



# 知识图谱发展与构建的研究进展

## 摘要

近年来知识图谱技术引起了广泛的关注和研究,本文介绍了近期知识图谱的发展、构建方法、详细的构建过程,并对知识图谱在交叉学科领域的应用和未来的研究方向做了总结.本文详细介绍了构建文本知识图谱、视觉知识图谱、多模态知识图谱的关键技术,比如信息提取、知识融合、知识表示等.作为知识工程的重要组成部分,知识图谱,尤其是多模态知识图谱的发展对大数据时代的高效知识管理、知识获取、知识共享有着重要的意义.

## 关键词

知识图谱; 语义网; 自然语言处理; 视觉知识图谱; 多模态

中图分类号 TP182

文献标志码 A

## 0 引言

知识图谱的起源可以追溯到 20 世纪 50 年代末 60 年代初语义网 (semantic network) 的诞生.当时的语义网是一种基于图的用于存储知识的数据结构,图的节点代表实体或者概念,图的边代表实体或概念间的关系,主要应用于机器翻译和自然语言处理中.到 20 世纪 80 年代,知识工程和基于规则的专家系统被提出并成为研究的重点.20 世纪 90 年代,机构知识库的概念被提出,知识表示和知识组织开始被深入研究,并广泛应用到各机构单位的资料整理工作中.

互联网在 21 世纪蓬勃发展,随着信息量的爆炸式发展和搜索引擎的出现,人们开始期望更加快速、准确地获得所需的信息.2012 年 11 月,谷歌公司提出知识图谱 (knowledge graph) 项目,该项目的关键技术包括从互联网的网页中抽取实体、实体属性信息及实体间的关系,旨在解决自动问答、个性化推荐和智能信息检索方面的问题.谷歌公司宣布要以此为基础构建下一代智能搜索引擎,既能分析也能理解用户真正的需求,而不仅仅只是基于关键词的搜索.在谷歌知识图谱的基础上,知识图谱的概念正式提出,定义为把所有不同种类的信息连接在一起得到的一个关系网络.通过对知识进行这种更加有序、有规律的组织,可以给用户提供更加智能的信息获取和管理.

知识图谱本质上可以算作一个语义网,具有语义网的很多特性,但又有所不同.知识图谱和语义网的目标都是实现人与计算机的无障碍沟通,让计算机能理解语言和概念,能够理解逻辑并拥有判断和推理能力,也就是使计算机更加智能化、自动化和人性化.语义网通过建立能够被计算机所理解的语义“元数据 (meta data)”,使网络上的信息能通过统一的标准成为一个通用的信息交换媒介,从而大大提高网络的搜索效率和推理理解能力,知识图谱则是以图谱的形式达到相同的目的.但是知识图谱的数据来源可以是百科类的半结构化数据和各种网络数据,关注的更多是知识清洗、知识融合和知识表示技术.早期语义网则多是靠人工构建,关注重点也不一样<sup>[1]</sup>.

目前一些比较大的知识图谱有谷歌公司的“谷歌知识图谱”,用多种语言对知识图谱中的实体、属性和关系进行描述,以达到提高搜索质量的目的;百度公司的“知心”,通过筛选、整合搜索结果,达到直接搜索到答案的效果;搜狗公司的“知立方”通过“语义理解”网上的碎片信息和优化结果,可以向用户显示最符合的信息<sup>[2]</sup>.

收稿日期 2017-08-28

资助项目 国家自然科学基金 (61572503、61432019); 北京市自然科学基金 (4152053)

作者简介

朱木易洁,女,主要研究方向为模式识别. zhumuyj@126.com

徐常胜 (通信作者),男,博士,研究员,研究方向为模式识别. csxu@nlpr.ia.ac.cn

<sup>1</sup> 中国科学院自动化研究所 模式识别国家重点实验室,北京,100190

本文的第1部分将进一步阐述知识图谱的一些基本概念和概况,第2部分至第4部分依次详细介绍文本知识图谱、视觉知识图谱、多模态知识图谱的构建框架、关键技术及应用,第5部分介绍知识图谱目前面临的挑战及未来的研究方向并对全文内容进行总结。

## 1 知识图谱简介

### 1.1 基本概念

知识图谱作为结构化的语义知识库,使用“资源描述框架(RDF, Resource Description Framework)”来描述数据,包括一系列的节点、边和属性,基本模型是三元组,一般为“实体-关系-实体”三元组。其中,节点表示物理世界中的实体或概念,边表示各种关系。实体(entity)指的是实际存在的事物,比如一个人物——奥巴马、一个国家——中国、一个城市——广州。概念(concept)是比实体更抽象的表达,比如人、颜色、天气。关系(relation)是实体与实体、实体与概念之间的结构或联系,包括层级关系,比如“属于”、“一部分”,和非层级关系,比如属性等。本体(ontology)在知识图谱的概念中也总是被提及,它属于知识表示的概念,来源于哲学的“本体论”。“知识本体”最初研究世界上各种事物以及代表这些事物的范畴的形式特性,并对其进行分类、建立规范,后来在计算机科学中,“知识本体”有了科学定义,表示概念体系的明确规范,使其概念明确、形式简单、容易共享,比如生物学领域知识本体就是生物学领域知识的抽象。同样地,“本体”是对领域实体存在本质的抽象,强调实体间的关联,并用多种知识表示元

素表达出来。构造本体的目的就是为了实现一定程度的知识共享和重用。本体强调实体间或概念间的关联关系,知识图谱在本体的基础上,增加了更多关于实体的信息。也就是说,本体描述了知识图谱的数据模式(scheme),给知识图谱构建数据模式相当于给它建立了本体。

### 1.2 数据存储

知识图谱的数据存储主要有2种方案: RDF存储和图数据库存储。

RDF存储又称三元组(triple)存储,使用唯一的URI标示一个资源(resource),一个资源通常又标示一个事物(thing),资源与资源间的关系用属性(property)来描述。3种资源可以组成三元组,一般格式为(主语(subject),谓语(predicate),宾语(object))。三元组模式的查询、归并和连接都非常高效,但因为其自身索引方式的问题,空间开销比较大,更新维护的代价也大,目前更多的RDF数据库使用列式存储,查询效率也逐渐得到了提高。

相对来说,图数据库可以提供完善的图查询语言、支持各种图挖掘算法,但是图数据库的分布式存储使得大节点处的处理开销高,数据更新速度慢,整体的实现代价比较高。图1展示了各种数据存储模式近几年的发展情况<sup>[3]</sup>,可以看出基于图存储的数据库的数量一直保持领先,是存储方式中的主流,其中Neo4j最为流行,而Jena是RDF领域内目前最流行的存储框架。

### 1.3 构建方式和方法

知识图谱的构建主要包括自底向上(bottom-up)

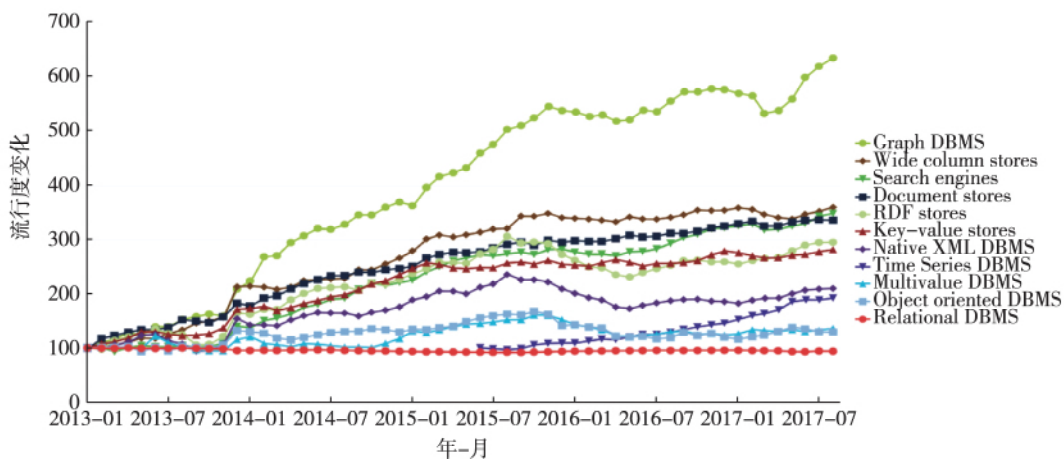


图1 各种数据模式的发展情况<sup>[3]</sup>

Fig. 1 Development trend of each data structure<sup>[3]</sup>

和自顶向下(top-down) 2 种方式. 自底向上就是先获得知识图谱的实体数据, 然后再构建本体, 即先得到具体再得到抽象的概念, 最典型的是谷歌的 Knowledge Vault. 自顶向下的方式则是先定义或得到本体的数据, 再逐渐将具体的实体加入到知识图谱中, 典型的有 Freebase. 目前大多数知识图谱都是采用自底向上的方式进行构建, 知识图谱大致的构建流程如图 2 所示. 知识图谱的构建主体包括知识获取、知识融合、知识加工和知识更新 4 个方面, 根据数据来源的不同构建流程也会有些许区别.

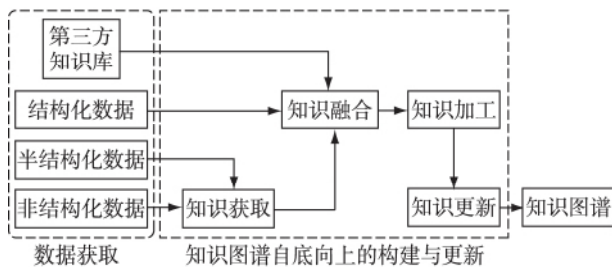


图 2 知识图谱的构建流程

Fig. 2 Flowchart of knowledge graph construction

知识图谱的构建方法有完全的专家手工构建、利用众包构建(比如 Wikipedia 和 Freebase)、半自动构建(比如预先人工设定规则或正则, 然后从半结构化的信息中再进行自动构建)和完全从非结构化的信息中自动构建.

#### 1.4 数据来源与知识库

知识图谱的数据来源主要包括结构化数据、半结构化数据、非结构化的数据, 结构化的数据指已经组织好的 RDF 数据. 目前有很多已经构建好的知识库, 它们包含半结构化、非结构化的数据, 是知识图谱的重要数据来源之一, 比较有名的有 Freebase<sup>[4]</sup>、Wikidata<sup>[5]</sup>、DBpedia<sup>[6]</sup>、YAGO<sup>[7]</sup>、IMDB<sup>[8]</sup>. 前 4 个是开放的链接知识库, IMDB 是典型的垂直行业知识库, 即只描述特定领域知识的知识库.

Freebase 知识库现在是谷歌知识图谱的重要组成部分, 它的数据主要依靠人工构建, 其他数据主要来自维基百科、NNDB 和 MusicBrainz 等网站或语料库. 2015 年 6 月宣布整体迁移至 Wikidata. Wikidata 是维基百科基金会主持的一个自由的协作式多语言辅助知识库, 作为 Wikipedia、Wikivoyage、Wikisource 中结构化数据的中央存储器, 数据主要以文档的形式存储, 每个文档都有唯一的数字标识. DBpedia 是由莱比锡大学和曼海姆大学共同创建的多语言综合

型知识库, 它从多种语言的维基百科中抽取结构化的信息, 并将其以关联数据的形式发布到互联网上, 提供给在线网络应用、社交网站或其他在线知识库使用. YAGO 是德国马普研究所(MaxPlanck Institute, MPI) 的科研人员构建的综合型知识库, 它的数据来自维基百科、WordNet、GeoNames 等网站, 通过将维基百科中的分类体系和 WordNet 的分类体系相融合, YAGO 构建了一个复杂的类别层次结构体系, 2012 年推出的 YAGO2s 拥有超过 1 000 万个实体和超过 1.2 亿个事实. IMDB 是一个关于电影演员、电影、电视节目、电视明星、电子游戏以及电影制作的在线数据库, 资料按类型进行组织.

#### 1.5 分类

1) 按照研究对象的规模分类, 知识图谱可以分为基于单样本的知识图谱和基于样本集的知识图谱, 比如文本上的就可以分为基于单文本的知识图谱和基于文本集的知识图谱, 视觉上可以分为基于单幅图片的知识图谱和基于图片集的知识图谱.

2) 按照研究内容的领域来分类, 知识图谱可以分为一般化知识图谱, 比如百度公司的知心、搜狗公司的知立方, 和领域知识图谱, 比如影视领域的 IMDB、金融领域的文因互联.

3) 按照研究内容来分类, 知识图谱可以分为文本知识图谱、视觉知识图谱、多模态知识图谱, 后面将逐个进行详细介绍.

### 2 文本知识图谱

#### 2.1 文本知识图谱的构建

文本知识图谱指以文本为主要研究内容, 由文本样本构造, 且节点和边均有文本表示的知识图谱. 如图 3 所示, 文本知识图谱在构建中分为信息表示、知识融合、知识加工、知识更新 4 个部分.

信息表示包括知识抽取和知识表示 2 部分, 知识抽取又包括实体抽取、关系抽取和属性抽取, 针对不同的数据有不同的知识抽取方式, 对于结构化的数据一般基于规则进行抽取, 对于非结构化的数据则一般基于学习抽取, 比如使用 SVM 分类或结合使用 RNN 和 CRF. 知识表示现在比较多的使用翻译嵌入(Translating Embedding), 即将知识嵌入到一个低维空间中, 获得知识的对应向量后就可以使用各种数学工具进行分析, 比如一个三元组实例(中国, 首都, 北京) 将其分别换成特征向量  $(A \ B \ C)$  后, 可以研究  $A + B = C$  这种或其他翻译的情况并分析.

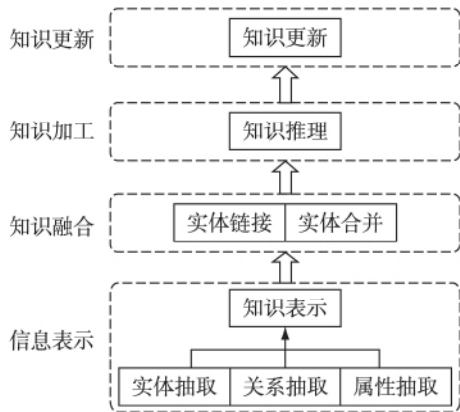


图3 文本知识图谱的构建

Fig. 3 Flowchart of textual knowledge graph construction

2013年 Bordes 等<sup>[9]</sup>提出了基于实体和关系的分布式向量表示的 TransE, 它将每个三元组实例中的关系看作是从头实体到尾实体的翻译(Translating), 并将实体和关系投影到一个平面中。虽然 TransE 简单又高效, 但在自反性上却有致命缺陷。因此, TransH 在 2014 年被提出, 它采用超平面的思路, 直接将关系翻译向量放在关系超平面上, 使实体在不同关系三元组中有不同的角色, 也就是有分布式的表示, 可以进行一对多、多对一和多对多关系的建模<sup>[10]</sup>。2015 年又衍生出了 TransD<sup>[11]</sup>、TransA<sup>[12]</sup>、TransG<sup>[13]</sup>、TransR<sup>[14]</sup>、KG2E<sup>[15]</sup>, 2016 年 Ji 等<sup>[16]</sup>提出了 TranSparse, 利用稀疏矩阵研究知识图谱关系和实体的不均匀性, 作者对不同难度的实体和关系使用不同稀疏程度的矩阵进行表示, 从而防止了对简单关系的过拟合或对复杂关系的欠拟合。

知识融合包括实体链接和实体合并。就文本语义来说, 存在诸如“苹果”既可能指“一种水果”也可能指“苹果公司”这种歧义, 在实体链接部分就要将这种具有歧义的实体链接到给定的确切的知识上, 这一步有时也被称作“实体消歧”。实体合并则是针对一些不同的词汇实际上是一个语义的情况, 将语义相同的实体合并到一起, 比如“贝克汉姆”、“Beckham”、“碧咸”其实指的是同一个人, 具体操作是将多异构的数据源实体归并为一个具有全局唯一标识的实体对象。在判断是否需要合并的实体过程中一般使用基于规则或基于上下文提取词特征向量的方法。

知识图谱并非一开始就是完整的, 而是随着知识加工步骤, 也就是随着知识推理和知识更新步骤来进行补全、扩充的, 这样既可以丰富知识图谱, 也

可以增强机器的理解力。知识推理是利用现有的知识进行推理, 可以基于模板推理, 也可以利用关系机器学习。关系机器学习中常用的方法有矩阵分解、因子图和神经网络等。知识更新则是对新来的知识进行处理, 可以分为全面更新和增量更新。全面更新需要根据新增量重新计算所有的知识表示, 增量更新则只需要将新数据添加到已有的知识表示当中去。Despande 等<sup>[17]</sup>专门写了一篇文章以 Kosmix 知识库和 WalmartLabs 为例, 从工业角度展示了一个真实的知识图谱案例, 详细系统地介绍了如何在数据管理应用方面构建、维护和使用一个知识库, 其中知识数据的补充使用的就是增量更新。

## 2.2 文本知识图谱的应用

1) 文本知识图谱的主要应用是语义检索, 这在谷歌搜索和百度搜索中有鲜明的体现。传统的基于关键词的搜索是将用户给予的输入进行切分得到关键词, 再用关键词和数据进行匹配, 最终将排序后的匹配结果返回给用户以供选择, 这种方法一旦遇到稍微复杂一些的语句就无法理解用户的输入而返回一些不相关的结果。基于知识图谱的语义检索则能更好地理解用户的输入, 并且在输入复杂的情况下也能返回准确的信息, 甚至直接返回答案。用户提供输入后, 语义检索分为以下步骤: 识别输入中的概念、实体、属性和关系, 结合知识图谱对识别的结果进行理解, 在数据集上搜索理解得到的结果并返回信息。比如, 当用户搜索“姚明的身高”, 搜索引擎第 1 条可以直接返回身高的具体数值, 而不只是提供一些相关网页让用户自己选择。

2) 文本知识图谱的另一项应用是深度搜索, 或叫知识导航。相比于传统的基于关键词的搜索方式, 使用知识图谱的深度搜索可以不只是提供用户输入的相关答案, 还能提供用户输入的相关深度信息, 也就是提供相关的知识导航供用户了解学习。比如在搜索“罗志祥”后, 搜索引擎会在右侧栏显示人物相关的知识卡片, 提供用户输入的“实体”、“概念”的相关关联信息, 比如“与罗志祥合作过的艺人”、“罗志祥主演的作品”、“台湾歌手”等。在搜索电影时, 同样, 相关的影视作品、相关的主演以及电影的相关武器都会被直接显示出来供用户选择。

3) 文本知识图谱在情报分析方面也有广泛的应用。在股票投研中, 可以通过建立公司间的知识图谱以供券商分析师等进行深层次情报分析并进行更好的决策。公安人员可以利用企业和个人的资金交



易、出行、住宿、税务等信息建立“公司-人-资金帐户”的知识图谱,辅助进行刑侦、线索侦查、同伙挖掘等。通过检测来自不同数据源信息构建的知识图谱的一致性可以识别潜在的欺诈风险,也就是可以进行反欺诈情报分析等。

### 3 视觉知识图谱

视觉知识图谱是以图像为主要研究内容,由视觉样本构造,且节点和边均有视觉表示的知识图谱。如图4所示,视觉知识图谱由于视觉和图像信息的特点,相比于文本知识图谱在构建中少了知识融合,只有信息表示、知识加工、知识更新3个部分,这3部分的步骤和文本知识图谱的构建没有差别,只是研究对象和方法有所不同。

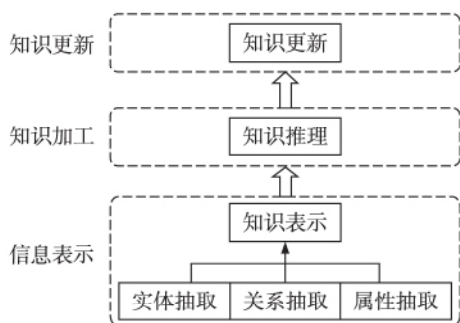


图4 视觉知识图谱的构建

Fig. 4 Flowchart of visual knowledge graph construction

视觉知识图谱在实体抽取上存在以下几个难点:实体列表难以获取、实体对应的视觉图像难以获取、一词多义(也就是需要实体链接)、概念主体在图片上的位置未知。实体列表一般采取预定义或数据挖掘的方式获得,实体对应的视觉图片一般利用搜索引擎的返回图片筛选得到,一词多义可以使用聚类解决。概念主体在图片上的位置则需要用到物体检测。2013年提出的NEIL<sup>[18]</sup>通过对所有的图片进行物体检测,再通过聚类找到每个实体的聚类中心点,最后用训练分类器对所有的实体样本进行分类,这个过程中使用的就是预先定义的实体,所以如果要添加新的图片就需要更新检测器。2015年Johnson等<sup>[19]</sup>提出先对图像中的每一个物体区域和区域内对应的实体进行标注,再利用所有标注区域来训练物体检测器的方法,但这种方法的所有数据都需要完全的手工标注。预先定义的实体也可以不很具体,Divvala等<sup>[20]</sup>就使用了预先定义的较为笼统的实体列表,他们利用文本搜索引擎获得丰富的实

体列表,再利用图片搜索引擎得到实体的图片样本,随后对图片进行聚类、筛选实体,最后训练分类器。

视觉内容的关系抽取方面的难点包括实体间的关系复杂,难以建模;实体间关系的视觉样本呈现多样化;概念主体在图片上的位置未知。针对复杂的视觉实体间关系,Chen等<sup>[18]</sup>使用了混淆矩阵(Confuse Matrix)对关系进行建模,但这种方法只能抽取较为简单的关系。Sadeghi等<sup>[21]</sup>提出了视觉知识提取系统(Visual Knowledge Extraction, VisKE),使用了预定义主谓宾关系的方法,利用关系短语中名词间的动词分析文本和图像,研究其空间一致性。关系建模过程中,作者用搜索引擎分别检索不同的关系结构,对每一类搜索结构训练一个分类器,最后构建了一个因子图来代表主谓宾的视觉关系。针对一般化的视觉关系,Lu等<sup>[22]</sup>结合视觉特征和语言模型先验知识(即可能的视觉关系)来确定视觉关系,可以实现一张图中探测多种视觉关系,并且在只有几张训练样本的前提下也能探测,甚至是用于zero-shot预测中。但该方法不仅需要质量较高的标注数据,而且只能表示单幅图像中的2个物体对应的关系。

在视觉属性抽取上,对于形状、颜色等物体属性,Krishna等<sup>[23]</sup>对每一类属性都训练了一组分类器。对于整幅图像的属性,可以使用深度卷积网络和评价损失函数相结合的方法抽取,Souri等<sup>[24]</sup>在2016年提出了一种预测图像属性相关性的方法,结合使用了ConvNet和ranking layer,作者使用ConvNet来学习图像特征,ranking layer用来给图像评分。该方法在各种或粗糙或精细的数据集上,在相关属性预测的实验中都取得了很好的效果。

对于视觉内容,因为在信息抽取时就已经将其转化为视觉特征向量,所以不需要显示特征嵌入,可以直接将视觉特征作为信息表示,也可以将所有特征组合成因子图的形式进行知识表示。视觉内容的知识推理可以使用基于视觉特征的标签传播,也就是根据相似性矩阵进行标签传播。Lu等<sup>[22]</sup>在进行视觉知识推理时则是使用基于检测和分类模型的知识补全,也就是将未知的实体、关系、属性等视觉内容提取视觉特征后送到对应的检测和分类模型中去预测结果。进行视觉方面的知识更新时,如果新来的数据没有标注,就需要重新进行实体抽取、关系抽取和属性抽取步骤,如文献[18];如果新来的数据有标注,就可以根据标注数据更新视觉实体、关系、属性及其影响到的推理。

视觉知识图谱的主要应用是语义图像检索,即根据用户的描述检索图像.传统的图像检索也是基于关键词的检索,类似于文本知识图谱,在进行基于视觉知识图谱的语义图像检索时,通过先分析输入的“实体”、“属性”,理解输入的含义结果,再利用含义结果进行图像检索,可以为用户返回更加相符和准确的目标图像.

利用视觉知识图谱,还可以进行文本关系真假的判断.一般来说,对于绝大部分文本描述的真实关系都可以找到相匹配的图像与之对应,在进行文本关系真假判断时,如果对应的图像并不存在或与同类图像的一致性不符,就可以判断该文本关系为假,相反则为真.

#### 4 多模态知识图谱

如图5和视觉知识图谱基本相似,多模态知识图谱在构建中分为信息表示、知识加工、知识更新3个部分.

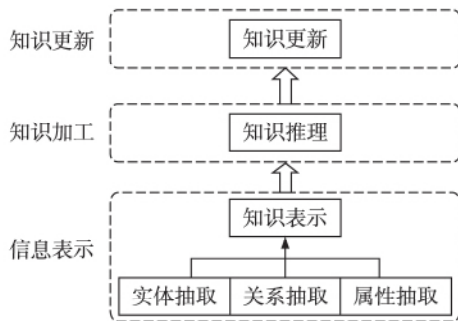


图5 多模态知识图谱的构建

Fig. 5 Flowchart of multi-modal knowledge graph construction

实体抽取一般是从多模态样本中自动抽取实体列表.Sun等<sup>[25]</sup>提出利用相似文本和视觉集合来自动提取视觉概念的方法.在视觉概念挖掘中,提供一组相似的图像和文本描述后,作者首先挖掘文本数据选出候选的概念.因为视觉世界和人类语言的丰富性,候选概念池会非常大.然后用视觉数据过滤那些没有视觉可辨识度的措辞,最后整合留下来的措辞到简洁的概念聚类中去,并且使用语义相似度、视觉相似度进行评分.作者在文本过滤时使用视觉上相关图像的区别,组织文本时使用视觉和文本的相似性.这种方法在图像和文本的双向概念获取任务、图像标记任务上的表现都非常好.

Fang等<sup>[26]</sup>则提出了一个基于公众分类图像来自动建立视觉实体的框架,作者根据文本和视觉的

聚类结果来确定最终的视觉概念.利用大规模的用户生成的图像自动建立视觉实体有3个关键方面:概念搜索、概念关系提取、概念分级建立.概念搜索时,作者基于维基百科从Flickr的标签中分辨筛选概念集.对那些已分辨得到的概念,就使用相关标签的图学习出模型,这样可以通过扩展和识别新图来自动更新结构化实体.概念关系提取时,不能直接获取的概念关系先利用视觉模范相似度和标签一致性提取出来.之后,类别关系就可以直接通过验证频度差异、概念标签的分布计算出来.建立概念分级时,作者通过计算概念熵来评估概念的语义广度.该框架能有效应对用户生成的噪声标签,通过利用文本和视觉两方面的信息实现了图片和概念的不停更新.

多模态样本中的关系分为同时出现关系和层级关系,抽取关系时一般利用通用的概念比具体的概念出现频率要高这一思想,通过计算实体的文本和图片特征的统计关系进行抽取.Fang等<sup>[26]</sup>就通过验证频度差异和概念标签的分布来计算类别关系.目前没有专门研究多模态属性抽取的提取方法,一般是将属性当作实体概念的一种,和实体抽取采用相同的方法.

多模态样本的知识推理可以使用基于多模态特征的标签传播,比如Fang等<sup>[26]</sup>就根据相似矩阵和图片相似矩阵进行标签传播;也可以使用因子图进行推导和学习,比如Zhu等<sup>[27]</sup>就采用马尔科夫随机场结合Gibbs采样学习因子图的权重进行知识推导与学习.

因为多模态知识图谱的每一步构建过程都需要所有的多模态样本,因此如果增加新样本就需要全面更新,目前多模态知识图谱方面还没有增量更新的相关论文.

相比于文本知识图谱和视觉知识图谱,多模态知识图谱在生活中有更加广泛的应用,比如可以实现基于本体的图像检索,也可以做视觉和文本相结合的视觉知识问答.传统的基于分类的视觉问答,由于进行的操作是分类和目标探测,这样对于简单问题能够回答得不错,但只要问题稍稍复杂,回答就不令人满意.目前针对大规模、多样性视觉问题任务的研究,都是将视觉识别任务扔给能进行多样性推理的模型,这种方法最大的好处就是避免了每次一出现新一类问题就得训练新分类器的困扰.

Zhu等<sup>[27]</sup>出于类似目的建立了一种新的大规模

多模态知识库结构.作者通过将整幅图像及其具有的所有文字信息当作一个整体,直接将图片特征及其标注的文本内容作为实体,可以将大规模知识库应用于视觉问答.其好处是知识库一旦建立好就可以处理各种各样的视觉问答,而不用像以前一样为了每一个专门的任务和目的单独训练模型.

## 5 总结

近年来知识图谱逐渐受到了广泛的关注并获得了众多的研究进展,通过将知识图谱应用于诸如语义检索、深度搜索、信息推荐、自动问答等领域,很多以前生活中的设想都已经成为可能.但是知识图谱依然面临着很多问题,尤其是多模态知识图谱方面,比如如何进行多模态的属性表达、如何进行复杂的多模态关系的挖掘和统一表示、如何进行多模态知识图谱的增量更新等.

作为知识工程的重要内容,知识图谱以语义网为理论基础,结合数据挖掘、自然语言处理、机器学习、知识表示等方面的内容,对大数据时代的高效知识管理、知识获取、知识共享有重要的意义.由于文本知识图谱和视觉知识图谱本身的局限性,未来知识图谱的发展趋势主要会在多模态知识图谱上,这方面的主要研究内容包括多模态的概念挖掘、概念的统一表示、概念的简单关系的构建等.

## 参考文献

### References

- [1] 漆桂林,高桓,吴天星.知识图谱研究进展[J].情报工程,2017,3(1):4-25  
QI Guilin,GAO Huan,WU Tianxing. The research advances of knowledge graph[J].Technology Intelligence Engineering,2017,3(1):4-25
- [2] 李涛,王次臣,李华康.知识图谱的发展与构建[J].南京理工大学学报,2017,41(1):22-34  
LI Tao,WANG Cichen,LI Huakang. Development and construction of knowledge graph[J].Journal of Nanjing University of Science and Technology,2017,41(1):22-34
- [3] DB-Engines.DBMS popularity broken down by database model [EB/OL]. [2017-06-28]. [https://db-engines.com/en/ranking\\_categories](https://db-engines.com/en/ranking_categories)
- [4] Bollacker K,Cook R,Tufts P. Freebase: A shared database of structured general human knowledge [C]// Proceedings of the 22nd National Conference on Artificial Intelligence,2007,2:1962-1963
- [5] Wikidata.Main page of wikidata [EB/OL]. [2017-06-28]. [https://www.wikidata.org/wiki/Wikidata:Main\\_Page](https://www.wikidata.org/wiki/Wikidata:Main_Page)
- [6] Bizer C,Lehmann J,Kobilarov G,et al.DBpedia-A crys-
- tallization point for the web of data [J].Web Semantics: Science Services and Agents on the World Wide Web,2009,7(3):154-165
- [7] Suchanek F M,Kasneci G,Weikum G.Yago: A large ontology from wikipedia and wordnet [J].Web Semantics: Science Services and Agents on the World Wide Web,2008,6(3):203-217
- [8] IMDB.Homepage of IMDB [EB/OL]. [2017-06-28]. <http://www.imdb.com>
- [9] Bordes A,Usunier N,Garcia-Duran A,et al.Translating embeddings for modeling multi-relational data [C]// International Conference on Neural Information Processing Systems,2013:2787-2795
- [10] Wang Z,Zhang J W,Feng J L,et al.Knowledge graph embedding by translating on hyperplanes [C]// AAAI Conference on Artificial Intelligence,2014:1112-1119
- [11] Ji G L,He S Z,Xu L H,et al.Knowledge graph embedding via dynamic mapping matrix [C]// Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing,2015,1:687-696
- [12] Xiao H,Huang M L,Hao Y,et al.TransA: An adaptive approach for knowledge graph embedding [J].arXiv e-print,2015,arXiv:1509.05490
- [13] Xiao H,Huang M L,Hao Y,et al.TransG: A generative mixture model for knowledge graph embedding [J].arXiv e-print,2015,arXiv:1509.05488
- [14] Lin Y K,Liu Z Y,Sun M S,et al.Learning entity and relation embeddings for knowledge graph completion [C]// AAAI Conference on Artificial Intelligence,2015:2181-2187
- [15] He S Z,Liu K,Ji G L,et al.Learning to represent knowledge graphs with Gaussian embedding [C]// ACM International on Conference on Information and Knowledge Management,2015:623-632
- [16] Ji G L,Liu K,He S Z,et al.Knowledge graph completion with adaptive sparse transfer matrix [C]// AAAI Conference on Artificial Intelligence,2016:985-991
- [17] Deshpande O,Lamba D S,Tourn M,et al.Building, maintaining, and using knowledge bases: A report from the trenches [C]// ACM SIGMOD International Conference on Management of Data,2013:1209-1220
- [18] Chen X L,Shrivastava A,Gupta A,Neil: Extracting visual knowledge from web data [C]// IEEE International Conference on Computer Vision,2014:1409-1416
- [19] Johnson J,Krishna R,Stark M,et al.Image retrieval using scene graphs [C]// IEEE Conference on Computer Vision and Pattern Recognition,2015:3668-3678
- [20] Divvala S K,Farhadi A,Guestin C.Learning everything about anything: Webly-supervised visual concept learning [C]// IEEE Conference on Computer Vision and Pattern Recognition,2014:3270-3277
- [21] Sadeghi F,Divvala S K,Farhadi A.VisKE: Visual knowledge extraction and question answering by visual verification of relation phrases [C]// IEEE Conference on Computer Vision and Pattern Recognition,2015:1456-1464
- [22] Lu C W,Krishna R,Bernstein M,et al. Visual

- relationship detection with language priors [C] // European Conference on Computer Vision 2016: 852-869
- [23] Krishna R ,Zhu Y K ,Groth O ,et al.Visual genome: Connecting language and vision using crowdsourced dense image annotations [J].International Journal of Computer Vision 2017 ,123( 1) : 32-73
- [24] Souri Y ,Noury E ,Adeli E.Deep relative attributes [C]// Asian Conference on Computer Vision 2016: 118-133
- [25] Sun C ,Gan C ,Nevatia R. Automatic concept discovery from parallel text and visual corpora [C]//IEEE International Conference on Computer Vision 2015: 2596-2604
- [26] Fang Q ,Xu C S ,Sang J T ,et al.Folksonomy-based visual ontology construction and its applications [J]. IEEE Transactions on Multimedia 2016 ,18( 4) : 702-713
- [27] Zhu Y K ,Zhang C ,Ré C ,et al.Building a large-scale multimodal knowledge base system for answering visual queries [J].arXiv e-print 2015 ,arXiv: 1507. 05670

## Research progress on development and construction of knowledge graph

ZHU Muyijie<sup>1</sup> BAO Bingkun<sup>1</sup> XU Changsheng<sup>1</sup>

<sup>1</sup> National Lab of Pattern Recognition ,Institute of Automation ,Chinese Academy of Sciences ,Beijing 100190

**Abstract** Knowledge graph technology is widely concerned and studied during recent years ,in this paper we introduce the construction methods ,recent development of knowledge graph in details ,we also summarize the interdisciplinary applications of knowledge graph and future directions of research.This paper details the key technologies of textual ,visual and multi-modal knowledge graph ,such as information extraction ,knowledge fusion and knowledge representation.As an important part of the knowledge engineering ,knowledge graph ,especially the development of multi-modal knowledge graph ,is of great significance for efficient knowledge management ,knowledge acquisition and knowledge sharing in the era of big data.

**Key words** knowledge graph; semantic network; natural language processing; visual knowledge graph; multi-modal