数据集中的一个子数据集。对基底数目在模型中产生的影响进行分析,结果如图 7 所示。



(a) AUC-PR



(b) Hits@n

图 7 不同基底数目对比

Fig. 7 Comparison of the number of different substrates

由图 7 可以发现,不同的子数据集对不同的碱基数有

同的性能结果,其中 wn8rr 的子数据集在基底数为 3 或 8 时可以达到最佳效果,然后随着碱基数的增加,模型的效果逐渐变差。fb15k-237 的子数据集在基底数为 32 时效果最佳,而 nell-995 的子数据集在基底数为 16 时效果最佳。本文分析出现这种情况的原因与不同子数据集中包含的关系数不同有关,其中 wn18rr 的子数据集至少包含 9 条关系数目,因此,只需 3 个或 8 个基底就足以表达不同关系的特征。更多的基底将使模型难以熟悉不同特征下的关系。

**结束语** 基于能够在羽毛球领域获得更多发展的目的,本文基于深度学习技术和补全方法,对羽毛球知识图谱构建进行了研究。首先对目前的研究工作进行了总结分析,然后构建了基于子图结构解耦的知识图谱补全模型,最后对模型进行了改进有效性评估。本研究为知识图谱的发展提出了未来的研究方向,即该补全知识图谱将推广至羽毛球领域。但本研究还存在一定不足。在与基线模型的对比分析中,本研究选择了 3 个常见基线模型,对比模型的数量较少;对于羽毛球领域的应用阐述较少,未结合羽毛球知识图谱进行案例分析。未来研究中,将加入更多的基线模型进行对比,并添加更多关于羽毛球知识图谱的内容,使知识图谱相关研究在羽毛球领域得到进一步发展。

## 参考文献

- [1] YAO B, LIANG N. A smart position optimization scheme for badminton doubles based on human-computer interactive training in wireless sensor networks [J]. EURASIP Journal on Wireless Communications and Networking, 2020, 2020(1): 1-16.
- [2] CAO Z, LIAO T, SONG W, et al. Detecting the shuttlecock for a badminton robot: A YOLO based approach [J]. Expert Systems with Applications, 2021, 164: 1-11.
- [3] WANG H, KLÄSER A, SCHMID C, et al. Dense trajectories and motion boundary descriptors for action recognition [J]. International Journal of Computer Vision, 2013, 103(1): 60-79.
- [4] SANG S, YANG Z, WANG L, et al. SemaTyP: a knowledge graph based literature mining method for drug discovery [J]. BMC Bioinformatics, 2018, 19(1): 1-11.
- [5] WANG C, MA X, CHEN J, et al. Information extraction and knowledge graph construction from geoscience literature [J]. Computers & Geosciences, 2018, 112(MAR.): 112-120.
- [6] ZHU H, TIAN F, WU K, et al. A multi-constraint learning path recommendation algorithm based on knowledge map [J]. Knowledge-Based Systems, 2018, 143: 102-114.
- [7] NATTHAWUT K, RYUTARO I. An Automatic Knowledge Graph Creation Framework from Natural Language Text [J]. Ice Transactions on Information & Systems, 2018, 101(1): 90-98.
- [8] YU J, ZHANG Y, WU Y, et al. Research on the Practical Application of Visual Knowledge Graph in Technology Service Model and Intelligent Supervision [J]. Journal of Physics: Conference Series, 2021, 1982(1): 012040.
- [9] CAI Z D, TANG J L, YU R J, et al. Situation and Fronts of International Sports Law Research Based on Knowledge Mapping [J]. Journal of Beijing Sport University, 2016, (6): 44-50.
- [10] WANG S, LEE J, KIM W J, et al. LiDAR Depth Completion

- Using Color-Embedded Information via Knowledge Distillation [C] // IEEE Transactions on Intelligent Transportation Systems, 2021.
- [11] YANG B, YANG M F. Overview of Atlas and Its Application in Risk Management [J]. Journal of Chinese Mini-Micro Computer Systems, 2021, 42(8):9.
- [12] MEI Y, GE Y, ZHANG Y, et al. Research on Drilling and Completion Design Knowledge Base System based on Knowledge Map[C]//2022 IEEE 2nd International Conference on Electronic Technology, Communication and Information. IEEE, 2022: 493-496.
- [13] HWANG S, LEE J, KIM W, et al. LiDAR Depth Completion Using Color-Embedded Information via Knowledge Distillation [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 16:221.
- [14] WANG H, YANG X, SUN Y. Research on relationship error detection and knowledge completion methods of knowledge atlas [J]. Computer and digital engineering, 2022, 50(7):1534-1539.
- [15] TAN Z, CHEN Z, FENG S, et al. KRACL: Contrastive Learning with Graph Context Modeling for Sparse Knowledge Graph Completion[J]. arXiv:2208.07622, 2022.
- [16] HOU J, CHOU R, XUE J, et al. Elevator fault prediction based on knowledge map embedding and completion[J]. Computer Engineering and Design, 2022, 43(1):224-230.
- [17] LATINJAK A T, HATZIGEORGIADIS A. The knowledge map of sport and exercise psychology: An integrative perspective[J]. Frontiers in psychology, 2021, 12:2320.
- [18] ZHANG D, WANG H, ZHONG W, et al. WGAN knowledge map completion method integrating soft reward and exit mechanism [J] Journal of Zhengzhou University (Science Edition), 2022, 54(2):67-73.
- [19] LOVELACE J, NEWMAN-GRIFFIS D, VASHISHTH, et al. Robust knowledge graph completion with stacked convolutions and a student re-ranking network[J]. arXiv:2106.06555, 2021.
- [20] REN Y, JIANG J, NI Q. Research on financial crisis early warning based on graph neural network model --from the perspective of information spillover across the industry [J] Statistical research, 2022, 39(8):141-160.
- [21] JIANG W, LUO J. Graph neural network for traffic forecasting: A survey [J]. Expert Systems with Applications, 2022, 28: 117921.
- [22] YANG X, JIN X, TAO J, et al. Text classification based on graph neural network and dependency parsing [J] Computer science, 2022, 11:1-13.
- [23] ASIF N A, SARKER Y, CHAKRABORTTY R K, et al. Graph neural network: A comprehensive review on non-euclidean space [J]. IEEE Access, 2021, 9:60588-60606.
- [24] WANG D, XIANG C, PAN Y, et al. A deep convolutional neural network for topology optimization with perceptible generalization ability[J]. Engineering Optimization, 2022, 54(6):973-988.
- [25] SULAIMAN A, OMAR K, NASRUDIM M F. Two streams deep neural network for handwriting word recognition[J]. Multimedia Tools and Applications, 2021, 80(4):5473-5494.
- [26] KIM M, BAEK S. ComDense: Combined Dense Embedding of Relation-aware and Common Features for Knowledge Graph Completion[J]. arXiv:2206.14925, 2022.
- [27] LI Z, DING Y, HUA Z, et al. A knowledge map completion model combining the importance of triplets [J]. Computer Science, 2020, 47(11):231-236.
- [28] CHEN H, WANG W, LI G, et al. The model of knowledge graph complement for Quaternion relation rotation [J]. Computer Science, 2021, 48(5):225-231.
- [29] NAWAZ M S, SHOAB B, ASHRAF M A. Intelligent Cardiovascular Disease Prediction Empowered with Gradient Descent Optimization[J]. Heliyon, 2021, 7(5):e06948.
- [30] HU Y, GE L. An improved differential privacy parameter setting and data optimization algorithm [J] Computer Engineering and Science, 2021, 43(10):1758-1765.
- [31] ESFAHANI A N, MOGHADDAM N B, MALEKI A, et al. The knowledge map of energy security[J]. Energy Reports, 2021, 7: 3570-3589.



**CHEN Yujue**, born in 1998, Ph.D. Her main research interest is sports and artificial intelligence.



**HU He**, born in 1999, postgraduate. His main research interest is natural language processing.