

家庭服务机器人领域知识图谱构建与应用

吴培良^{1,2}, 王天成^{1,2}, 金鑫龙^{1,2}, 闫鹏宇³, 张云川^{1,2}, 陈雯柏⁴,
毛秉毅^{1,2}, 高国伟⁴

(1. 燕山大学 信息科学与工程学院, 河北 秦皇岛 066004;

2. 河北省计算机虚拟技术与系统集成重点实验室, 河北 秦皇岛 066004;

3. 上海工业自动化仪表研究院有限公司, 上海 200233;

4. 北京信息科技大学 自动化学院, 北京 100192)

摘要:对于家庭服务机器人,能否准确快速地获取到家庭场景中实体的语义信息是决定其智能化水平的关键。为了增强其语义信息获取能力与知识推理能力,针对家庭场景提出了一种面向服务机器人的领域知识图谱自动化构建流程。首先,利用词频-逆向文件频率算法(Term Frequency - Inverse Document Frequency, TF-IDF)从文本信息中提取服务策略关键字,构建服务策略图谱;其次,通过预训练的場景分割模型(Scene Segmentation Model, SSM)识别出场景内的实体;之后,根据当前场景的实体信息,运用场景分类模型(Scene Classification Model, SCM)来预测房间类别,生成结构化数据;再次,根据结构化数据构建家庭实体图谱;最后,将服务策略图谱与家庭实体图谱合并为家庭服务领域知识图谱,并将其存储到 neo4j 图数据库中。实验结果表明,所提出的方法可以根据非结构化数据自动生成领域知识图谱,通过查询知识图谱,检索家庭服务所需的语义信息,可以帮助机器人生成适用于当前工作环境的服务策略,使其更加智能地完成服务任务,证明了方法的可行性。

关键词:家庭服务机器人;服务策略;语义信息;知识图谱;深度学习

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2023)08-0172-08

doi:10.3969/j.issn.1673-629X.2023.08.025

Domain Knowledge Graph Construction and Application of Home Service Robot

WU Pei-liang^{1,2}, WANG Tian-cheng^{1,2}, JIN Xin-long^{1,2}, YAN Peng-yu³, ZHANG Yun-chuan^{1,2},
CHEN Wen-bai⁴, MAO Bing-yi^{1,2}, GAO Guo-wei⁴

(1. School of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, China;

2. Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province,

Qinhuangdao 066004, China;

3. Shanghai Industrial Automation Instrument Research Institute, Shanghai 200233, China;

4. School of Automation, Beijing Information Science and Technology University, Beijing 100192, China)

Abstract: For the home service robot, whether it can accurately and quickly obtain the semantic information of the entity in the home scene is the key to determine its intelligence level. In order to enhance its semantic information acquisition and knowledge reasoning ability, we propose an automatic construction process of knowledge graph based on family scene. Firstly, TF-IDF algorithm is used to extract service policy keywords from text information and construct the service policy graph. Secondly, the pre-trained scene segmentation model (SSM) is used to identify the entities in the scene, and the scene classification model (SCM) is used to predict the room category and generate structured data. Then, the family entity map is constructed according to the structured data. Finally, the service policy graph and the family entity graph are merged into the family service knowledge graph and stored in the Neo4J graph database. The experimental results show that the proposed method can automatically generate domain knowledge graph according to un-

收稿日期:2023-01-13

修回日期:2023-04-17

基金项目:国家自然科学基金项目(62276028, U20A201584);河北省自然科学基金项目(F2021203079);河北省创新能力提升计划项目(22567626H)

作者简介:吴培良(1981-),男,研究方向为机器学习、多智能体系统;通信作者:王天成(1998-),男,研究方向为知识图谱。

structured data. By querying the knowledge graph, the semantic information needed by service can be searched, and the service policy suitable for the current working environment can be generated by the robot, the robot can complete the service more intelligently, which proves the effectiveness of the proposed method.

Key words: home service robot; service policy; semantic information; knowledge graph; deep learning

0 引言

随着人口老龄化的加剧和用工成本的上涨,家庭服务机器人的市场需求在不断扩大。在2021年的德国国际机器人联合会(International Federation of Robotics)发表的世界机器人-服务类机器人的报告中指出:全球市场在专业型服务机器人领域的营业额已经达到67亿美元,同比过去的2020年增长率可达到12%;专业性服务机器人中细分的家居服务机器人的营业额同比增长可达到16%,累计44亿美元^[1]。由此可见,家庭服务机器人拥有迅猛的发展势头和巨大的市场前景。

然而,当前服务机器人的智能化程度还不能够满足室内场景下人机交互的需求,主要原因为当前的环境信息表示方法无法使机器人产生对家庭环境的深入认知。目前,常用的环境信息表示方法有谓词逻辑表示法^[2]、产生式规则表示法^[3]和语义网本体表示法^[4]。这些环境信息表示方法努力增强了对环境信息的表达,但是对信息记录的同时,忽略了各信息之间的关系,使得机器人在使用已获取的信息时存在一定的困难,导致其无法智能化地执行服务任务。例如,机器人要执行“倒垃圾”任务,那么它不仅需要知道垃圾与垃圾桶的位置信息,还需要根据它们的功能来判断它们与任务之间的关联。

近年来,知识图谱的研究受到越来越多的关注。知识图谱的研究价值集中地体现在它是实现认知智能的基础。“理解”和“解释”是机器认知智能的两个核心能力,二者均与知识图谱有着密切关系^[5]。“理解”视作建立从数据(包括文本、图片、语音、视频等数据)到知识图谱中的实体、概念、属性之间映射的过程。这一过程的本质就是将知识图谱中的知识与问题或者数据相关联。有了知识图谱,机器有望可以重现人类的这种理解与解释的过程。

当前,研究人员已经针对国防、能源、金融等领域展开知识图谱研究,但针对家庭服务机器人的领域知识图谱构建方面的研究十分稀少。为此,该文提出了一种家庭服务领域知识图谱的自动化构建方法。基于深度学习的神经网络框架实现场景分割,获取家庭场景中实体的语义信息;基于知识图谱的共享性与拓展性构建知识图谱,实现机器人的推理功能。通过知识图谱领域知识与家庭服务场景机器人的场景分割识别逻辑相结合,完成交叉学科跨领域知识运用,进一步提

高家庭服务机器人在感知与认知部分的可靠性与适应性。

1 相关研究

1.1 知识图谱与环境信息表达

2012年,知识图谱(Knowledge Graph)的概念由Google公司正式提出,旨在实现更智能的搜索,并且于2013年后开始在学术和业界普及,在智能问答^[6]、智能推荐^[7]、反欺诈^[8]等领域中发挥重要作用。由于其出色的环境信息表达能力,在其它领域知识图谱也得到了广泛应用。在自动驾驶领域,自动驾驶场景的实体及其关系表示通过知识图谱构建,生成的场景语义网络可用于实体预测的任务,从而提高驾驶过程中对于周边场景的理解^[9];在图像遥感领域,利用遥感知知识图谱推理与深度数据学习的融合,来提升遥感影像的解译性能^[10];在社会交通领域,基于开源数据构建交通知识图谱和事理图谱,利用自然语言处理技术对网络文本进行分析,从而实现知识图谱对于网络交通事件说明的识别^[11]。

由此可知,知识图谱拥有良好的环境信息表达能力,能够帮助智能体理解环境中实体之间的关系。对于家庭服务领域,知识图谱可以帮助机器人更好地理解用户、实体、服务任务之间的关系,使得其可以更加准确地完成有关服务任务的逻辑推理。

1.2 室内场景识别

一般地,家庭服务机器人通过摄像设备进行信息采集,通过对于场景图片的处理来获取场景的物品参数。如何对场景图片进行快速有效的信息抽取,是亟需解决的重要环节。目前常用的方法主要有基于物品识别实现场景信息汇总以及基于场景分割实现场景物品分模块识别。其中,基于物品识别实现场景信息汇总可用的方法常为级联分类器框架^[12]、卷积神经网络^[13]以及基于回归的YOLO(You Only Look Once)算法^[14]。基于场景分割实现场景物品分模块识别则是采用神经网络模型进行深度学习。

在家庭场景中,实体的形态特征在画面中因摆设位置不同将会发生形体抽象变形,并因为物品间覆盖和放置大小差异,从而使得物品无法被有效地进行多角度特征抽取。通过YOLO物品检测算法进行物品识别的准确率最高只能达到42.5%。因此,对于家庭场景,信息理解的重心将放在基于场景分割实现场景

物品的分模块识别。Zhou Liguang^[15]和 Miao Bo^[16]等皆采用场景分割的思路来进行场景中实体的识别。

2 家庭服务场景知识图谱构建

由于各个家庭的环境具有较大差异,所以家庭服务机器人应该现场采集数据,构建家庭服务领域知识图谱,基于此,设计了如下自动构建流程:

(1) 获取服务策略的文本信息,通过词频-逆向文件频率算法 (Term Frequency - Inverse Document Frequency, TF-IDF) 提取服务策略关键字,构成服务策略图谱。

(2) 家庭服务机器人现场采集家庭场景图片,将其输入到训练好的场景分割模型 (Scene Segmentation Model, SSM) 中,得到场景中的实体信息。

(3) 将实体信息输入到场景分类模型 (Scene Classification Model, SCM) 中,预测当前场景类别,如厨房、卧室、餐厅等。

(4) 将 SSM 模型得到的物品信息和 SCM 模型得到的场景类别进行信息汇总,得到对应的文本文件。得到家庭实体知识图谱。

(5) 将服务策略图谱与家庭实体图谱进行拼接,得到家庭服务知识图谱。

家庭服务场景系统框架如图 1 所示。

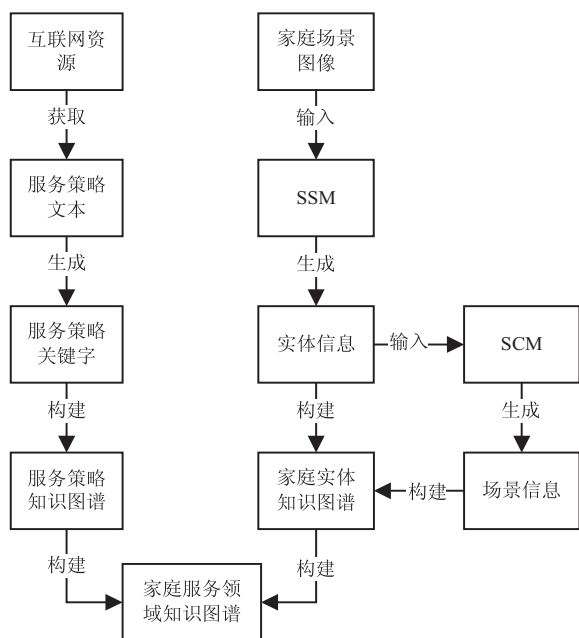


图 1 系统框架

2.1 服务策略知识图谱构建

服务策略文本主要来源于 WikiHow^[17]网站,它是一个开放的互联网网站,包括多个与家庭服务任务相关的知识模块,例如兴趣与手艺、家居与园艺、饮食与休闲等,具体服务知识包括如何洗碗、如何做饭等。该网站的信息利用半结构化文本表示,易于提取,故采用

该网站的文本数据作为服务策略文本。

之后,利用 TF-IDF 算法^[18]对策略文本进行处理,生成服务策略关键字。首先,根据公式(1)计算词频 (Term Frequency, TF)。

$$tf_{ij} = \frac{n_{ij}}{\sum_k n_{kj}} \quad (1)$$

其中, n_{ij} 表示该词出现的次数,分母部分表示文本 d_j 中所有词汇的数量, tf_{ij} 表示文本 d_j 中第 i 个关键词出现的频率,这个值越大,代表这个词出现的频率越高。

之后,根据公式(2)计算逆向文件词频 (Inverse Document Frequency, IDF)。

$$idf_i = \log \frac{|D|}{|\{j:t_i \in d_j\}|} \quad (2)$$

其中, $|D|$ 是文件总数, $|\{j:t_i \in d_j\}|$ 表示包含该词语的文件总数,包含该词语的文档越少, idf_i 值越大,表示该词语具有良好的类别区分能力。

最后,如公式(3)所示,将 tf_{ij} 与 idf_i 相乘,即可得到该词条的权值 w_{ij} ,数值越大,代表该词汇越关键。

$$w_{ij} = tf_{ij} \times idf_i \quad (3)$$

得到服务步骤的关键字后,利用自然语言处理工具包^[19] (Natural Language Toolkit, NLTK) 对原文本进行处理,根据词性标注判断该关键字是动作还是物品词条,将其输入到知识图谱中,构成服务策略知识图谱。

2.2 场景分割模型 (SSM)

在构建家庭服务场景知识图谱时,首先需要利用 SSM 模型对场景内的实体进行分割,并将其识别出来。

2.2.1 ADE20K 数据集

SSM 模型采用 ADE20K 进行训练, ADE20K 是一个密集注释的数据集,总共有 25K 张复杂的日常场景的图像,其中包含日常生活中常见的 150 种物体类别,平均每个图像有 19.5 个实例和 10.5 个对象类。数据集根据训练集和检验集分成两大部分,每部分中根据 a-z 的相同首字母进行场景分类。

2.2.2 场景编码器 (Encoder): ResNet50

ResNet50 为残差神经网络,50 代表该残差神经网络层数。为了解决梯度消失与梯度爆炸问题, ResNet 提出了两种映射:一种是恒等映射,另一种残差映射^[20],将两种映射组合,最后的输出是 $y = F(x) + x$ 。其中,恒等映射是指本身,也就是公式中的 x ,而残差映射指的是 $y - x$,所以残差指的就是 $F(x)$ 部分。残差网络的组合单元如图 2 所示。

将输入数据记为 x ,则中间网络层对于数据的操作可以记为 $F(x)$,通过残差网络的构件输出结果为 $F(x) + x$,即将残留下来的数据与原数据合并,再进行

运算输出,故该模型称为残差神经网络。这时需要考虑对 $F(x)$ 和 x 的数据格式进行分类讨论,如果 $F(x)$ 的数据通道数与 x 相同则输出结果见式(4)。

$$y = F(x, \{W_i\}) + x \quad (4)$$

其中, $F(x, \{W_i\})$ 指的是中间网络层对于数据使用权值 ω_i 进行计算, x 为原输入数据。 y 为最终的输出结果。另一种情况为 $F(x)$ 的数据通道数与 x 不同,则需要将 x 原数据转换成相同的数据格式,结果见式(5)。

$$y = F(x, \{W_i\}) + W_s x \quad (5)$$

其中, W_s 为通过权值的计算将 x 为原输入数据转化为相同的数据通道数量。 y 为最终的输出结果。

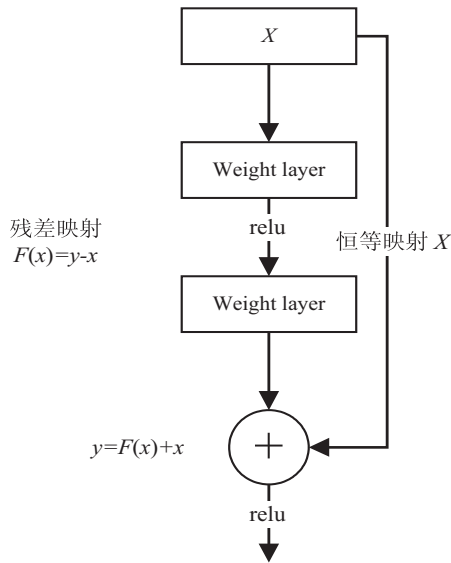


图2 残差网络组合单元(示例)

在家庭场景分割的过程中,通过 ResNet50 将 ADE20K 数据集的图片转化成物品的区域预测概率序列,将图像编码成高维特征区块的物品种类分布,以此来实现家庭服务场景的编码过程。

2.2.3 场景解码器(Decoder):PPM

高维特征区块物品种类预测序列并不是最终的数据结果,需要对该序列进行反向编码,即解码的过程。将高维特征还原成对应的原图片区域,并根据该种类查询对应色卡进行上色,完成对于识别的物品种类在图中进行标记。为了实现这一解码过程,使用场景解码器:金字塔池化模型(Pyramid Pooling Module, PPM)。

采用 1×1 、 2×2 、 3×3 和 6×6 四种不同尺寸的卷积核,对 ResNet50 获取到的特征图进行卷积。得到多个尺寸的特征图,并对这些不同尺寸的特征图再次进行 1×1 的卷积操作来减少通道数。然后采用双线性插值进行上采样,以此获得金字塔模块前相同尺寸的特征图,并在通道上进行拼接,将不同层级的特征图拼接为最终的金字塔池化全局特征。从直觉上来看,这种多尺度的池化确实可以在不同的尺度下保留全局信息,

比起普通的单一池化更能保留全局上下文信息。

2.3 场景分类模型(SCM)

家庭场景中实体分布具有一定的规律,如床大概率出现在卧室、沙发大概率出现在客厅、厨房中很可能存在厨具等。故而可以根据 SSM 模型识别出的实体来对当前所处的场景进行预测分类。

2.3.1 PLACE365 数据集

PLACE365 数据集有 365 个场景类别,用于各类场景识别的比赛和训练。该家庭服务场景系统所作用的场景为家庭常见环境,因此选取 PLACE365 - standard 数据集中的浴室(bathroom)、卧室(bedroom)、衣帽间(closet)、餐厅(dining room)、书房(home office)、厨房(kitchen)和客厅(living room)共 7 个数据集进行分类测试。

通过训练好的 SSM 模型生成家庭功能区的标签和功能区中的前 10 种像素占比大于 0.1% 的物品(不足 10 种补空物品样本)样本数据集,用于后续 SCM 模型进行场景分类的训练。

2.3.2 场景数据预处理与标签加载

由 SSM 模型生成的数据格式可得,每个检测出来的物品都是通过字典的形式进行存储,字典中包含对应物品的中文名、英文名和功能关键词,该文只需要选取物品的英文名作为每条数据样本的属性即可。

模型的数据属性一般要求是数字格式,所以需要物品列表文件进行反向字典生成,即通过物品的英文名将物品的标签还原,让每一条场景数据都是 10 个标签数字组成的数字序列(如果检测物品不足 10 个,则对应的物品标签用 -1 代替)。

在完成对 7 个场景类别图片的识别后,将文本文件用于基于梯度提升决策树(Gradient Boosted Decision Tree,GBDT)算法的 SCM。其中,GBDT 算法是通过梯度提升的方式不断地优化参数,直到达到局部最优解。梯度提升算法每一步优化的对象是不同的。第一步是以 label 值为目标来优化参数,第二步优化的目标就是第一步的残差,第三步优化的目标是第二步的残差。以此类推,不断地对权重进行调整,是一种带权的决策树形式。

2.4 模型测试

2.4.1 评价指标

SSM 与 SCM 的性能通过响应速度、准确率(Accuracy)、F1-score 和 Cohen's Kappa 等参数来进行判定。

其中, F1 - score 是基于准确率、精确率(Precision)、召回率(Recall)加权调和后得到的。以二分类问题为例,最终结果应当被分为正类或者负类。如果原标签标记为正类,并且分类结果为正类,则记为

TP (True Positive)。同理,负类标签与分类结果一致就可以得到 FN (False Negative)。TP 和 FN 都是分类正确的结果。反之,正类被预测为负类、负类被预测为正类,分别记为 TN (True Negative) 和 FP (False Positive),TN 和 FP 是分类错误的结果。根据上述得到的四种分类情况,可以得出精确率、准确率、召回率、F1-score 的计算方式,具体见式(6)~式(9)。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (6)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (7)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

除了上述的性能参数外,再引入 Cohen's Kappa 进一步检验模型的分类精度与性能。当标签数据存在样本失衡的情况时,准确率就无法反映出模型的分类效果。因而需要引入一个基础准确率(P_e)来均衡这种样本数据失衡的情况。基础准确率计算公式见式(10)。

$$P_e = \frac{1}{N^2} \sum_{k=1}^N n_{k1} n_{k2} \quad (10)$$

其中, k 表示第 k 个类别, n_{k1} 和 n_{k2} 的乘积表示为真实占比与预测占比的乘积。在得出基础准确率后,就可以得出 Cohen's Kappa 的数值,数值越大代表分类效果越好,具体如式(11)所示。

$$\text{Cohen's Kappa} = 1 - \frac{1 - \text{Accuracy}}{1 - P_e} \quad (11)$$

2.4.2 SSM 模型测试

通过场景编码器 ResNet50 进行场景解析物品的分类,然后场景解码器 PPM 充分利用好场景的上下文信息,将场景中的物品解析得更加准确与全面。到此为止,通过场景编码器和场景解码器完成场景分割识别物品的功能。训练模型最终像素准确率为 80.13%,总分为 61.14,响应速度为 2.6 s。与其他训练模型的结果对比如表 1 所示^[21]。

表 1 SSM 模型训练结果统计

| 模型 | Accuracy /% | F1-score | 响应速度 /s |
|-------------------------|----------------|----------|------------|
| MobileNetV2+C1_deepsup | 76.80 | 55.32 | 10.3 |
| MobileNetV2+PPM_deepsup | 78.26 | 57.27 | 6.7 |
| ResNet18+C1 | 77.41 | 56.38 | 5.8 |
| ResNet18+PPM | 79.29 | 59.05 | 4.2 |
| ResNet50+PPM | 80.13 | 61.14 | 2.6 |

通过测试结果可知,SSM 模型性能良好,与其它场景分割模型相比,能够较为准确地获取场景中的信

息,且图像处理速度快,能够满足家庭服务机器人连续采集并处理场景图像信息的需求。

2.4.3 SCM 模型测试

选取 PLACE365 数据集中 7 种场景的部分数据集作为测试集,通过 sklearn 的机器学习库选取多种机器学习算法与 GBDT 算法进行数据的训练与预测比较,使用 Accuracy 和 Cohen's Kappa 系数来进行各模型的评估。从模型对比可得,GBDT 算法能够在使用 10 个物品特征的情况下,取得较好的场景分类效果。在 SSM 的像素识别准确率为 80.13% 的情况下,将场景分类的准确率维持在 79%。因此选择该模型作为分类模型使用。各模型训练后,对测试集进行测试的结果如表 2 所示^[22]。

表 2 SCM 模型训练结果统计表

| 模型名称 | 最佳属性数 | Accuracy /% | Cohen's Kappa |
|------------------------------|-------|----------------|------------------|
| KNeighborsClassifier() | 7 | 0.43 | 0.34 |
| LinearDiscriminantAnalysis() | 10 | 0.30 | 0.19 |
| GaussianNB() | 10 | 0.31 | 0.19 |
| DecisionTreeClassifier() | 7 | 0.63 | 0.57 |
| RandomForestClassifier() | 10 | 0.70 | 0.64 |
| SVC(gamma='scale') | 10 | 0.40 | 0.30 |
| Perceptron(max_iter=2 000) | 10 | 0.22 | 0.09 |
| GBDT | 10 | 0.79 | 0.75 |

2.5 家庭服务领域知识图谱结构设计

通过对家庭服务的场景分析,家庭环境可以划分成若干个房间,每个房间中拥有若干个物品,每个物品拥有若干关联词,关联词可以为功能词汇,也可以是人的需求词汇。基于此,家庭服务机器人便可以通过查询知识图谱来找到服务所需的实体信息。

之后,将构建好的策略图谱与实体图谱合并,生成家庭服务知识图谱,结合三元组的知识,创建家庭服务场景的本体结构。功能区(Room)和家庭(Home)拥有位于(Located in)的关系;工具物品(Object)和功能区(Room)拥有从属(Belong to)的关系;服务任务(Service)和任务步骤(Service Step)拥有方法(Method)的关系;任务步骤(Service Step)和工具物品(Object)拥有动作(Affordance)的关系;功能需求关联词(Function)和工具物品(Object)拥有属于(Owned to)的关系。所构建的知识图谱结构的三元组表示形式如式(12)至式(16)所示。

$$\langle \text{Room}, \text{Located in}, \text{Home} \rangle \quad (12)$$

$$\langle \text{Object}, \text{Belong to}, \text{Room} \rangle \quad (13)$$

$$\langle \text{Service}, \text{Method}, \text{Service Step} \rangle \quad (14)$$

$$\langle \text{Service Step}, \text{Affordance}, \text{Object} \rangle \quad (15)$$

<Function, Owned to, Object> (16)

除了本体结构外,仍需要考虑构建成本体后,每个节点需要的属性值。一个知识图谱数据库中,可能存在多个家庭环境信息,所以 Home 节点需要拥有名称(name)和该家庭的唯一信息标识符(id)。Room 与 object 同理也需要拥有 name 和 id 两个属性。一种服务任务可能有多种策略方法,所以任务与服务步骤之间用 method i ($i = 1, 2 \dots$) 来区别不用策略方法。

以 wash clothes 服务任务为例,家庭服务知识图谱片段如图 3 所示(工具物品的功能节点已隐去)。

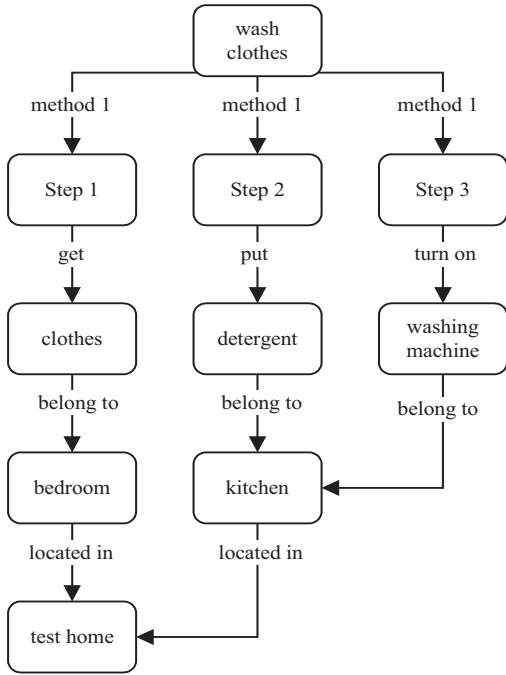


图 3 家庭服务知识图谱片段

3 家庭服务领域知识图谱应用

3.1 实验与应用

为了模拟验证所提方法的有效性,从 PLACES365 数据集中选取 7 张场景图片当作机器人在家庭中采集的图像信息,以此来模拟家庭场景,包含的 7 个常见家庭场景有餐厅、厨房、客厅、书房、卧室、衣帽间和浴室。

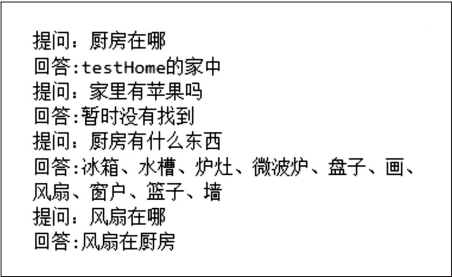


图 4 图谱搜索结果

通过训练好的 SSM 模型与 SCM 模型,获取实验场景图中的实体、场景与功能的语义信息,构建家庭服务领域知识图谱,所构建的知识图谱共包含 337 个实

体,783 条关系。机器人通过搜索知识图谱,可以获取家庭环境中的实体与场景信息,与图谱的交互结果如图 4 所示。通过结果可知,所构建的知识图谱能够帮助机器人对家庭环境产生深入认知。

3.2 反馈与更新

由于家庭环境中的实体是复杂多变的,且构建图谱时可能会匹配错误,所以需要定期对家庭服务领域知识图谱进行实时更新。为了达成这个目的,加入了反馈响应实体关系更新功能,该功能主要通过用户间的交互完成家庭知识图谱数据的更新。即在进行查询后,用户可以对上一步交互结果进行反馈,当接收到用户关于“错”之类的关键词后,通过用户的选择与上一步查询结果中反馈的领域词信息对图谱进行修改,例如增加或删除实体的功能信息,以及修改实体的位置信息等,以此来达到更新知识图谱的效果。反馈更新效果图如图 5 所示。

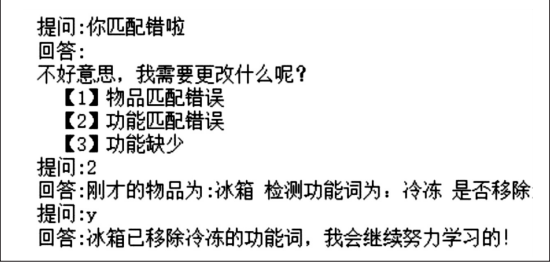


图 5 反馈更新效果

3.3 服务策略生成

当用户对机器人发出任务指令时,机器人可以通过匹配功能需求节点以及任务节点,遍历家庭服务领域知识图谱,选择适合当前环境的服务策略,根据服务任务的步骤、动作、实体以及实体的位置信息来制定相应的策略。生成的服务策略示例如表 3 所示。

表 3 服务策略示例

| 服务指令 | 服务策略 |
|--------------|---|
| Wash clothes | Wash clothes step 1: <get, clothes> <Clothes, belong to, bedroom> |
| | Wash clothes step 2: <put, detergent> <Detergent, belong to, kitchen> |
| | Wash clothes step 3: <turn on, washing machine> <Washing machine, belong to, kitchen> |
| Turn on TV | Turn on TV step 1: <turn on, TV> <TV, belong to, sitting room> |

现有的家庭服务机器人策略生成方法主要有两种:一种为基于语义解析器的方法,从用户指令中获取动作序列,这种方法不适合多物品操作的复杂任务;另一种为基于强化学习的方法,设置基于先验知识的奖

励,让机器人探索生成策略,这种方法没有考虑到当前的环境信息,所给出的服务策略可能包含环境中不存在的物品,导致任务失败。由于所构建的知识图谱能够反映出服务任务与实体之间的关系,所以用户只需下达任务指令,机器人便可以根据家庭环境生成服务策略。与现有的几种策略生成方法对比如表 4 所示。

表 4 服务策略生成方法对比

| 方法 | 多物品操作 | 环境信息 |
|-------------------------|-------|------|
| 语义解析 ^[23-24] | × | ✓ |
| 强化学习 ^[25] | ✓ | × |
| 知识图谱(文中) | ✓ | ✓ |

4 结束语

为了增强家庭服务机器人的信息获取能力与知识推理能力,设计了一种自动构建家庭服务领域知识图谱的方法流程。所构建的 SSM 模型及 SCM 模型能够快速准确地识别出家庭环境中的场景及实体,生成相应的文本文档,将其存储 neo4j 图数据库中,完成家庭服务领域知识图谱的构建。通过对知识图谱进行查询,可以得到服务所需的语义信息,进而生成服务策略。除此之外,还可以通过与用户的问答反馈来实现知识图谱的自更新。

所提出的方法的创新点如下:

(1)将知识图谱应用到家庭服务领域,使得家庭服务机器人拥有了一定的推理能力,可以更加准确地获取服务所需的语义信息。

(2)所提出的家庭服务领域知识图谱为完全自动化构建,只需要机器人在工作区内采集图像便可以自动生成知识图谱,不需要人工参与构建。

除此之外,在家庭服务领域知识图谱研究的过程中,仍存在一些不足之处。未来将通过以下几个方面进行改进:

(1)数据集扩充与模型优化。当前的 ADE20K 数据集中标注的物品为 150 种,且对于家庭环境而言,诸多细小物品并不存在相应的数据。因此,使用该数据训练出的模型识别物品类别数量与此相同,对部分家庭物品无法实现分割与分类。在未来的工作中,可以通过使用场景分割的标注工具,采用相同的数据格式对于 ADE20K 数据集的训练与检验样本进行扩充,以便完成 SSM 与 SCM 的优化。

(2)改进策略生成算法。目前服务策略是基于模板匹配算法生成,模板较为单一,无法处理多元化的服务指令。下一步可以利用所构建的知识图谱辅助强化学习来生成服务策略,使机器人能够更好地利用图谱中的信息,为用户提供更加宜人化的服务。

(3)改用更加真实的实验场景,当前采用场景图像来模拟真实场景,与真实家庭环境相差甚远。未来可以采用 AI2-THOR 等虚拟环境来模拟机器人的工作环境,使其具有更强的实用性。

综上所述,通过系统的研究实现,完成家庭服务领域知识图谱的构建、应用与创新。通过当前系统存在的不足提出未来的发展规划,力求在原基础上不断地更新与完善,来推进家庭服务机器人对于场景信息的感知与处理,从而推动服务型机器人智能化的发展。

参考文献:

- [1] Frankfurt. World robotics 2021 - service robots report released[EB/OL]. 2021. <https://ifr.org/ifr-press-releases/news/service-robots-hit-double-digit-growth-worldwide>.
- [2] BADIA G, COSTA V, DELLUNDE P, et al. Syntactic characterizations of classes of first-order structures in mathematical fuzzy logic[J]. Soft Computing, 2019, 23(7): 2177-2186.
- [3] JONASSEN D H. Changes in knowledge structures from building semantic net versus production rule representations of subject content[J]. Journal of Computer-Based Instruction, 1993, 20(4): 99-106.
- [4] BRENAS J H, SHIN E K, SHABAN-NEJAD A. Adverse childhood experiences ontology for mental health surveillance, research, and evaluation: advanced knowledge representation and semantic web techniques[J]. JMIR Mental Health, 2019, 6(5): e13498.
- [5] 阮利, 温莎莎, 牛易明, 等. 基于可解释基拆解和知识图谱的深度神经网络可视化[J]. 计算机学报, 2021, 44(9): 1786-1805.
- [6] 郑泳智, 朱定局, 吴惠彝, 等. 知识图谱问答领域综述[J]. 计算机系统应用, 2022, 31(4): 1-13.
- [7] GUO Q, ZHUANG F, QIN C, et al. A survey on knowledge graph-based recommender systems[J]. Scientia Sinica Informationis, 2020, 50(7): 937-953.
- [8] 袁俊, 刘国柱, 梁宏涛, 等. 知识图谱在商业银行风控领域的研究与应用综述[J]. 计算机工程与应用, 2022, 58(19): 37-52.
- [9] WICKRAMARACHCHI R, HENSON C, SHETH A. Knowledge-infused learning for entity prediction in driving scenes[J]. Frontiers in Big Data, 2021, 4: 759-772.
- [10] LI Y, KONG D, ZHANG Y, 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.
- [11] TAN J, QIU Q, GUO W, et al. Research on the construction of a knowledge graph and knowledge reasoning model in the field of urban traffic[J]. Sustainability, 2021, 13(6): 31-50.

- [12] MENG Xianpeng, SHANG Chaoyuan, DONG Jian, et al. Automatic modulation classification of noise-like radar intrapulse signals using cascade classifier [J]. ETRI Journal, 2021, 43(6):991-1003.
- [13] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Columbus:IEEE, 2014:580-587.
- [14] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas:IEEE, 2016:779-788.
- [15] ZHOU L, CEN J, WANG X, et al. Borm: Bayesian object relation model for indoor scene recognition [C]//2021 IEEE/RSJ international conference on intelligent robots and systems (IROS). Prague:IEEE, 2021:39-46.
- [16] BO M, LIGUANG Z. Object-to-scene: learning to transfer object knowledge to indoor [C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). [s. l.]:IEEE, 2021:2069-2075.
- [17] 田国会, 陈浩然, 张梦洋, 等. 面向环境适应性的机器人服务策略生成方法 [J]. 华中科技大学学报: 自然科学版, 2023, 51(2):102-108.
- [18] LAN F. Research on text similarity measurement hybrid algorithm with term semantic information and TF-IDF method [J]. Advances in Multimedia, 2022(4):442-452.
- [19] WANG Meng, HU Fanghui. The application of NLTK library for python natural language processing in corpus research [J]. Theory and Practice in Language Studies, 2021, 11(9):1041-1049.
- [20] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas:IEEE, 2016:770-778.
- [21] HANG Z. Pytorch implementation for semantic segmentation/scene parsing on MIT ADE20K dataset [EB/OL]. 2020. <https://github.com/CSAILVision/semantic-seg-mentation-pytorch>.
- [22] 胡栋鹏, 曾坚毅. 数据挖掘常见分类算法比较分析 [J]. 计算机产品与流通, 2018(8):107-110.
- [23] THOMASON J, PADMAKUMAR A, SINAPOV J, et al. Improving grounded natural language understanding through human-robot dialog [C]//Proc of the 2019 international conference on robotics and automation. Montreal:IEEE, 2019:6934-6941.
- [24] WANG Zhongli, TIAN Guohui, SHAO Xuyang. Home service robot task planning using semantic knowledge and probabilistic inference [J]. Knowledge-Based Systems, 2020, 204:106174.
- [25] ZHANG Mengyang, TIAN Guohui, ZHANG Ying, et al. Service skill improvement for home robots: autonomous generation of action sequence based on reinforcement learning [J]. Knowledge-Based Systems, 2021, 212:106605.

(上接第171页)

- [14] JUAN Y, ZHUANG Y, CHIN W S, et al. Field-aware factorization machines for CTR prediction [C]//Proceedings of the 10th ACM conference on recommender systems. [s. l.]:ACM, 2016:43-50.
- [15] RENDLE S. Factorization machines [C]//2010 IEEE international conference on data mining. Sydney:IEEE, 2010:955-1000.
- [16] JÄHRER M, TÖSCHER A, LEE J Y, et al. Ensemble of collaborative filtering and feature engineered models for click through rate prediction [C]//Proceedings of KDD cup 2012 workshop at the 18th ACM. Beijing:ACM, 2012:1-7.
- [17] KETKAR N. Deep learning with python: stochastic gradient descent [M]. Berkeley: Apress, 2017.
- [18] ZHANG X Y, LIU C L, SUEN C Y. Towards robust pattern recognition: a review [J]. Proceedings of the IEEE, 2020, 108(6):894-922.
- [19] VILLARRUBIA G, DE PAZ J F, CHAMOSO P, et al. Artificial neural networks used in optimization problems [J]. Neurocomputing, 2018, 272(10):10-16.
- [20] 史加荣, 王丹, 尚凡华, 等. 随机梯度下降算法研究进展 [J]. 自动化学报, 2021, 47(9):2103-2119.
- [21] 梁杰, 陈嘉豪, 张雪芹, 等. 基于独热编码和卷积神经网络的异常检测 [J]. 清华大学学报: 自然科学版, 2019, 59(7):523-529.
- [22] CHIN W S, ZHUANG Y, JUAN Y C, et al. A learning-rate schedule for stochastic gradient methods to matrix factorization [C]//Pacific-Asia conference on knowledge discovery and data mining. Vietnam:PAKDD, 2015:442-455.
- [23] RENDLE S. Factorization machines with LIBFM [J]. ACM Transactions on Intelligent Systems and Technology, 2012, 3(3):1-22.
- [24] CHIN W S, YUAN B W, YANG M Y, et al. LIBMF: a library for parallel matrix factorization in shared-memory systems [J]. Journal of Machine Learning Research, 2016, 17(1):2971-2975.