

## 基于场景图知识融入与元学习的视觉语言导航

胡成纬 江爱文 王明文  
江西师范大学计算机信息工程学院



## 摘要

视觉语言导航任务涉及到对视觉图像和自然语言的综合理解，以及路径优化，具有较大挑战性，是近年来人工智能领域比较热门的研究课题。目前大多数的研究工作使用模仿学习或强化学习方式来解决的视觉导航问题。视觉语言导航本质上是一个目标搜索定位的问题，机器人对探寻环境的认知是提升算法搜寻效率的重要因素。为了能够自动构建适应于目标环境的先验知识，本文采用场景图生成结合图卷积网络学习的知识图谱构建方式。为了进一步提高模型在测试阶段能适应强化学习中的策略和先验知识，本文加入模型无关的元学习过程，在导航的强化学习框架引入可训练的自监督交互损失，提高模型的泛化推广能力。我们在主流 AI2-THOR 仿真环境中进行实验，相比标准的强化学习方法及仅加入手动构建的知识图谱方法，本文模型具有优越的导航效率，以及良好的知识自动化抽取与构建能力。

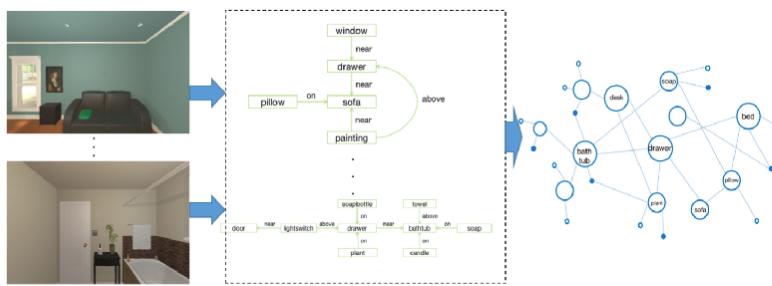


图.1 场景图生成获取具体环境知识并创建知识图谱的过程

我们采用对目标环境进行场景图生成的方式，获取先验知识。通过对每个目标环境进行场景图生成，最终获得相应物体间的关系。这种方式获得的先验知识，能够较好地适应探索环境，尤其当在我们的探索环境（如医院、仓库等）中，物体间的位置关系在日常中很难获取时，使用场景图生成先验知识的方式可以很好的提取到相关信息来适应新场景，并由这些信息创建知识图谱的。

整个算法的过程如图2，在得到先验知识信息后，使用图卷积网络进行处理，其中每个结点的初始化输入的是一个联合特征，由对象的描述特征和当前的观察场景特征组成。

视觉特征和目标对象特征联合先验信息特征，最终拼接得到一条总的联合特征送入A3C网络中。同时为了导航

模型在测试阶段能够更好地适应所学的知识，我们借鉴模型无关的元学习策略MAML，在训练过程中，我们将训练集划分为元训练(meta-training)域和元验证(meta-validation)域。因此总的优化目标表示为如公式所示：

$$\min_{\theta, \phi} L_{a3c}([\theta, \phi] - \alpha_{meta} \nabla_{\theta} L_{\phi}^D(\theta, D), D')$$

## 实验结果

当加入了知识信息的方法后，相比基线模型A3C，的确可以提高导航效率。更明显的，我们的结果在知识信息模型的基础上，使用了元学习的方式，使得在测试阶段更好适应所学到的知识，使得效果有所提升。尤其是，当距离目标对象距离较远时，导航性能提高更为显著，在成功率上提高近5个百分点，SPL也提高了近2个百分点。

算法介绍

我们认为，视觉导航所需的知识同样应该是基于特定的探索环境，而不需要加入所有信息。因此，受到场景图生成的启发，我们对目标环境通过场景图生成的方式，自动获取相应的知识信息用以构建对应的知识图谱，从而获得了特定场景下的导航相关知识。

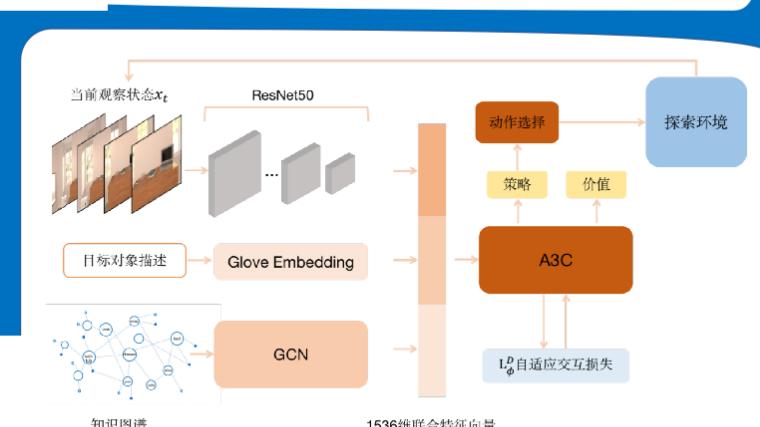


图 2. 本文最终的网络模型框架图。当前观察状态  $x_t$ 、目标对象描述以及场景图知识三者联合特征作为强化学习网络的输入特征。自适应交互损失帮助模型在测试阶段更好的适应所学习的知识。

表 1. 实验结果表.

方法	成功率	SPL	成功率(L≥5)	SPL(L≥5)
随机导航	7.86	3.58	0.27	0.09
强化学习(A3C)	32.89	14.43	21.03	11.24
知识信息[4]	35.01	15.38	21.89	11.68
我们的方法	<b>38.8</b>	<b>16.17</b>	<b>26.95</b>	<b>13.57</b>

表2. 仅使用知识信息的实验结果表.

方法	成功率	SPL	成功率(L≥5)	SPL(L≥5)
知识信息[4]	35.01	15.38	21.89	11.68
我们的方法(仅使用知识)	35.05	15.35	21.83	11.51