

# 具体化视觉导航研究的深度学习调查

朱丰达、李珠、李文胜、梁小丹、张晓军

摘要-“体现视觉导航”问题要求代理在三维环境中导航，主要依赖于其第一人称观察。这一问题由于其在真空吸尘器、救援机器人等领域的广泛应用，近年来越来越受到人们的关注。导航代理应该具有各种智能技能，如视觉感知、映射、规划、探索和推理等。建立一个能够观察、思考和行为的主体，是掌握真正智力的关键。深度学习方法显著的学习能力使主体能够完成具身的视觉导航任务。尽管如此，具体化的视觉导航仍然是如此

在起步阶段，因为需要许多先进的技能，包括感知部分观察到的视觉输入，探索看不见的区域，记忆和建模所看到的场景，理解跨模式指令，适应新环境，等等。近年来，具体化视觉导航引起了社区越来越多的关注，许多作品被提出学习这些技能。本文试图通过提供一个全面的文献调查，来建立一个当前在具身视觉导航领域的工作大纲。我们总结了基准测试和度量标准，回顾了不同的方法，分析了挑战，并强调了最先进的方法。最后，我们讨论了具身视觉导航领域尚未解决的挑战，并为未来的研究提供了广阔的发展方向。

索引术语——深度学习、具体化环境、具体化视觉导航、跨模式导航、导航机器人技术。-F

## 1介绍

利用机器人自动地完成任务 人类是长期研究的课题[2], [3]。一些复杂的应用程序，如真空清洗、禁用帮助和救援，需要代理导航以在3D实体环境的不同位置完成几个子任务。因此，导航是在现实世界中构建智能导航机器人的关键能力之一。在导航过程中，机器人需要通过感知具身视觉输入来四处移动，寻找目标位置，这被称为“具身视觉导航”。通过该环境中的物理实体与环境交互的代理被命名为“具体化代理”[4]。图2演示了一个导航过程。代理首先收到“把客厅里的椅子放到第二个阳台上”的指示。然后它导航以找到目标椅子。特工拿起椅子，导航到阳台上，把它放下。

早期关于机器人导航[5]的工作，[6]主要依靠手工制作的功能，如光流和传统的算法，如马尔可夫定位[7]，增量定位[8]，或地标跟踪[9]。这些方法涉及大量的超参数，不能在看不见的环境中很好地推广。深度学习的最新发展显示了它从大规模数据中学习鲁棒模型的能力。采用端到端深度学习方法训练的视觉机器人更具鲁棒性，具有更少的超参数，并且在不可见的环境中具有更好的泛化能力。

然而，在实现具身视觉导航的深度学习方面需要解决一些挑战：1)从现实世界收集数据代价昂贵；2)从部分观察学习的模型是不稳定的；3)难以学习探索、记忆等长期导航技能；4)由于自然语言格式多样，感知自然语言指令具有挑战性；5)模拟环境与现实环境之间的大域差距阻碍了导航策略的适应等。

本文讨论了机器人导航的相关工作，为构建真实导航机器人提供了一个广阔的方向。本文的结构如图所示。1. 在现实世界中进行训练和测试有许多缺点：1)现实世界中的数据采样效率很低，因为真实机器人只能一次采样轨迹，而模拟器可以在多处理中有效地采样轨迹；2)现实环境的复杂性和动态性决定着再生能力；3)不同环境之间存在较大的域偏差等。三维仿真技术的发展使研究人员能够构建一个模拟环境[10]、[11]、[12]、[13]、[14]，以研究在其中构建一个强大的导航代理。模拟器渲染这些3D资产来生成RGB-D图像，并提供传感器，如物理传感器，GPS传感器来模拟一个现实的具身机器人环境。在模拟环境中学习导航是一个广泛的领域，需要解决许多挑战的问题。在解决目标驱动的导航问题时，研究人员提出了无模型方法[10]、[15]、[16]、自监督方法[17]、[18]、[19]、基于规划的方法[20]、[21]、[22]。由于感知自然语言的多样性和复杂性，因此感知它是一项具有挑战性的任务。它要求代理不仅可以一步一步地遵循一个句子指令[12]，[23]，[24]，而且还可以理解对话[20]，[21]，[22]或导航来回答问题[20]，[21]，[22]。在构建一个真实世界的导航机器人时，一些工作人员[25]，[26]，[27]提议训练一个代理

●朱丰达和李文胜在数据科学系和人工智能，莫纳什大学信息技术学院。

●易珠就职于中国科学院大学。

●梁晓丹，孙大学智能系统工程学院教授中山大学。

●张晓军就职皇家大学计算机技术学院。稿件于2021年7月8日收到。

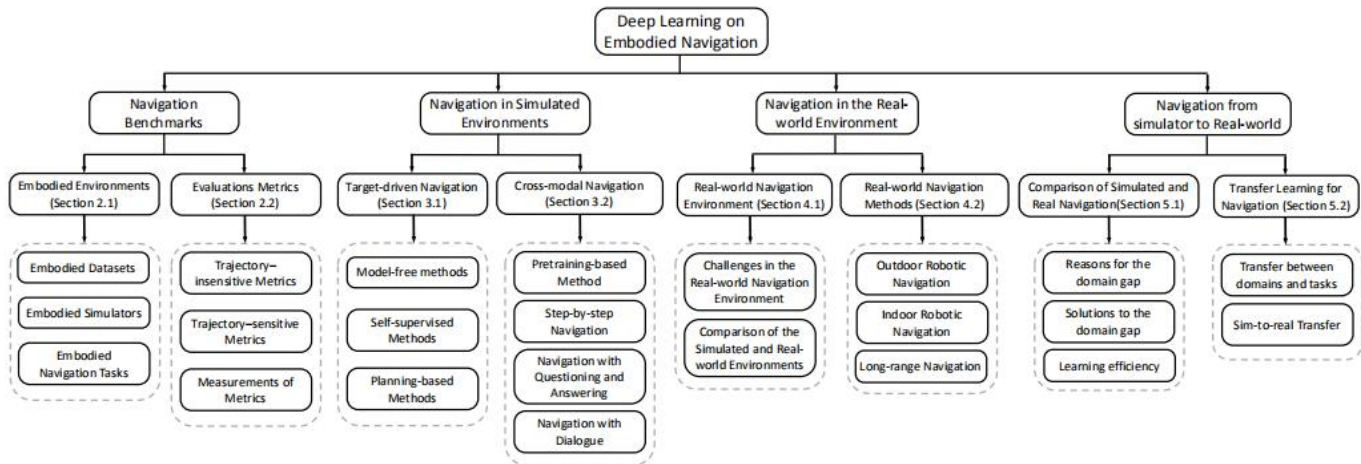


图1：一种用于具体化导航的深度学习方法的分类法。

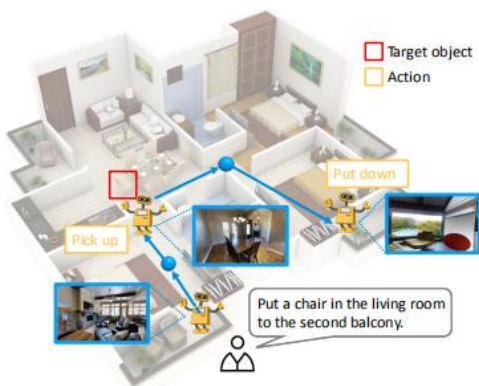


图2：一个导航过程的演示，其中一个机器人移动到几个地方来完成一个任务。

在现实环境中，其他工作[28]，[29]，[30]提出引入迁移学习，将学习到的导航策略从模拟环境转移到现实环境。

与以往对机器人导航[5]，[6]的调查相比，我们的论文主要关注解决具体化导航问题的深度学习方法：

- 1) 据我们所知，我们的论文是第一个全面研究在具体化导航任务的深度学习方法的进展。
- 2) 本文总结并比较了他们对最近提出的具身导航数据集、模拟器和任务的独特见解。
- 3) 本文介绍了具体化视觉导航的深度学习方法，包括它们的动机和贡献。
- 4) 本文对近年来的研究成果进行了分类，并给出了一些很有前景的具身导航方向。

本文的组织结构如下。秒2讨论了当前的具体化的数据集和具体化的模拟器。秒3介绍了包含的导航基准测试，包括导航任务和导航指标。秒4列出了在一个模拟的具体化导航环境中训练代理导航的方法，其中，秒。4.1列出了目标驱动任务和Sec的方法。4.2介绍了跨模态任务的处理方法。秒5

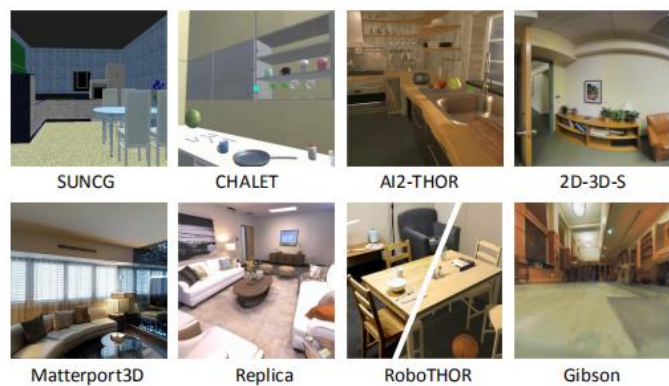


图3：每个数据集的渲染场景。

总结了构建一个导航机器人的工作。秒6说明了模拟环境和真实环境之间的领域差距，并介绍了解决这些挑战的方法。在秒。7、我们强调了最近最先进的作品，讨论了当前工作的局限性，并提出了构建现实世界导航机器人的发展方向。

## 2个具体化的导航环境

在这里，我们将讨论用于具体化导航的环境。我们总结了提供3D资产的数据集和呈现资产并为导航代理提供交互式接口的模拟器。

### 2.1完善的数据集

一个具体化的数据集包含诸如纹理和网格等三维资产用于渲染和其他配置数据，如对象位置、对象类别和高级任务的相机姿态。所提数据集的比较如表1100所示。1.

早期的工作集中于渲染复合RGB视图[31]。它训练了一个概率模型来基于手工创建的场生成合成数据。之后，SceneNet [32]引入了一个生成器模型来注释二维语义。由于深度通道被证明有助于导航代理[19]，[38]，2D-3D-S [33]为资产提供了深度信息。不同于这些同时呈现一个单一房间的作品，后来的作品[34]，[39]，[40]提供了大量的场景，包括卧室，生活

数据集	年	场景	房间	对象目录	RGB	深	二维语义
斯坦福场景*[31]	2012	130	130	—	合成的	7	7
SceneNet*[32]	2016	57	57	218	合成的	7	3
2D-3D-S*[33]	2017	270	270	13	合成的	3	3
SUNCG [34]	2017	45,622	775,574	84	合成的	3	3
小木屋[35]	2018	10	58	150	合成的	7	7
Matterport3D [36]	2017	90	2,056	40	现实的	3	3
吉布森[14]	2018	572	8,854	84	现实的	3	3
副本[37]	2019	18	35	88	现实的	3	3

表1: 现有的具体化数据集的比较 (\*: 数据集仅将一个房间作为场景呈现)。

房间、浴室、厨房等。然而,上述数据集所使用的合成视图与现实世界的场景有很大的不同,这限制了数据集的应用。为此,马特波特3D[36]通过3D重建和这些视图的2D和3D语义提供了逼真的全景视图。Gibson [14]提供了一个更多样化的数据集和572个房子。复制品[37]提出了一个包含18个室内场景的数据集,由密集的网格和高分辨率的纹理组成。一些工作,如AI2-THOR [40],RoboTHOR [39]和小木屋[35],都依赖于目前没有发布的数据集。一些数据集的渲染场景如图所示。3.

### 2.2 体式模拟器

一个具体化的模拟器为代理提供了一个与环境交互的接口。我们在Tab中比较了现有模拟器的不同特征。2. 一个模拟器配备了許多传感器,如RGB传感器、深度传感器、物理传感器和位置传感器。由于3D渲染技术的限制,早期的模拟器提供了较低的RGB分辨率和不现实的图像。由于缺乏视觉细节,限制了代理的导航性能。之后,为了解决这个问题,模拟,如马特波特3D模拟器[12]、Gibson模拟器[14]和栖息地[13]提出了高分辨率的逼真全景视图来模拟一个更真实的环境。渲染帧率对于具体化的模拟器也很重要,因为它对训练效率至关重要。MINOS [11]运行超过100帧每秒(FPS),比它之前的工作快10倍。在512 512 RGB-D图像上超过1000 FPS,使其成为现有模拟器中最快的模拟器。[12]中的离散状态空间简化了导航问题,使代理更容易学习复杂的视觉语言导航任务。然而,连续状态空间更受欢迎,因为它有助于将学习到的代理转移到现实世界的机器人中。一个可定制的模拟器能够通过移动对象、改变对象的纹理和重新配置灯光来生成更多样化的数据。多样的数据几乎没有偏差,因此,使深度学习学习健壮的导航策略。尽管在静态的房间里导航找到目标物体,但互动是现实世界机器人的另一项关键技能。一些复杂的任务可能需要机器人与物体互动,比如拿起杯子、移动椅子或打开一扇门。AI2THOR [40],iGibson [41]和RoboTHOR [39]提供了交互式环境来训练这种技能。多智能体强化学习[42],[43]是一个新兴的代理合作和竞争问题。AI2-THOR和iGibson也支持在研究合作任务时的多智能体训练。

### 3个具体化的导航基准测试

在这里,我们介绍了几个任务来研究具体化的视觉导航问题。这些任务可以分为三个部分

注释: 目标驱动的导航任务、跨模式导航任务和交互式导航任务。

### 3.1 目标驱动的导航任务

**点目标导航**,首先由安德森等人定义。[44]是一个任务,一个代理被初始化为一个随机的起始位置和方向,然后被要求导航到一个目标位置。目标位置由其与起始位置的相对坐标表示。这个任务需要代理估计距离起始位置的累积距离,以便代理知道距离目标的距离。理论上,这项任务能够应用于所有具体化的环境。

**目标目标导航**是由Zhu等人提出的。[15].在这个任务中,一个代理被初始化为一个随机的起始位置,并且需要找到一个特定的对象,比如一张桌子或一张床。一旦导航代理找到了该对象,它就会停止。如果代理位于与目标对象的距离内,则认为导航过程成功。除了房间结构外,目标目标导航任务还需要对象标签和位置。目标识别和探索是目标目标导航的关键技能。**屋顶目标导航**是由Wu等人提出的。[10].在这个任务中,一个在一个随机位置初始化的代理被要求导航到一个房间(例如。卧室或厨房)。如果代理在目标房间内停止,则导航过程被视为成功。房间目标导航需要房间注释。房间的概念是一个高级的语义概念。因此,旋转目标导航代理需要根据视觉细节来了解场景,如家具类型和房间布局。

**多对象导航 (MultiON)** 近年来,越来越多的研究人员开始关注长期导航,即代理记忆所有访问的场景。受此启发,Wani等人。[45]提出了一个用于多对象导航的基准测试——MultiON。在MultiON中,一个代理被要求逐个导航到多个目标对象,这使得导航轨迹会很长。当到达指示目标时,代理引发找到的操作。在部分观察下进行感知和有效的规划将是解决这一任务的关键。

### 3.2 跨模式导航任务

**视觉和语言导航 (VLN)** VLN是一项代理按照自然语言指令[12]逐步导航的任务。以前的任务,如对象目标和房间目标,将对象和房间语义硬编码为一个热向量。相反,VLN引入了自然语言句子来指导导航过程,比如“上楼,穿过前面的拱门走过钢琴。”当走廊到达图片和桌子时向右转。在驼鹿的鹿角旁等待



模拟器	年	使用数据集(s)	分辨率	物理学	FPS	可定制的交互式多代理程序		
米诺斯[11]	2017	SUNCG, Matterport3D	84 84×	3	100	7	7	7
AI2-THOR* [40]	2017	-	×300 300	3	120	3	3	3
House3D [10]	2018	桑克	120 90×	7	600	3	7	7
小木屋[35]	2018	小木屋	×800 600	7	10	7	7	7
Matterport3D [12]	2018	匹配端口3D	×512 512	7	1,000	7	7	7
吉布森[14]	2018	吉布森, 马特波特3D, 2D-3D-S	×512 512	3	400	7	7	7
iGibson [41]	2018	吉布森鸡尾酒	×512 512	3	400	3	3	3
栖息地[13]	2019	马特波特3D, 吉布森, 副本	×512 512	3	10,000	7	7	7
RoboTHOR* [39]	2020	-	×300 300	3	1200	3	3	7

表2: 现有的具体化模拟器的比较 (\*: 模拟器使用的数据集目前尚未发布)。

墙”。如果代理按照指令停止到接近预期目标, 则VLN任务将成功完成。

有一些数据集已经被提出用于视觉语言导航: R2R [12]、R4R [46]和RxR [47]。在[12]中提出了房间到房间(R2R)数据集来研究视觉语言导航。R2R数据集包含21,567条导航指令, 平均长度为29个单词。然而, R2R数据集有几个缺点: 1) 引用路径直接指向目标, 因此R2R指令缺乏描述复杂路径的能力; 2) 指令由几个句子组成, 没有细粒度; 3) 训练数据小, 模型容易过拟合; 4) 教学语言仅使用英语, 不包括其他语言。为了解决这些问题, 人们提出了更多先进的数据集。Jain等。[46]交叉连接R2R中的轨迹和指令, 并生成一个名为R4R的新数据集。FGR2R [48]用子指令及其相应的轨迹丰富R2R。RxR [47]是一个时间对齐的数据集, 它缓解了已知的轨迹偏差并引出了对R2R中可见实体的更多引用。

对话历史(NDH)导航在不熟悉的环境中导航时, 人类通常会寻求帮助, 并根据其他人的反应继续导航。然而, 建立一个能够自主地提出自然语言问题并对答案做出反应的代理仍然是机器人导航的一个长期目标。在NDH [49]中, 需要一个代理根据对话历史进行导航, 该对话历史由几个问答对组成。研究NDH是构建真实世界对话导航机器人的基础。

**具体化问题和回答(EQA)视觉问题回答(VQA) [50]是一种跨模态任务, 其中系统用给定的图像回答一个基于文本的问题。**VQA很快成为最受欢迎的计算机视觉任务之一, 因为它揭示了人类和人工智能代理在自然语言[51], [52], [53]之间互动的可能性。与VQA相比, 一个更高级的活动是在一个看不见的环境中通过自我探索来回答问题。嵌入式询问和回答(EQA) [54]是一个任务, 在3D环境中的一个随机位置生成一个代理, 并问一个问题。EQA是一项具有挑战性的任务, 因为它需要广泛的人工智能技能: 视觉感知、语言理解、目标驱动导航、常识推理等。除了在其他任务中的导航精度外, EQA还提出了EQA精度来衡量代理是否正确地回答了问题。

幻想最近, Qi等人。[55]提出了真实室内环境中的远程体现视觉参考表达, 简称为遐想, 以研究自然语言指令与视觉语义的关联。不同于VLN, 它给出了一个指令来描述一步一步的轨迹

目标, 沉思中的自然语言指令是指一个远程的目标对象。与目标目标导航相比, 幻想提供了丰富的语言描述, 使代理能够在房子里找到一个独特的目标。

**视听导航, 由Chen等人提出。**[56]为实体导航环境引入了音频模态。此任务要求代理通过查看和听到来导航到声音对象。它鼓励研究人员研究音频在导航中的作用。这项工作还为视听导航任务提供了声音空间[56]数据集。声音空间数据集是建立在两个模拟器, 复制品和矩阵端口3D。它包含了102种自然的声音, 种类广泛: 铃声, 开门, 音乐, 人说话, 电话等。

#### 多目标体现的问卷和回答

EQA) EQA中的自然语言问题很简单, 因为每个问题都描述一个对象, 缺乏多个目标之间的属性和关系。在MT-EQA [57]中, 说明就像“卧室里的梳妆台比厨房里的烤箱大”, 即梳妆台和烤箱位于不同的地方, 具有不同的属性。因此, 代理必须导航到多个地方, 找到所有目标, 分析它们之间的关系, 并回答问题。

### 3.3交互式导航任务

交互式询问和回答(IQA)构建一个能够与动态环境进行交互的代理是人工智能社区的一个长期目标。最近提出的交互式模拟器[39], [40], [41]提供了基本的功能, 如打开门或移动椅子, 这使研究人员能够构建一个交互式导航代理。交互式问答(IQA) [58]要求代理通过与环境中的对象交互来回答问题。IQA包含76,800个训练问题, 包括存在问题、计数问题、空间关系问题。

“救命的, 安娜!” (HANNA) HANNA [59]是一个对象查找任务, 它允许代理在它丢失时向自动自然导航助手(ANNA)请求帮助。与NDH提供作为指令的全局对话框历史记录记录的NDH不同, HANNA提供了一个环境, 其中指令会根据情况动态变化。环境创建一个接口, 使人员能够在测试时间中丢失时帮助它。

#### .43评价指标

已经提出了许多评估指标来评估导航代理的性能。我们将它们分为两类: 轨迹不敏感度和轨迹敏感度量。**轨迹不敏感度量, Zhu等人。**[15]使用“平均值”

公制名称	“#配方PS SP UO SI OS CC
路径长度 (PL)	$-\sum_{1 \leq i <  P } d(p_i, p_{i+1})$
导航错误 (网元)	$\# \sum_{i=1}^{ P } d(p_i, r_i)$
Oracle导航错误 (ONE)	$\# \min_{p \in P} \sum_{i=1}^{ R } d(p, r_i)$
成功率 (SR)	$* [NE(P, R) \leq d_{th}]$
Oracle成功率 (OSR)	$* [ONE(P, R) \leq d_{th}]$
PL (SPL)	$SR(P, R) \cdot \frac{\sum_{i=1}^{ P } d(p_i, r_i)}{\sum_{i=1}^{ R } d(r_i, p_i)}$
按编辑距离 (SED) 路径覆盖范围 (PC)	$R) 1 - \frac{ED(P, R)}{\max\{ P ,  R \} - 1} \cdot \frac{1}{ R }$
加权成功	$* \sum_{r \in R} \exp\left(-\frac{d(r, P)}{d_{th}}\right)$
长度分数 (LS)	$-\frac{1}{1 +  P  - PC(P, R) \cdot LS(P, R)}$
通过LS (CLS) 标准化动态时间扭曲 (nDTW)	$0 \cdot \sum_{i=1}^{ P } \sum_{j=1}^{ R } \exp\left(-\frac{d(r_{ik}, p_{jk})}{ R  \cdot d_{th}}\right)$
加权的覆盖范围	$\sum_{i=1}^{ P } \sum_{j=1}^{ R } \exp\left(-\frac{d(r_{ik}, p_{jk})}{ R  \cdot d_{th}}\right)$

表3: 我们从几个方面比较了现有的指标, 包括性能 (“越高越好, t越低”)、公式、路径相似性 (PS)、软惩罚 (SP)、唯一最优 (UO)、尺度不变性 (SI)、顺序敏感度 (OS) 和计算复杂度 (CC)。假设我们有一个预测的轨迹P和一个地面真实的轨迹R。p<sub>i</sub>和r<sub>i</sub>是轨迹P和R上的第i个节点。|P|和|R|分别代表P和R的长度。房子的Dijkstra距离经过了预处理, 任何d(p<sub>i</sub>, r<sub>j</sub>)是0 (1)。

从一个随机的起点到达一个目标所需的步数 (即平均轨迹长度)。然而, 当导航环境变得更加复杂, 导航任务变得更加具有挑战性时, 有很大一部分轨迹失效。后来的工作[10], [11], [60]提出了成功率 (SR) 来测量代理成功达到目标的频率, 其他工作[54], [60]报告导航错误 (NE), 即代理最终停止时到达目标的平均距离。Oracle成功率 (OSR) 建议评估代理是否按照Oracle停止规则[12], [61]正确地停止。这些指标可以衡量代理是否完成任务的概率, 但是, 它们并不能衡量它完成任务的比例。

#### 以路径长度加权的轨迹敏感指标的成功度

(SPL) 是第一个同时评估导航代理的效率和有效性的度量, 它被认为是VLN中的主要度量。SPL会忽略转向动作和代理标题。通过编辑距离 (SED) 加权成功[62]考虑行动并解决这个问题。SED是为在基于图的环境中的指令遵从性而设计的, 在该环境中存在一定的正确路径。然而, 在一些任务中, 如R4R [46]和R6R [63], 被指示的路径并不是直接到目标的。因此, 不适合评价SPL的导航性能。因此, 我们提出了用长度评分 (CLS) [46]加权的覆盖范围来衡量代理的行为对所描述路径的保真度。CLS是两个变量的乘积: 路径覆盖范围和长度分数。Ilharco等。吸收了动态时间扭曲[64]的思想, 这是一种广泛应用于[65]、[66]、[67]等各个领域的方法, 并提出了标准化的动态时间扭曲 (nDTW) 度量[68]来评估导航性能。与CLS类似, nDTW计算预测路径与地面真实路径之间的距离。此外, nDTW对该顺序也很敏感

而CLS是顺序不变的。nDTW可以在一种有效的动态规划算法中实现。路径敏感指标, 如CLS和nDTW, 当它们被用作奖励功能时, 在强化学习中比目标导向的奖励功能来导航[46], [68]。每个度量根据其公式都有其独特的特点。我们比较了Tab中现有指标的公式和特征。3. 在这部分中, 我们将介绍度量来评估一个度量的函数:

- 1) 路径相似性 (PS) 表征了P和R之间的相似性概念。这意味着度量值应该依赖于P中的所有节点和R中的所有节点。PS惩罚偏离地面真实路径, 即使它们导致相同的目标。这不仅是谨慎的, 因为如果不强制执行, 代理可能会在不希望的地形上徘徊, 而且还明确地衡量了关于所提供的语言指令的预测的保真度。
- 2) 软惩罚 (SP) 根据一个取决于图中距离的差的软概念来惩罚与地面真实路径的差异。这确保了较大的差异比较小的差异受到更严重的惩罚, 并且SP不应该仅仅依赖于交集的二分法视图。当且仅当参考路径和预测路径精确匹配时, 3) 唯一最优 (UO) 才会产生一个完美的分数。这确保了完美的分数是明确的: 因此, 参考路径R被视为一个黄金标准。任何其他路径都不应该与引用路径本身具有相同或更高的分数。
- 4) 尺度不变性 (SI) 度量的是, 如果一个度量在不同的数据集上是独立的。如果一个在数据集上的度量变量, 如导航错误, 它的分数不能直接在不同的数据集上进行比较。
- 5) 顺序敏感 (OS) 表示一个度量是否对具有相同轨迹长度、成功率等的导航顺序敏感。

导航顺序显示了一些类型的导航策略，尽管它通常很难被评估。

6) 计算复杂度 (CC) 衡量了计算一对 (P, R) 的成本。设计一种快速的算法来计算自动验证和测试的分数是很重要的。

### 5.3 总结

嵌入式导航基准测试为不同的设置定义了任务和度量值。面向目标的任务，如目标目标、目标和目标目标导航，可以通过3D资产提供标签，不需要额外的人工注释。跨模态导航任务，如IR2R [12]、视觉对话导航[49]或幻想[55]，需要人类标记轨迹和相应的语言描述。交互式交互任务[41]、[59]要求代理学习操作对象，由于其在现实世界场景中的广泛应用而引起了越来越多的关注。

#### 在模拟环境中采用的4种方法

在本节中，我们主要讨论模拟环境中的两个问题：目标驱动导航和跨模态导航。并介绍了解决这些问题的方法。

#### 4.1 目标驱动导航

解决这个问题的方法集中于从一个随机的起始位置导航到一个目标。目标可以由RGB图像、向量或单词来指定。代理预测动作，如左转，右转，向前移动以在实体环境中导航，并预测停止指示停止动作。解决这个问题的方法多种多样，包括：1) 无模型的方法；2) 基于规划的方法；和3) 自我监督的方法。

##### 4.1.1 无模型方法

无模型方法学习了不建模环境的端到端导航，如图所示。4. 学习目标包括模仿学习或强化学习。学习对象的公式是：

$$L = \sum_t a_t \log(p_t) - \sum_t a_t \log(p_t) A_t, \quad (1)$$

\*其中 $a$ 是地面真相行动， $p_t$ 是动作概率，和 $A_t$ 是A3C [69]的优势。虽然广泛的强化学习工作[61]，[70]，[71]长期以来一直在研究二维导航问题，即代理接收每一步的全局状态，但具有部分观察的具体化导航问题仍然具有挑战性。许多机器人控制工作的[72]、[73]、[74]、[75]都关注避障，而不是轨迹规划。

朱等人。[15]首先提出使用深度学习进行特征匹配，使用深度强化学习进行策略预测，这使agent能够更好地推广到不可见的环境中。然后，提出了后续表示 (SR) [76]，以使代理能够与对象交互。这个框架考虑了对象的状态和对场景的离散描述。后续表示对语义信息进行编码，并将其与视觉表示连接起来，如[15]中所示。与[15]只使用强化学习来学习策略预测器的[15]不同，[15]通过模仿学习引导强化学习。以往的模型缺乏编码时间信息的能力。凭借

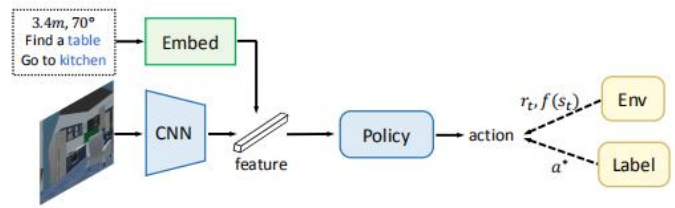


图4：一个无模型的视觉导航模型的说明。该模型从模仿学习和强化学习中学习。 $r_t$ 是奖励和 $f(s_t)$ 表示从状态 $s$ 中计算出的标签 $t$ 。和一个 $a^*$ 是标签代表最佳的行动。

引入了一个LSTM层来编码历史信息，Wu等人。[10]能够构建一个能够泛化到看不见的场景的代理。在消融研究中，本工作证明了A3C [69]在视觉导航任务中优于DDPG [77]，从语义掩模学习的模型优于从RGB输入学习的模型。尽管通过策略上的深度强化学习算法解决了视觉导航问题，但一些工作采用了其他算法。李等人。[78]提出了一种基于q学习的端到端模型，该模型学习视点不变的视觉导航，用于本地移动机器人导航。

有很多工作都是使用物体的分割掩模来增强视觉输入。穆萨维安等人。[16]通过引入在MSCOCO数据集[79]上训练的Faster-RCNN检测器和由[80]定义的分段器来利用视觉输入中的实例特征。沈等人。[81]通过融合不同的视觉表示，包括RGB特征、深度特征、分割特征、检测特征等，改进了导航代理的零镜头泛化。不同的视觉表示是自适应加权的融合。为了进一步提高鲁棒性，他们提出了一种任务间亲和正则化方法，鼓励代理选择更多互补和更少冗余的表示来进行融合。尽管执行良好的检测器和分段器，学习一个健壮的导航策略仍然具有挑战性。例如，为了寻找杯子，人会搜索咖啡机附近的橱柜和水果，人可以先尝试冰箱。为了解决这个问题，Lv等人。[82]将三维知识图和子目标集成到深度强化学习框架中。为了增强跨场景的泛化，Wu等人。[83]在RL目标中引入了一个信息理论的正则化项，并通过学习一个变分生成模型来对动作-观察动力学进行建模。

一些工作调查的是室内导航以外的问题设置，如街景导航或结合其他模式。Khosla等。[84]首先尝试通过具体化的视觉导航方法来解决户外街道导航任务，其中代理纯粹基于全景街景进行导航。DeepNav [85]是建立在卷积神经网络之上的

(CNN) 在大城市使用当地可见的街景图像在大城市导航。这些工作依赖于使用地面真实指南针输入的监督训练，然而，指南针有时在现实世界中无法使用。[86]的另一项工作提出了一个端到端深度强化学习框架，该框架使用谷歌街景的街景作为视觉输入，但没有地面真相罗针。认识到本地特定知识对导航的重要性，他们提出了一种双路径架构，允许封装本地特定的特征。AV-WaN [87]被提出来解决视听导航中的挑战。这个模型学习视听路径点和



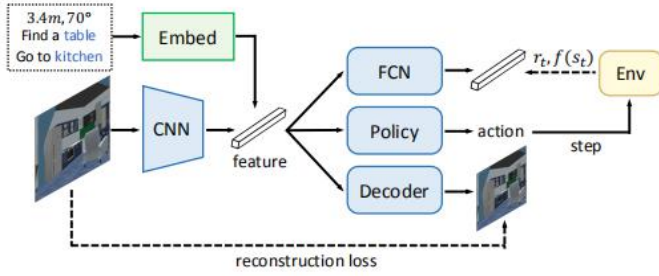


图5: 一个具有自监督目标的端到端视觉导航模型的说明。rt是奖励和 $f(s_t)$ 表示从状态 $s$ 中计算出的标签 $t$ 。

基于其视听观察和部分地图，以端到端的方式动态设置中间目标位置。

#### 4.1.2 自我监督方法

自监督学习是通过各种借口任务利用额外的训练信号的一个长期研究的课题。它使代理能够学习更多的知识，而无需添加额外的人工注释。在深度学习领域，各种自我监督的任务已经被提出，如上下文预测[88]、解决拼图游戏[89]、着色[90]、旋转[91]。在强化学习中，还提出了一些提高数据效率和泛化的辅助任务。谢等人。[17]将自监督学习与基于模型的强化学习相结合来解决机器人的任务。基于传统的UVFA架构[92]，通过特征学习来学习价值函数，贾德伯格等人。[18]发明了辅助控制和奖励预测任务，显著提高了数据效率和鲁棒性。

在具体化导航中，环境包含了难以端到端学习的非结构化语义信息。尽管使用SLAM或记忆机制明确地建模环境，但自监督学习提供了另一种可行的学习非结构化知识的方法。米罗斯基等人。[19]提出了一个具有两个自监督辅助目标的在线导航模型，通过RGB视图预测当前的深度视图，并检测回路闭合。模拟思想[93]已应用于游戏应用[94]进行快速探索。辅助任务也可以加快学习速度。Ye等人。[95]，[96]通过将强化学习与各种辅助任务相结合，在点目标和目标目标导航方面取得了巨大的成功，表述为：

$$L_{total} = L_{RL} + \sum_i^n \beta_i L_{Aux,i}. \quad (2)$$

视觉感知对视觉导航至关重要。但强化学习提供的训练信号包含过多的噪声，无法训练一个鲁棒的特征感知网络。编码解码器结构被证明是有益的[97]视觉编码和分割预测。此外，还采用了一个辅助任务来惩罚分割误差，这有利于特征感知的学习。然而，这种自监督的辅助任务只学习两个相邻状态之间的低级动态函数，而不能学习高级语义信息。为了保证轨迹中动作的语义一致性，Liu等人。[98]提出了一个辅助的正则化任务来惩罚表征的不一致性。这个正则化任务鼓励策略网络从每个传感器中提取显著的特征。现实世界中的机器人运动还很遥远

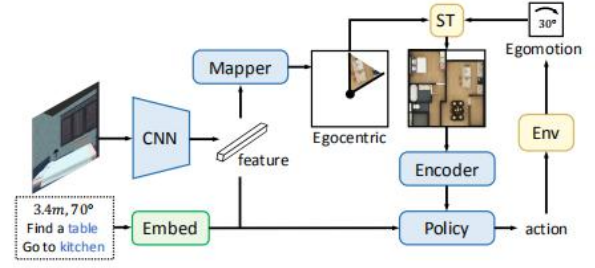


图6: 基于“神经SLAM”的模型的常见实践的概述。“ST”是空间变换函数。

确定性是由于驱动噪声的存在，这可能是由车轮滑动、运动传感器误差、反弹等引起的。为了降低噪音，Datta等人。[99]引入了一种基于时间差分的定位估计的辅助任务。辅助任务用于训练一个CNN网络，并使用估计的运动作为策略网络的输入。一个好奇心驱动的自我监督目标[100]被用来鼓励探索，同时惩罚重复的行为。一个稳定的由好奇心驱动的政策，没有重复的行动，可以提高勘探效率。自我监督的辅助任务也有助于导航的跨模态理解。迪恩等人。[101]使用音频作为自我监督探索的额外模态。它包括一种好奇心驱动的内在奖励，它鼓励主体探索不同的感觉模式（音频和视觉）之间的新联系。自监督导航方法的管道概述如图所示。5. 该代理首先嵌入一个视觉图像和一个指令作为特征。然后将视觉特征和指令特征融合来预测动作。辅助任务使用融合的特征来进行预测，如预测奖励，或重建输入的视觉图像。

#### 4.1.3 基于规划的方法

在解决定位问题的同时，未知环境下的映射构建问题被称为同步定位和映射（SLAM）[102]，[103]。早期对视觉导航的研究是使用立体相机[104]、[105]和单眼相机，如MonoSLAM [106]进行的。在过去的十年里，传统的基于几何的方法[107]，[108]，[109]仍然主导着这个领域。随着深度学习的发展，人们提出了CNNSLAM [110]、DVO [111]和D3VO [112]等方法。提出了一些室内任务来研究SLAM，如KITTI [113]和EuRoC [114]。然而，这些任务不同于具体化导航任务。测程基准是估计给定的视觉输入序列的位置，而导航任务是将指令与环境语义对齐。

近年来，研究人员发现定位能力对导航非常重要，特别是对于长期的路径规划。因此，一些工作引入了SLAM方法来对房屋进行建模，并提高了代理的定位能力。神经地图[115]将这一想法推广到所有的深度强化学习代理中，而不仅仅是导航。然而，这项工作假设代理的位置始终是已知的，并且没有利用这个内存的二维结构。神经SLAM [116]通过将类似于SLAM的程序嵌入到软注意[117]中来解决这个问题。为了避免与反复扭曲相关的空间模糊，MapNet [118]建议使用以世界为中心，而不是以自我为中心的地图。与以前的工作不同，MapNet维护了一个2.5D的表示

通过一个深度神经网络模块，学习从三维具体化的视觉输入中提取视觉表征。戈登等人。[58]提出了层次交互式记忆网络（HIMN），一个基于IQA任务的层次控制器框架。高级控制器是决定长期导航目标的规划器，低级控制器预测动作，与环境交互，并回答问题。Gupta等人。[20]介绍了具体化导航中的神经-slam方法。这项工作包括两个部分：测绘和规划。映射机制维护一个二维内存映射。对于每一步，它都将所体现的场景转换为一个二维特征，并使用该特征更新地图。规划机制使用一个值函数来输出一个策略。

有效的探索被广泛认为是强化学习（RL）[119]中的主要挑战之一，[120]，[121]。同样，这在导航中也很重要，因为目标并不总是从起始位置可见，而且代理需要探索看不见的场景并搜索目标。最近，基于显式建模的语义记忆的探索被证明是有效的。为了学习具有空间记忆的策略，Chen等人。[122]通过模仿学习来引导模型，并使用纯粹来自车载传感器的覆盖奖励来微调它。主动神经SLAM（ANS）[22]是一种成功的神经SLAM方法，实现了CVPR 2019人居点目标导航挑战。ANS提出了一种规划的层次结构。受层次RL [123]、[124]思想的启发，ANS通过强化学习学习高级规划器，通过模仿学习学习低级规划。该映射器是通过一个预测二维映射的辅助任务来实现的。

地图的第一个通道表示是否有障碍，而第二个地图表示该位置是否已经被探索过。然而，预定义的二维地图不能帮助长期导航，因为在不同视点下的场景的语义信息没有被编码在地图中。神经拓扑SLAM [125]提出了一种更先进的方法来存储所观察到的特征表示。该方法引入了一个图更新模块来利用语义。图形更新模块维护一个拓扑特征内存。对于每个步骤，该模块都会将当前的观察结果本地化到内存节点中。如果观察结果没有局限位于内存的任何节点，图更新模块将在拓扑特征内存中添加一个新节点。面向目标的语义探索（Sem Exp）[126]处理现实环境中的目标目标导航任务。该方法首先构建一个情景语义映射，并利用其基于目标对象的类别来探索环境。这种方法实现了最先进的栖息地目标导航挑战2020。“神经SLAM”的常见实践的概述-

基于模型的模型如图所示。6.除了视觉编码器和指令编码器，如图所示。5、基于“神经松弛”的模型有一个独特的模块来投射一个具体化的视觉视图来进行特征表示，并将其存储在一个2D自上而下的地图中：

$$m_t, x_{t-1:t}^* = f_{\text{猛击}}(s_t, x_{t-1:t}^*, m_{t-1} | \theta_S) \quad (3)$$

其中 $x_{t-1:t}^*$ 代表之前的姿势， $m_{t-1}$ 是以前的地图，和 $\theta_S$ 代表参数。这张地图模拟了房间结构和场景的视觉表示。将投影特征表示与视觉特征和指令特征融合，共同预测一个动作。

#### 4.1.4总结

与传统的机器人技术方法相比，无模型方法能够通过对实体模拟器的大规模数据进行采样来获得鲁棒的导航模型。一些工作采用了检测和分割的方法来获得更好的视觉视图。尽管有室内场景，无模型方法在街景和多模态环境中取得了很大的成功。提出了自监督方法，利用辅助任务的额外知识来提高学习效率和泛化能力。基于规划的方法利用二维映射或拓扑内存在导航过程中建模环境。

### 4.2跨模式导航

一个懂自然语言的导航机器人可以完成更复杂的任务，比如“在厨房里拿起杯子”或“帮我找到楼上的杯子”。在本节中，我们将介绍三种解决跨模式导航任务的工作：1) 一步一步的方法；2) 基于预训练的方法；3) 基于规划的方法。

#### 4.2.1序列到序列的导航

安德森等人。[12]首先提出了一个类似于[127]的序列到序列模型来解决视觉语言导航问题。该模型依次对语言指令逐字进行编码，将句子特征与视觉图像特征连接起来，并对动作序列进行解码。然而，序列到序列模型缺乏稳定性和泛化性，因为它没有考虑到现实环境中的动态。RPA [23]被提出通过装备一个“前瞻性”模块来解决泛化问题，该模块可以学习预测未来的状态和奖励。为了提高指令-轨迹对齐的泛化能力，Fried等人。[24]提出了一种名为“说话者-追随者”的数据增强方法来改进模型的泛化性。为了生成增强数据，说话者首先将一个随机的轨迹转换为一个指令，跟随者其次将一个指令转换为一个轨迹：

$$\arg\max_{r \in R(d)} P_S(d|r) \cdot P_F(r|d) \quad (4)$$

其中PS扬声器是P吗F为跟随者，d表示指令，r表示轨迹，V表示加权因子。本文[24]的另一个贡献是定义了一个高级动作，它朝着全景空间中的方向向前移动，而不是像向左转、向右转和向前走这样的低级动作。与低级动作的定义相比，这种方法大大减少了描述相同轨迹的动作序列的长度。在高级动作空间中导航需要更少的预测时间，这使得模型更容易训练，测试也更鲁棒。然而，以往的方法只通过教学轨迹数据对学习模仿学习导航，它监督最短路径，而忽略次优轨迹，从而使leds过拟合。为了解决这个问题，Wang等人。[128]提出通过模仿学习和强化学习来联合学习一个导航代理。此外，该方法还引入了一种LSTM来对视觉特征的时间信息进行编码，并引入了一种跨模态机制，以实现更好的视觉语言导航能力。马等人。[129]提出了一种具有视觉-文本共接地模块和进度监控器的自我监控代理。进度监测器采用了来自共接地的跨模态特征



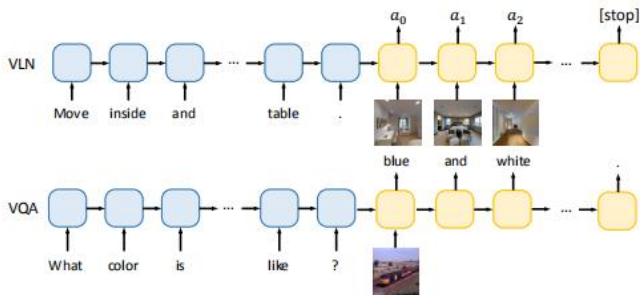


图7: VLN和VQA中序列序列模型的比较。

模块和估计已完成的进度。由于视觉语言任务中的指导指导代理一步一步地到达目标，因此进度信息包含了丰富的知识，有助于提高对代理的感知。马等人。在[130]中提出遗憾代理，其中的遗憾模块使用估计的进度来指示代理是否导航到错误的地方并需要返回。类似于遗憾的代理，Ke等人。提出了[131]一个使用异步搜索来通过启用显式回溯来提升VLN导航器的框架。安德森等人。[132]将逐步导航过程视为一个视觉跟踪任务。该方法在贝叶斯状态跟踪[3]的框架内实现了导航代理，并建立了一个具有可学习观测和运动模型的端端可微直方图滤波器[133]。缓解视觉过拟合的一种常用方法是在视觉特征上应用辍学[134]层，该层是由预先训练好的网络如VGG[135]或ResNet[136].t.)提取的。Tan等人。[137]认为，简单地在视觉特征上应用一个辍学层就会导致不一致，e.g. 这个框架里的一把椅子可以掉在下一个框架里。为了解决这个问题，他们提出了一个环境辍学层，在轨迹中随机退出一些固定的通道。朱等人。[138]提出了一个AuxRN框架，该框架引入了自监督的辅助任务，从几个方面利用环境知识。除了引入了在其他具身视觉导航方法中广泛使用的时间差辅助任务外，AuxRN还引入了轨迹复述任务和指令-轨迹匹配任务来学习轨迹的时间语义。Fu等人没有生成低质量的增强数据。[139]引入了反事实思维的概念，对具有挑战性的路径进行采样，以增强训练数据集。他们提出了一个模型不可知论的对抗性路径采样器（APS）来选择困难的轨迹，并且只考虑有用的反事实条件。

与早期基于数据增强和其他经典导航方法的作品不同，一些作品发现了自然语言对VLN的重要性。托马森等人。[140]发现单峰基线优于随机基线，甚至优于一些多模态基线。因此，该工作主张，废除单峰模式来评估偏差对于提出一个数据集是重要的。黄等人的研究。[141]表明，在[24]中，只有有限数量的这些增强路径是有用的，并且在使用了60%的增强数据后，改进会随着额外的增强数据的增加而减少。为了避免在奖励工程方面的广泛工作，Wang等人。[142]提出了一种软专家奖励学习模型，其中包括两部分：1) 软专家蒸馏，鼓励代理以软方式表现为专家；2) 自我感知，尽可能快地将代理推向最终目的地。夏等人。[143]利用多个指令作为对同一轨迹的不同描述



图8: 一个基于预训练的框架的示例。

解决语言歧义，提高泛化能力。这项工作表明，VLN中的人类注释在一定程度上根据具体的场景和轨迹有很大的偏差。视觉特征的质量是提高具身导航性能的关键。之前的工作通过预先训练的ResNet101[136]等CNN网络从全景中提取全局视觉特征。Hong等。[144]引入Faster-RCNN来检测导航中的对象，并在视觉和语言实体之间的关系图，用于视觉-语言对齐。尽管有视觉输入，但结构信息也有助于导航。胡等人。[145]发现，语言指令包含高级的语义信息，而视觉表示则是一种低级的模态，这使得视觉-语言对齐变得困难。基于此，他们将接地过程分解为一组具有不同模式的专家模型，并在预测时对它们进行集成。为了更好地研究语言理解在VLN任务中扮演什么角色，Hong等人。[48]认为，中间监督在视觉-语言对齐中很重要。因此，他们提出了FGR2R，这是一种使导航过程能够被可跟踪的方法，并鼓励代理在子指令的级别上移动。

## 2.24 基于预训练的方法

在视觉语言导航的研究中发现了几个挑战：1) 训练效率低；2) 大数据偏差（包括视觉和语言）；3) 从看到的场景到看不见的场景缺乏泛化。为了解决这些挑战，我们提出了基于训练前的模型，以从来自其他来源的大规模数据集中学习，并快速适应看不见的场景。

训练效率低，传统的编解码器框架首先通过教师强迫或学生强迫对总轨迹进行采样，然后对梯度进行反向传播。在其他深度学习任务中，如图像分类[1]或文本识别[146]，该模型直接预测结果。然而，在视觉语言导航任务中，agent通过与环境的逐步交互来预测轨迹，这非常费时，降低了训练效率。

大数据偏差视觉语言导航场景是这样的

R2R中的61个房子不能覆盖所有的房子。从

在自然语言方面，在R2R任务中，只有69%的双字句在训练和评估之间共享。

由于缺乏泛化，尽管提出了轨迹训练、视觉特征训练和自然语言训练等增强方法，但缺乏不同的训练数据仍然在很大程度上限制了泛化。因此，从其他任务和数据集中引入额外的知识成为一个很有前途的话题。

基于预训练的方法通过大规模数据[51]学习，很大程度上提高了模型的泛化能力，

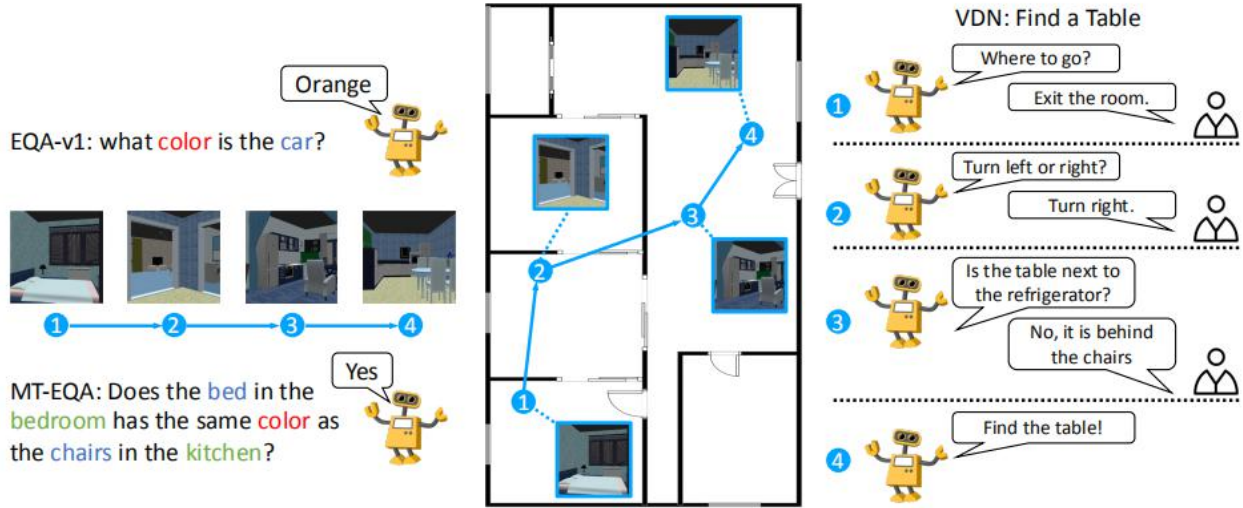


图9: 比较了三个具体化的视觉语言导航任务: EQA-v1 [54], MT-EQA [57]和VDN [49]。

[136].此外,基于bert的方法[147], [148]对变压器网络进行预训练,在视觉、语言和跨模态任务方面取得了巨大的成功。许多研究人员考虑通过基于保留的方法来解决视觉语言导航问题。李等人。[149]建议出版社首先引入一个预先训练好的语言模型来学习指令表示。他们提出了一种随机抽样方案,以减少训练中的专家行动和测试中的采样行动之间的差距。马琼达尔等人。[150]提倡通过利用大规模的网络数据来改进模型。但是,很难将静态图像数据传输到VLN任务中。因此,他们提出了VLN-bert,一个基于转换器的模型,它由静态图像及其字幕进行预训练。流行的[151]是对一个具体化环境中进行大量的预训练,有两个预测目标,掩蔽语言建模(MLM)和动作预测(AP):

$$\begin{aligned} L_{MLM} &= -\mathbb{E}_{s \sim p(\tau), (\tau, x) \sim D_E} \log p(x_i | \mathbf{x}, s), \\ L_{AP} &= -\mathbb{E}_{(a, s) \sim p(\tau), (\tau, x) \sim D_E} \log p(\mathbf{x} | x_{[CLS]}, s), \end{aligned} \quad (5)$$

其中(s, a)一个状态-动作对。流行被证明对一些视觉语言导航数据集有效,包括R2R, CVDN和HANNA。具体化导航代理接收部分观测而不是全局观测,最好建模为部分可观测的马尔可夫决策过程。与编解码器模型不同,以往基于预训练的模型在导航过程中不记忆之前看到的场景,并利用时间知识,导致动作预测中的信息丢失。基于此, Hong等人。[152]提出了一种可用于VLN的含时特性的循环多层变压器网络。该方法引入了一个变压器,它保持一个特征向量来表示时间上下文。

#### 带有提问和回答的4.2.3导航

Das等人并没有被动地感知来自人类指挥官的自然语言指令。[57]建议一个智能代理应该能够通过导航来回答一个问题。因此, Das等人。提出一个名为EQA(具体化问卷和回答)的新任务,在三维环境中随机生成一个代理,并被要求回答一个问题。为了回答,代理必须首先导航来探索环境,通过以自我为中心的视觉收集信息,然后回答

问题。为了解决这一具有挑战性的任务, Das等人。提出了一个带有自适应计算时间(ACT)模块[153]的CNN-RNN模型PACMAN,以决定重复执行一个动作[57]的次数。PACMAN通过最短路径演示进行引导,然后使用RL进行微调。然而,这种方法缺乏高水平表示的能力。在后来的工作中, [154], Das等人。提出了一种名为神经模块化控制器(NMC)的分层策略,该策略在多个时间尺度上运行,其中高级主策略提出了由低级子策略执行的子目标。Anand等。[155]在EQA上发现了一个蒙眼(只有问题)的基线,并发现该基线执行了之前最先进的模型。他们认为,以前的EQA模型在利用来自环境中的上下文方面是无效的,而且EQAv1数据集有很多噪声。

吴等人。[156]提出了一个简单的监督学习基线,它可以与最先进的EQA方法竞争。为了提高EQA在未知环境下的性能,在本文中,他们提出了一个允许代理回答适应问题的设置。余等人。[57]认为, EQA任务假设每个问题都有一个目标,这限制了它的应用。因此, Yu等人。提出了多目标EQA(MT-EQA),一个EQA的广义版本。这个任务的问题包含多个目标。它要求代理对多个目标进行比较推理,而不是简单地感知一个目标的属性。维曼斯等人。[157]将EQA问题扩展到光现实环境。在这种环境中,他们发现点云表示对导航更有效。罗等人。[158]认为,视觉感知能力限制了EQA的表现。他们引入了Flownet2 [159],一个高速视频分割框架,作为辅助导航和问答的骨干。李等人。[160]提出了一个MIND模块,它对环境图像进行建模,并生成被视为短期子目标的心理图像。Tan等人。[161]调查多个目标之间的问答问题。在这个任务中,代理必须导航到多个地方,找到所有目标,分析它们之间的关系,并回答问题。受视觉问题回答(VQA) [50]和视频问题回答(VideoQA) [162]最近的进展的推动, Cangea等人。[163]提出了VideoNavQA,这是一个包含成对的问题和视频的数据集。这是在房子的3D环境中生成的。这个数据集填补了这个空白



在VQA和EQA之间。VideoNavQA任务代表了EQA范式的另一种观点：通过向代理提供接近最优的轨迹，导航问题比推理问题更容易解决。邓等人。[164]提出了操作问题回答（MQA），其中要求机器人通过操纵主动探索环境来找到问题的答案。为了提出解决MQA的一个有前途的方向，他们提供了一个由QA模块（VQA框架）和操作模型（Q学习框架）组成的框架。尼尔森等人。[165]构建一个代理，它在三维环境中进行探索，并偶尔在导航过程中请求注释。同样，Roman等人。[166]提出了一种合作视觉和对话导航的双代理范式。他们的模型学习多种技能，包括导航、提问和问答组件。

## 2.44 导航与对话

人类使用对话框来指导机器人已经有很长的历史了。在实体导航领域，班纳吉等人。[59]建议“帮助你，安娜！”（HANNA），一个交互式的逼真模拟器，其中一个代理通过请求和解释自然语言和视觉辅助来完成寻找对象的任务。阮等人。[169]提出了一个名为VLNA的任务，其中一个代理通过语言被引导来寻找对象。然而，这两个任务中的语言教学远离了现实世界的问题：hanna的反应是由一个训练过的模型自动生成的，而VLNA的指导是以模板语言的形式，编码黄金标准的计划行动。Vries等。[170]提出“谈论散步”（TtW），即两个人进行交流，在户外环境中达到目标位置。然而，在TtW中，人类使用了一个抽象的语义映射，而不是一个以自我为中心的环境视图。

托马森等人。[49]提出了视觉和对话导航（VDN），这是一个以导航为中心的问答和问答任务的支架，在这里代理按照多轮对话历史而不是指令进行导航。与R2R数据集中的单轮指令相比，VDN提供了多轮注释，其中每轮对话都描述了一个子轨迹。更细粒度的对话注释有助于研究自然语言的导航问题。朱等人。[138]提出了一个具有跨模态记忆机制的框架来捕捉对话回合和子轨迹之间的层次相关性。更普遍的是，一些方法，如流行的[151]和婴儿行走[63]，使用句子指令和对话指令来验证它们的导航能力。不幸的是，这些工作严重依赖于劳动密集型的对话注释。为了缓解这种情况，罗曼等人。[166]利用基于视觉视图生成对话问题的答案。这项工作解决了基于回合制的对话建模中面临的四个挑战，其中包括：1) 决定何时问一个问题；2) 生成导航员的问题；3) 生成问题

答案对作为指导；4) 生成导航器操作。为了实现这一目标，罗曼等人。[166]引入了一个双代理范式，其中一个代理导航和询问问题，而另一个引导代理回答问题。与之前用模板语言指导导航器的工作不同，这项工作通过对CVDN对话的预训练来初始化oracle模型来生成自然语言。

对话框并不总是描述一步一步的导航过程。相反，神谕描述了目标场景，并让导航员找到它，这通常发生在有人在新建筑中迷路的时候。哈恩等人。提出一个LED任务

（从对话历史记录中定位观察者）以实现它何时丢失。受此启发，他们提供了一个名为“你在哪里”[171]的数据集，该数据集由两个人类的6k个对话组成。由于多智能体通信系统[172]、[173]在现实世界中的广泛应用，研究人员对在物理环境中实现对话导航产生了兴趣。Marge等。[174]提出了一个MRDwH平台，在一个大型户外模拟环境中实现了两个模拟机器人的自主对话管理和导航。Banerjee等。[175]提出了RobotSlang基准，这是一个数据集，是通过配对控制物理机器人的人类“驾驶员”并向人类“指挥官”提出问题来收集的

我们比较了图中具体化问题回答（EQA）[54]、多目标体现问题回答（MTEQA）[57]和视觉对话导航（VDN）[49]的差异。9. 我们举例说明了相同的导航轨迹的三个不同的对话。与EQA相比，MTEQA中的问题更为复杂，因为它应该描述多个目标。代理必须获得高级技能，如推理、比较和多对象本地化技能，才能完成MT-EQA。在EQA和MT-EQA任务中，代理需要通过导航来回答来自人类的问题。然而，在VDN任务中，代理是导航员和提问者，它要求人类提供提示来找到目标。任务设置的不同导致了导航模型的设计的不同。

## 2.54 总结

自然语言为人类提供了一个与机器人交互的界面。一个具有跨模态理解的机器人能够完成复杂的任务，比如按照自然语言指令或对话进行导航，向神谕询问更多细节，等等。人们从视觉语言导航问题的各个方面提出了大量的研究工作。

## 在现实世界的环境中的5种方法

模拟环境中的嵌入式导航方法为解决现实世界中的导航问题提供了一个很有前途的方向。在本节中，我们将介绍1) 介绍用于实际应用的方法；2) 将它们与模拟器中的方法进行比较；3) 讨论模拟到真实传输的可能性。

## 5.1 个真实世界的导航方法

### 5.1.1 室内机器人导航

深度学习在室内导航的现实世界应用中起着重要的作用。LeCun等人。[176]首先采用卷积网络进行避障。哈德塞尔等人。[177]提出了一种自监督学习过程，该过程通过层次深度模型准确地分类远程视觉语义。该方法在一个应用于地面机器人（LAGR）[25]的学习系统上进行了验证。后来，越来越多的现实世界的机器人采用深度学习来感知和提取独特的视觉特征[178]。张等人。[179]关于真正的机器人在简单的迷宫环境中导航的研究。基于RL算法解决具有挑战性的控制任务的成功，[77]，[180]，Zhang等人。在学习过程中采用后继表示来实现快速适应。莫拉德等人。[26]提出了一种名为NavACL的室内对象驱动导航方法，该方法使用自动课程学习，并且很容易推广到新的环境和目标中。Kahn等人。[181]采用多任务学习和非策略RL



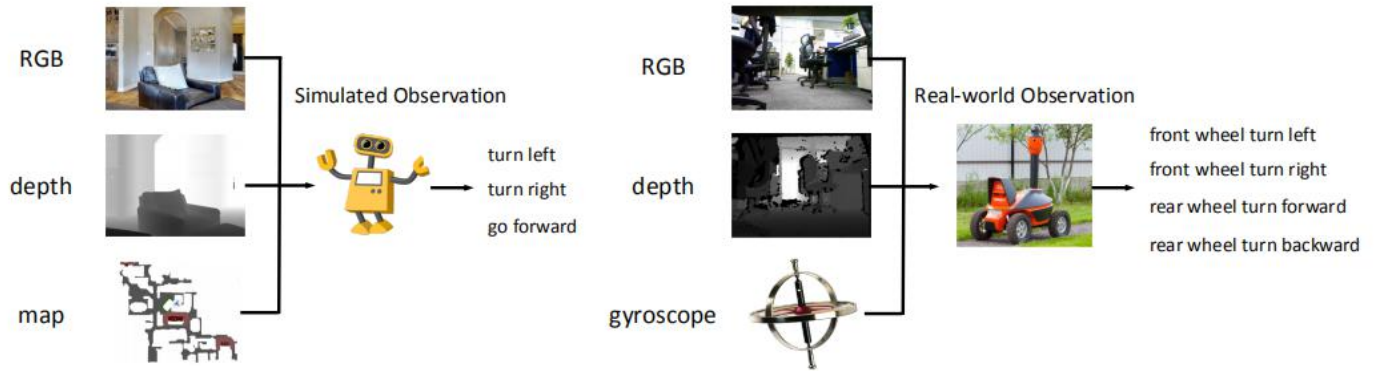


图10: 模拟环境与真实环境的输入空间和动作空间的比较。

学习直接从现实世界的事件中学习。这种方法使机器人能够自动学习, 并可以轻松部署在多个真实世界的任务上, 而无需任何人类提供的标签。

### 5.1.2 户外机器人导航

人类研究户外导航机器人的历史一直很久。索普等人。[27]提出了两种算法, 一种基于rgb的道路跟踪方法和一种基于3d的障碍物检测方法, 用于机器人在校园中学习导航。罗斯等人。[182]将深度学习和强化学习相结合, 进行无人机避障学习。莫拉德等人。评估NavACL在Gibson和栖息境两种模拟环境下的性能。我们将导航转移到涡轮机器人3轮式机器人 (AGV) 和DJI Tello四旋翼机器人 (UAV)。定量和定性的结果都表明, 在模拟环境中训练的NavACL策略在AGV和无人机中是惊人的有效的。曼德森等人。[183]使用条件模仿学习来训练水下航行器在没有任何预先地图的情况下导航到接近稀疏的地理路径点。

### 5.1.3 远程导航

该模型在真实的机器人平台上取得了最佳的性能, 并表现出了竞争性的泛化能力。博伦斯坦等人。[184]建议维护一个世界模型, 不断更新和实时更新, 以避免障碍。该世界模型学习并模拟了真实世界的环境, 并降低了数据采样的成本[185]。刘等人。提出了终身联邦强化学习 (LFRL), 一种用于云机器人系统导航的学习架构来解决这个问题。

远程导航对现实世界中的机器人来说是一个挑战。为了解决这一问题, 弗朗西斯等人。[186]提出了一种分层机器人导航的PRM-RL方法。PRM-RL模型包括一个从噪声传感器中学习短期避障的强化学习代理, 以及一个基于采样的规划器来绘制导航空间。沙h等人。[187]提出了一种基于学习的导航系统ViNG, 以实现视觉上指示的目标, 并在一个真正的移动机器人平台上演示了该系统。与之前的工作不同, ViNG使用纯离线体验, 不需要模拟器或在线数据收集, 这大大提高了训练效率。映射[188]和路径规划[189]也被许多现实世界的应用程序广泛采用。戴维森等人。[190]建立了一个自动系统, 它能够在目标导向的导航过程中检测、存储和跟踪合适的地标特征。它们展示了机器人如何利用主动视觉来提供连续和准确的全球定位, 从而实现高效的导航。

Sim等人。[191]启用“机器人”

通过采用三维点地标的混合地图表示来准确地定位其位置。

## 6个导航从模拟器到现实世界

在本节中, 我们首先通过比较模拟环境和真实世界环境之间的差异, 来演示现实世界导航中所面临的挑战。然后, 我们介绍了专注于解决这些挑战的方法。

### 6.1 模拟导航与实际导航的比较

今天, 目前在模拟导航方面的成就仍然远远不能建立一个现实世界的导航机器人。与模拟环境相比, 真实世界的导航环境更加复杂和不断变化。模拟环境 (生境[13]) 和真实世界环境之间的输入比较如图所示。10.

#### 6.1.1 域差距的原因

总结了模拟实域差距的三个方面: 1) 观测空间2) 作用空间3) 环境动力学。观察差异。对模拟的envi-的观察

环境可以是RGB图像、深度图像或地面真实地图。RGB图像的质量和深度图像输入都较高。环境包含所有静态对象信息, 并使其能够提供地面真实信息, 如房间结构、分割或对象标签。模拟的环境提供了不真实的合成图像, 而真实世界的环境要复杂得多。现实环境中的传感器, 包括RGB、GPS和速度传感器, 通常是有噪声的, 而模拟环境中的传感器没有噪声。虽然一些模拟器[11]、[13]、[14]提供了物理传感器, 并模拟了一些物理相互作用 (如碰撞和加速度), 但它们的物理引擎的性能仍远非真实。

行动差异。与模拟环境中简单的动作空间由“左转”、“右转”和“向前”不同, 现实世界中的动作空间更具挑战性, 这取决于机器人的结构。在现实世界的导航过程中存在许多障碍, 阻碍机器人转弯或向前移动。现实世界的环境通常是动态的, 因为环境是如此复杂, 许多因素在长期或短期内发生变化, 如温度、湿度、摩擦力、障碍物和行人。另一个挑战也是

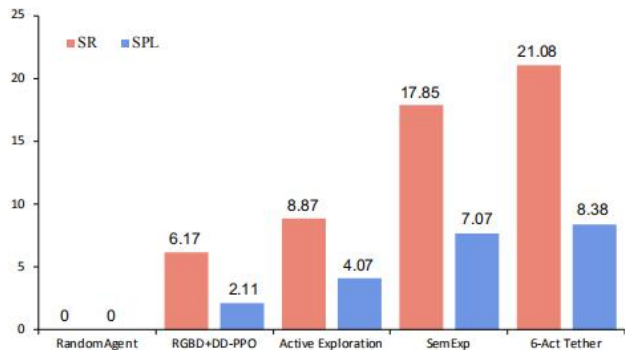


图11: 栖息地目标目标导航方法的性能, 包括DD-PPO [195]、主动勘探[22]、SemExp [126]和6-Act Tether [96]。

在模拟环境中, 被广泛忽视的是动作空间的复杂性和不稳定性。例如, 执行相同动作的结果是不确定的, 因为物理条件是不不断演变的, 如车轮正在打滑或被卡住。**环境动力学**。环境条件的变化, 如温度、湿度或部件磨损导致环境动态。一个没有在线适应能力的政策不能很好地处理这个问题。近年来, 学习动态环境的自适应策略受到越来越多的关注。一些工作人员[192], [193], [194]提出了模拟机器人环境来实现这一点, 然而, 模拟比现实世界要简单得多。

### 6.1.2 域差距的解决方案

领域的差距带来了关键的挑战, 研究人员提出了方法来填补这两种情况之间的差距。移动机器人导航被认为是一个几何问题, 它要求机器人感知环境的几何形状, 以规划无碰撞路径到达目标。避免障碍是最重要的挑战之一, 在之前的工作中已经提出了许多方法[27], [176], [186]来实现这一目标。然而, 模拟任务中的机器人导航被认为是一个策略学习问题, 它在具有多种可能路径的复杂环境中学习从起始位置到目标的鲁棒导航策略。基于slam的方法, 如[60], [116]对映射和路径规划有很大的贡献, 这对模拟导航和真实导航都是通用的。深度学习显示了其处理图像和机器人控制的学习策略的能力, 这是广泛应用于这两种设置。然而, 模拟导航和真实导航在深度学习中的应用是不同的。在现实导航中, 深度神经网络用于感知RGB输入[177], 预测未来[197], 学习导航策略[186]。然而, 由于采样效率低下和现实环境中复杂的动态因素, 该策略不够稳健。一些工作[184], [185]提出对环境进行建模, 而其他工作[186]采用手工制作的规则, 以提高导航策略的鲁棒性。数据采样在模拟环境中更为有效。大多数模拟器渲染RGB和深度图像在超过数百帧每秒(FPS), 其中最快的模拟器, Rabacer[13], 达到10万FPS。快速的数据采样使学习具有较大的批量规模。许多工作证明, 一个大的训练批量大小导致表示[45], [198]的鲁棒性。尽管渲染了RGB和深度

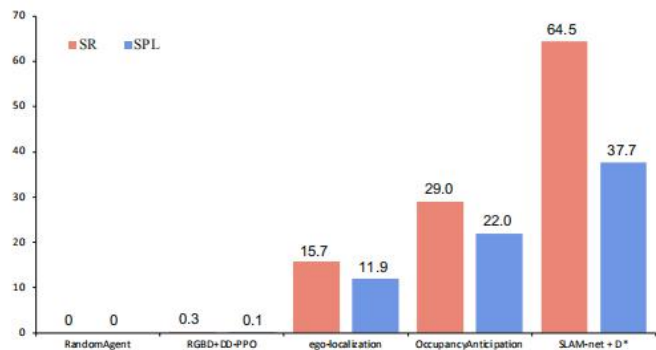


图12: 该方法在栖息地点目标挑战上的表现, 包括DD-PPO [195]、自我定位[99]、占用预测[199]和SLAM-net。

在图像中, 一些模拟环境能够提供语义分割掩码[13]、[14]、[157]。一个更精确的模拟器, 噪音少, 以方便训练。有了更丰富、无噪声的数据, 研究人员就可以在导航代理上应用更深层次的神经网络, 而不用担心过拟合。例如, 变压器[117]由于其特征表示的能力, 在模拟环境的导航工作中得到了广泛的应用, 而在噪声数据上进行训练, 容易过拟合。

### 6.1.3 学习效率

许多研究人员关注学习效率, 因为在现实世界中的数据采样是缓慢而昂贵的。Lobos-Tsunekawa等。[200]提出了一种针对两足类机器人的无地图视觉导航方法。该方法采用DDPG算法[77]从彩色图像中提取信息, 从而推导出运动命令。该方法在一个物理机器人上运行20 ms, 允许在实时应用中使用。布鲁斯等人。[201]提出了一种学习在移动机器人上导航到固定目标的方法。通过使用一次环境遍历和随机环境增强, Bruce等人。演示了在没有微调的真实环境变化下的零镜头转移。为了进一步提高采样效率, 法伊弗等人。[202]利用之前的专家演示进行培训前培训, 从而在微调过程中大大降低培训成本。

### 6.2 导航传输

迁移学习在具身导航中越来越受到关注。研究人员的动机来自两个方面: 1) 学习能够在不同领域和任务中执行准确和高效的导航的导航代理; 2) 在现实世界的导航机器人中部署在模拟环境中训练的agent。

训练一个模型来学习在不同领域的导航技能是具有挑战性的。此外, 由于模拟环境与真实环境之间的领域差距很大, 在模拟环境中训练的良好的导航策略不能轻易地转移到真实环境中。许多导航任务已经被提出来研究在不同场景下的不同导航能力。

在本节中, 我们将讨论从两个不同层次的导航中的迁移学习: 1) 任务级转移; 2) 环境级转移, 包括模拟到真实的转移。任务级别的传输要求代理学习一种适应不同输入模式或目标的策略; 环境级

方法	R2R验证已看到				R2R验证看不到				R2R测试看不见			
	TL	NE#	SR	SPL	TL	NE#	SR	SPL	TL	NE#	SR	SPL
随机的 人	9.58	9.45	16		9.77	9.23	16		.899	9.79	13	12
	—				—				11.8	1.61	86	76
	—		—		—		—		5			
Seq2Seq [12]	11.33	6.01	39	—	8.39	7.81	22	—	8.13	7.85	20	18
扬声器-追随者[24]	—	3.36	66	—	—	6.62	35	—	14.82	6.62	35	28
RPA [23]	8.46	5.56	43	—	7.22	7.65	25	—	9.15	7.53	25	—
SMNA [129]	—	3.22	67	58	—	5.52	45	32	18.04	5.67	48	35
RCM+SIL [128]	10.65	3.53	67	—	11.46	6.09	43	—	11.97	6.12	43	38
遗憾的[130]	—	3.23	69	63	—	5.32	50	41	13.69	5.69	48	40
PRESS* [149]	10.57	4.39	58	55	10.3	5.28	49	45	10.77	5.49	49	45
快速-短[131] EnvDrop	—	—	—	—	6	4.97	56	43	22.08	5.14	54	41
[137] AuxRN [138]流	11.00	3.99	62	59	21.1	5.22	52	48	11.66	5.23	51	47
行*[151]主动探索	—	3.33	70	67	7	5.28	55	50	—	5.15	55	51
[196] RelGraph [144]	10.3	3.67	69	65	10.7	4.71	58	53	10.5	5.30	54	51
VLN BERT* [152]	2	3.20	70	52	—	4.36	58	40	1.21	4.33	60	41
0	19.7	3.47	67	65	0	4.73	57	53	10.2	4.75	55	52
	0	2.90	72	68	—	—	—	—	9	4.09	63	57
	10.1				10.19	3.93	63	57				
	3				20.60				12.3			
	11.1				9.99				5.6			
	3				12.01							

表4：在单次运行设置下，R2R上的代理性能的比较。\*基于训练的方法。

转移要求模型对不同的动力学和转移函数不变。

DisCoRL [28]引入了一种策略蒸馏方法[203]来传输二维导航策略。除了导航策略外，视觉和语言嵌入层也可以传输[29]。受元学习[204]的成功推动，维变量技能转移（DVST）[205]获得了一个具有深度强化学习的元代理，然后使用一种称为维变量技能转移的方法将元技能转移到一个具有不同维度配置的机器人上。同样，Li等人。[206]提出了一种无监督的强化学习方法来学习可转移的元技能。朱等人。[63]将较长的导航指令分解为较短的导航指令，从而使模型能够很容易地转移到具有较长轨迹的导航任务中。夏普特等人。[207]提出了一种多任务模型，它可以联合学习多模态任务，并在整个任务中转移视觉语言知识。该模型采用双注意单元来分离视觉知识和语言知识，并相互协调。

王等人。[208]建议学习导航策略的环境无关表示，使模型能够在视觉语言导航（VLN）和对话历史导航（NDH）任务上执行。Yan等。[30]提出了MVV-IN，一种通过多模态输入获取可转移的元技能来处理新任务的方法。刘等人。[209]研究了如何让机器人融合和转移他们的经验，从而使它们能够有效地利用先验知识并快速适应新的环境。戈登等人。[210]提出将视觉感知和策略脱钩，以促进转移到新的环境和任务。

仿真转移在机器人控制[211]、[212]领域得到了很好的研究。Sadeghi等。[213]首先提出了一种基于学习的方法，该方法完全在模拟器中训练导航代理，然后将其转移到真实环境中，而不对任何真实图像进行微调。因此，袁等人。[214]采用一种模拟真实转移策略来学习导航控制器，使用端到端策略，将原始像素作为视觉输入映射到控制操作，没有任何形式的工程

特征抽取泰等人。[215]利用异步DDPG [77]算法对机器人进行仿真训练，并将学习到的控制器直接部署到真实的机器人上进行导航传输。鲁苏等人。[212]引入了一个渐进的网络，将学习到的策略从模拟转移到现实世界。同样，对抗性特征自适应方法[216]也适用于模拟到真实的策略转移[217]。深度强化学习策略的模拟实向转移可以应用于复杂的导航任务[218]，包括六条腿机器人[219]、足球比赛机器人[220]等。

### . 36总结

在本节中，我们首先比较了模拟环境和真实世界环境之间的差异。然后，我们推断导致域间隙的域间隙。最后，我们介绍了导航中的一些迁移学习工作，为解决这一问题提供了一个很有前途的方向。

### 7未来方向

虽然广泛的工作已经从不同的方面解决了导航问题，但目前的研究进展仍远非真正的人工智能。此外，目前的工作还不能建立一个强大的真实导航机器人。我们将解决具身人工智能的挑战总结到以下方面： 1)功能和性能受具体化环境的限制；2)导航问题定义不明确；3)具身人工智能代理在复杂环境中的性能仍然很差；4)感知自然语言难以学习；5)很难将训练有素的导航策略部署到现实世界的应用中。

**未来的环境。**高级的功能在

环境帮助导航模型获得高级能力。例如，与早期的实体环境相比，马特波特3D[36]中的大型场景首先需要导航模型来探索和记忆复杂的房间结构。视觉语言导航基准测试[12]使代理能够感知自然语言。像AI2-THOR [40]和iGibson [41]这样的交互式体现环境使代理能够执行



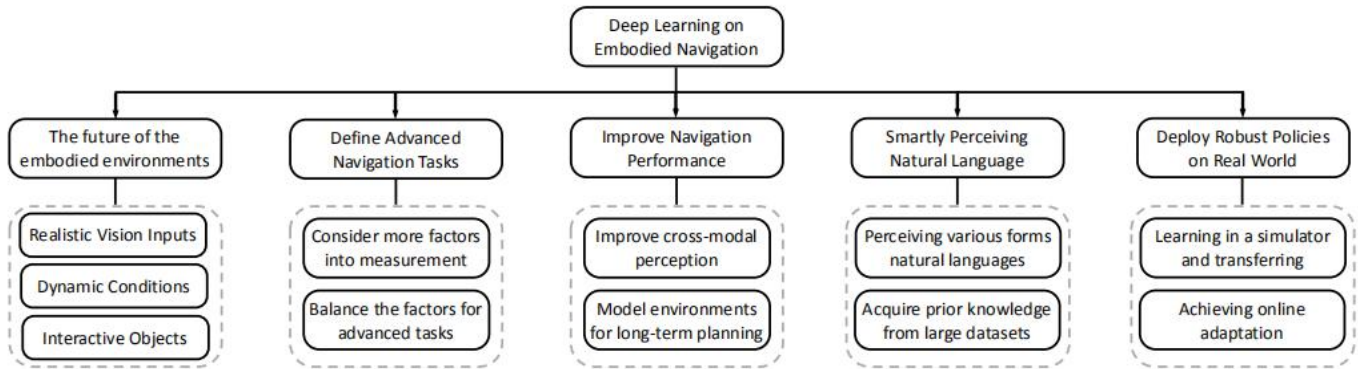


图13: 构建一个用于真实世界导航的先进机器人的未来方向的总结。

交互式操作。在交互式环境中学习到的代理能够移动对象、放置对象并打开门。

具有更多功能的环境是学习智能代理的基础。当房间中的对象具有演化条件时，代理必须能够处理动态环境。我们希望一个代理能够在有一个有许多障碍的房间内找到可能的道路，而不是在像 [12]、[13] 这样的可通航区域内导航。此外，我们需要一个交互式代理，它可以拾起和放下物体，移动椅子，并与人类互动。如果我们想在一个复杂的室内环境中建立一个强大的导航器，我们还需要考虑其他模式，如行走、跑步和攀爬。

**定义高级导航任务。** 尽管已经提出了许多具体的导航任务和导航度量，但什么是好的导航策略仍然不清楚。这个问题有两个方面：1) 需要考虑哪些因素；2) 如何平衡这些因素。正如我们在Sec 4.4 中分析的那样。3. 4. 精度和效率是评价导航性能的两个主要因素。然而，准确性和效率的重要性在不同的指标之间是不同的。最佳导航策略会根据不同的度量标准而有所不同。在 [221] 提出的交互式导航任务中，通过路径效率得分和努力效率得分来评估代理的性能。交互式导航任务在评估中改变分数权重，以测试代理在不同的设置中是否表现良好。然而，目前还不清楚这些因素的权重如何影响测试结果，以及在现实应用中是什么设置。

此外，在像 Help Anna [59] 或 RMM [166] 这样的问答设置中，会考虑到来自代理的询问或要求的频率。目前的方法认为，如果代理向人类请求更多的信息，则性能会较低。平衡提问的“成本”和导航的“成本”仍然是一个具有挑战性的问题。我们希望社区能够发布更多关于如何评估高级导航行为或比较代理和人之间的导航策略差异的著作。

**提高导航性能。** 目前的导航代理即使在简单的导航任务中，如点目标任务和目标目标任务仍然表现很差，如图所示。12 和无花果。11. 最先进的点目标任务模型在成功率 (SR) 和 37 上执行了 64.5%。在 SPL 上为 7%，而在目标目标任务中，最佳模型在成功率 (SR) 上为 21.08%，在 SPL 上为 8.38%。如图所示。4. 目前最先进的视觉语言导航任务模型 [152] 在 SR 中分别为 63% 和 SPL 中为 57%，而人类在 SR 中分别为 86% 和 76%

在 SPL。目前模型的导航性能仍远远不及人类的性能。此外，现有的模型在视觉和挑战性导航任务上通常表现较差。如选项卡所示。4. 在最先进的模型和人类基线之间大约有 19% 的性能差距。在交互式对话任务 [59], [166], [222] 中，代理用来与神谕交互的自然语言有许多错误，并且不流畅。最近提出的任务 [221] 中的基线很难完成任务。此外，导航机器人在现实世界中的性能不如在模拟环境中表现好。

通过参考最近的工作，有几个有希望的方向可以解决这些问题。变压器 [117] 具有特征提取和跨模态融合的能力。一些工作是 [151], [152] 建立了基于变压器的导航模型，在视觉语言导航任务中取得了巨大的成功，揭示了变压器结构有利于跨模态导航策略。夏普特等人。[125] 将神经 SLAM 模块构建成导航模型，并通过层次强化学习对模型进行训练。该模型能够学习房间的结构，并在具有连续状态空间的环境中执行一个稳健的低级导航策略。我们认为，基于模型的方法 [96]、[138] 和层次强化学习 [223] 是建立鲁棒导航模型的关键。**聪明地感知自然的语言。** 自然语言由于其多样性和复杂性，是机器人需要理解的一种复杂的形态。然而，在这个时候，教一个导航机器人学习理解语言需要大量的自然语言注释，并且每个注释都描述了一个轨迹、一个场景或一种行为的语义。语言注释可以是一个单词、一个句子、一个问答对或一个对话，这是相当昂贵和劳动密集型的。

即使我们有足够的语言注释来训练导航机器人，机器人仍然很难正确理解语言指令。例如，由于有许多自然语言变体来描述相同的轨迹或场景，使用轨迹-指令对监督代理可能会导致严重的过拟合。此外，感知自然语言的技能还需要先验知识。例如，“在客厅找到第四把椅子”要求代理能够计数，“安全导航到浴室”要求代理平稳转动，不碰任何物体。一些工作 [151], [152] 采用训练前的方法，在有先验知识的情况下获得更好的语言理解技能。基于这些工作的成功，我们认为从其他大规模语言数据集 [147]、[224]、[225] 学习并转移先验知识可能是解决理解自然的挑战的一个有前途的方向

语言说明。

**在现实世界中部署强大的策略。**尽管我们已经在模拟环境中获得了一个健壮的导航策略，但如何将该策略部署到现实世界中仍然具有挑战性。如Sec所示。6. 1、三个主要差异导致了巨大的模拟-实域差距：1）观测；2）行动空间；3）环境动态。巨大的模拟-真实领域差距阻碍了学习到的导航策略在现实世界中的直接部署。解决这个问题有两个方向。一种方法是构建真实的模拟器，包括逼真的视觉图像渲染机制、先进的物理传感器、障碍物物体、动态仿真、机器人车轮、齿轮等部件的仿真等。然而，这样一个真实的视觉模拟器是计算昂贵的。另一种解决模拟-真实部署问题的方法是通过迁移学习或元强化学习[204]，[226]来实现在线适应。这些方法使代理能够更改其策略以适应环境。该方法不仅具有较高的计算效率，而且在事故发生时具有较强的适应性。

## 8结论

本文通过总结数百部作品，对具体化的导航场景进行了全面的综述。我们彻底地调查了环境、任务和度量标准，以引入研究人员正在试图解决的问题。我们介绍了数百种方法来解决这些任务在具体的环境中，比较它们的差异。然后我们介绍了现实环境中的方法，并演示了巨大的领域差距是如何导致导航性能下降的。最后，分析了实体导航中存在的问题，给出了改善社区的四个方向。

## 参考文献

- [1] A. 克里耶夫斯基, 我. Sutskever G. E. 《影像学分类》与深度卷积神经网络, “ACM的通信”, 卷。60岁, 没有。6, pp. 84 - 90, 2017.
- [2] Y. Lecun Y. 孟加拉和G. 《深度学习》, 《自然》, 第1卷。521, pp. 436 - 444, 2015.
- [3] S. 特伦, 概率机器人学出版社, 2005年。
- [4] P. 悲哀, 行动是: 体现的基础交互, 2001年。
- [5] T. 克鲁斯。K. 潘迪, R. 阿拉米和A. Kirsch, “人类感知机器人导航: 一项调查”, “机器人和自动系统”, 卷。61岁, 没有。12, pp. 1726 - 1743, 2013.
- [6] G. N. 德苏扎和A. C. 《移动机器人导航的愿景》: A 调查, “TPAMI”, vol. 24日, 没有。2, pp. 237 - 267, 2002.
- [7] S. “机器人技术中的概率算法”, Ai杂志, 第1卷。21, pp. 93 - 109, 2000.
- [8] A. Kosaka和。 “快速视觉引导的移动机器人导航”使用基于模型的推理和不确定性的预测, ”在IROS, 卷。3, 1992, pp. 2177 - 2186.
- [9] M. 卡布卡和A. “移动机器人的位置验证使用”标准模式, ICRA, vol. 3, 没有。6, pp. 505 - 516, 1987.
- [10] Y. 吴, Y. 吴, G. Gkioxari和Y. 田, “构建可推广的代理”与一个现实和丰富的3d环境, ”在ICLR (工作室), 2018年。
- [11] M. 萨夫瓦。X. 张, A. Dosovitskiy, T. A. 放电室, 以及V. 科尔顿, “米诺斯: 在复杂环境下导航的多模式室内模拟器”。arXiv预印本, arXiv: 1712.03931, 2017年。
- [12] P. 安德森, 问。吴, D. Teney, J. 布鲁斯, M. 约翰逊, N. Sunderhauf, I. 里德, S. 古尔德和A. , “视觉与语言导航: 在真实环境中解释视觉基础的导航指令”, CVPR, 2018页, 页。3674 - 3683.
- [13] M. 萨夫瓦。科迪亚, 哦。Maksymets, Y. 赵, E. Wijmans, B. 耆那教徒J. 斯特劳布, J. 刘, 五。Koltun J. 马利克等人。 , 《栖息地: 体现ai研究的平台》, ICCV, 2019年, 页。9339 - 9347.
- [14] F. 夏, A. R. Zamir Z. 他, A. Sax, J. 马利克和S. 萨瓦雷斯, 吉布森环境: 具体化代理的现实感知”, CVPR, 2018, 页。9068 - 9079.
- [15] Y. 朱, R. Mottaghi, E. Kolve J. J. Lim. 古普塔, L. 费菲, 还有A. 法哈迪, “使用深度强化学习的目标驱动视觉导航”, ICRA, 2017年, 页。3357 - 3364.
- [16] A. 穆萨维安。Toshev, M. Fiser J. 科塞卡。瓦希德和J. 大卫“语义目标驱动导航的视觉表示”, 在ICRA, 2019, 页。8846 - 8852.
- [17] C. 谢, S. 帕蒂尔, T. 摩尔多瓦, S. Levine和P. Abbeel, “基于模型参数化物理模型和最优驱动探索的强化学习, ICRA, 2016, 页。504 - 511.
- [18] M. Jaderberg, V. Mnih, W. M. Czarnecki, T. Schaul, J. Z. Leibo, D. 银和K. 卡夫, “无监督辅助任务的强化学习”, ICLR 2017: ICLR 2017, 2017.
- [19] P. 米罗斯基, R. Pascanu, F. Viola, H. 索耶。巴拉德。Banino, M. Denil, R. Goroshin, L. Sifre, K. Kavukcuoglu, D. 库马兰和R. 哈德塞尔, 《学习在复杂的环境中导航》, 发表在ICLR, 2017年。
- [20] S. 古普塔, 五。Tolani, J. 戴维森。莱文, R. Sukhankar和J. 马利克“视觉导航的认知映射和规划”, IJCV, 第1卷。128年, 没有。5, pp. 1311 - 1330, 2020.
- [21] W. 齐, R. T. 穆拉普迪, S. Gupta和D. 拉马南, “学会搬家”“在ICLR, 2020年。”
- [22] D. S. Chaplot, D. 甘地, S. 古普塔。Gupta和R. 萨拉库蒂诺夫“学习探索使用主动神经大满贯”, 发表在ICLR, 2020年。
- [23] X. 王, W. 熊, H. 王和W. 杨王说, “看在你面前飞跃: 连接计划提前视觉和语言导航的无模型和基于模型的强化学习”, ECCV, 2018, 页。37 - 53.
- [24] D. 油炸, R. 胡, 五。Cirik, A. 罗巴赫, J. 安德烈亚斯, LP. .-莫伦西T. 伯格-柯克帕特里克, K. Saenko, D. 克莱因和T. 达雷尔, “视觉和语言导航的说话者追随者模型”, 发表在NeurIPS, 2018年, 页。3314 - 3325.
- [25] L. D. 杰克尔, E. 克罗特科夫, M. Perschbacher, J. 皮平和C. 苏里文“darpa lagr项目: 目标、挑战、方法和第一阶段的结果”, 《野外机器人杂志》, 第1卷。23日, 没有。1112, pp. 945 - 973, 2006.
- [26] S. D. 莫拉德, R. 麦加, R. P. K. Poudel S. Liwicki R. 西波拉, “在真实环境中与自动课程学习的体现视觉导航”, 在IEEE机器人和自动化字母, 卷。6、没有。2, 2021, pp. 683 - 690.
- [27] C. 索普, M. H. 赫伯特, T. Kanade和S. A. 剃须刀, “视觉和卡内基-梅隆导航实验室的导航”, “TPAMI”, 卷。10, 没有。3, pp. 362 - 373, 1988.
- [28] R. Traor, H. 卡塞尔斯-Dupr, T. 莱索特, T. 太阳, G. 蔡, N. 达兹-罗德·格斯和D. “不和谐: 通过政策蒸馏的持续强化学习”, NeurIPS深度强化学习研讨会, 2019.
- [29] H. 黄, 五。Jain H. Mehta. Ku, G. 马加尔海斯, J. Baldrige, 和E. 即, “视觉和语言导航中的可转移表示学习”, 发表在ICCV, 2019年, 页。7404 - 7413.
- [30] L. 燕, D. 刘, Y. 宋和C. 余, “多模式聚合方法”为记忆视觉-语音室内导航与元学习, ”在IROS, 2020, 页。5847 - 5854.
- [31] M. 费舍尔, D. 里奇, M. 萨夫瓦, T. Funkhouser和P. 汉拉恩“基于例子的3d对象安排的合成”, 计算机图形和交互技术国际会议, 卷。31日, 没有。6, p. 135, 2012.
- [32] A. 汉达, 五。Patraucean, S. 支架和R. Cipolla, “场景: 一个用于室内场景理解的注释模型生成器”, “在ICRA, 2016, pp. 5737 - 5743.
- [33] I. Armeni, S. 萨克斯, A. R. Zamir和S. 联合2d-3d语义用于室内场景理解的数据。”arXiv预印本arXiv: 1702.01105, 2017.
- [34] S. 宋, F. 余, A. 曾, A. X. 常, M. Savva和T. 芬克豪泽“来自单一深度图像的语义场景完成”, 发表在CVPR, 2017, pp. 190 - 198.
- [35] C. 燕, D. K. 米斯拉。贝内特。Walsman, Y. Bisk和Y. Artzi, “小屋: 康奈尔学院代理学习环境。”arXiv预印本, arXiv: 1801.07357, 2018年。
- [36] A. 张, A. 戴。Funkhouser, M. 哈尔伯, M. Niebner, M. 萨夫瓦, S. 宋, A. 曾和Y. 张, “马特波特3d: 室内环境中的rgb-d数据学习”, 第3DV, 2017, 页。667 - 676.
- [37] J. 斯特劳布, T. 惠兰, L. 妈妈, Y. 陈, E. Wijmans, S. 绿色, J. J. 恩格尔R. Mur-Artal, C. 任, S. Verma. 克拉克森, M. Yan, B. 稍微移动Y. Yan, X. 潘, J. Yon, Y. Zou, K. 莱昂, N. 卡特, J. 布里亚莱斯

T. 伊利诺斯州吉林厄姆。Mueggler, L. 佩斯奎拉, M. Savva, D. 巴特拉, H. M.



- Strasdat, R. D. Nardi M. Goesele, S. 洛夫格罗夫和R. A. “复制品数据集：一个室内空间的数字复制品。” *arXiv预印本arXiv: 1906.05797*, 2019年。
- [38] L. 泰和M. 刘, “深入进行认知探索对移动机器人的强化学习,” *arXiv预印本arXiv: 1610.01733*, 2016.
- [39] M. Deitke, W. 韩, A. 赫拉斯蒂. 肯巴维, E. Kolve, R. Mottaghi, J. 萨尔瓦多, D. 施文克, E. 范德比尔特, M. 沃林福德等人., “机器人：一个开放的模拟到真实体现ai平台”, CVPR, 2020, 页. 3164 - 3174.
- [40] E. Kolve, R. Mottaghi, D. 戈登, Y. 朱, A. Gupta和A. 法哈迪  
艾神：视觉ai的交互式3d环境。arXiv预印本, arXiv: 1712.05474, 2017年。
- [41] F. 夏, W. B. 沈, C. 李, P. Kasimbeg, M. E. 查普米, A. Toshev, R. 马丁和S. “交互式吉布森基准：混乱环境下交互式导航的基准”, 《IEEE机器人与自动化信件》, 第1卷. 5、没有. 2, pp. 713 - 720, 2020.
- [42] M. L. “马尔可夫游戏作为一个多智能体的框架“强化学习”, ICML, 1994, 页. 157 - 163.
- [43] M. Tan, “多智能体强化学习：独立vs”。“coopera-《主动代理》, 《ICML》, 1997页, 页. 487 - 494.
- [44] P. 安德森. X. 张, D. S. 夏普拉特. Dosovitskiy, S. 古普塔 V. Koltun J. Kosecka, J. 马利克, R. Mottaghi, M. Savva和. R. 扎米尔, “关于实体导航代理的评估问题。” *arXiv预印本, arXiv: 1807.06757*, 2018年。
- [45] S. Wani, S. 帕特尔, 美国Jain. X. 张和M. Savva, “组分：使用多对象导航测试语义映射记忆的基准测试. 33, 2020.
- [46] V. 耆那教, G. 马加尔海斯. Ku, A. Vaswani, E. Ie和J. Baldridge, “  
停留  
《路径：视觉与语言导航中的指令保真度》, ACL, 2019, 页. 1862 - 1872.
- [47] A. Ku, P. 安德森, R. 帕特尔. Ie和J. Baldridge, “房间间：具有密集时空基础的多语言视觉和语言导航”, 在EMNLP, 2020, 页. 4392 - 4412.
- [48] Y. 香港, C. 罗德里格斯, 问. 吴和S. 古尔德, “子指令感知”视觉和语言导航”, EMNLP, 2020, 页. 3360 - 3376.
- [49] J. Thomason, M. 穆雷, M. Cakmak和L. Zettlemoyer, “视觉和对话导航”, “CoRL, pp. 394 - 406, 2019.
- [50] S. 安托尔. 阿格拉瓦尔, J. 鲁, M. 米切尔, D. 巴特拉, C. L. Zitnick, 和D. 帕里克, “Vqa: 视觉问题回答”, 在ICCV, 2015, 页. 2425 - 2433.
- [51] P. 安德森, X. 他, C. 贝勒, D. Teney, M. 约翰逊, S. 古尔德和L. 张, “对图像字幕和视觉问题回答的自下而上和自顶向下的关注”, 发表在CVPR, 2018, 页. 6077 - 6086.
- [52] R. 克里希纳, Y. 朱, 哦. Groth, J. 约翰逊, K. 哈塔, J. 克拉维茨. 陈, Y. Kalantidis, L. J. -李, D. A. Shamma. S. 伯恩斯坦和L. “视觉基因组：使用众包密集的图像注释连接语言与视觉”, IJCV, 第1卷. 123, 没有. 1, pp. 32 - 73, 2017.
- [53] Z. 杨, X. 他, J. 高, L. 邓和A. Smola, “停滞不前的注意力图像问题回答的网络,” 在CVPR, 2016, 页. 21 - 29.
- [54] A. 达斯, S. 达塔, G. Gkioxari, N. 李, D. 帕里克和D. Batra, “体现回答问题,” 在CVPR研讨会 (CVPRW), 2018年, 页. 1 - 10.
- [55] Y. 齐, 问. 吴, P. 安德森, X. 王, W. Y. 王, C. 沈, 和 A. “可逆：真实室内环境中的远程体现视觉参考表达”, CVPR, 2020, 页. 9982 - 9991.
- [56] C. 陈, 美国耆那教, C. 五味子, S. V. A. 加里, Z. Al-Halah V. K. Ithapu, P. 罗宾逊和K. 格劳曼出版社, 《声音空间：三维环境中的视听导航》, ECCV出版社, 2020年, 第二页. 17 - 36.
- [57] L. Yu, X. 陈, G. Gkioxari, M. 坦萨尔, T. L. Berg和D. 巴特拉, “多目标体现了问题的回答,” 在CVPR, 2019, pp. 6309 - 6318.
- [58] D. 戈登. 肯巴维, M. Rastegari, J. 雷德蒙, D. 福克斯和 A. 法哈迪, “Iqa: 互动环境中的视觉问题回答”, CVPR, 2018年, 页. 4089 - 4098.
- [59] K. 阮和H. 多姆, “救命啊, 安娜! 基于视觉的导航通过回顾性好奇心鼓励模仿学习的自然多模态协助, EMNLP, 2019, 页. 684 - 695.
- [60] S. 古普塔, J. 戴维森. 莱文, R. Sukhankar和J. 马利克, “认知视觉导航的映射和规划”, CVPR, 2017, 页. 7272 - 7281.
- [61] D. K. 米斯拉, J. 兰福德和Y. “地图绘制说明和视觉效果”通过强化学习观察行为. “在EMNLP, 2017, 页. 1004 - 1015.
- [62] H. 陈, A. 苏尔, D. 米斯拉, N. Snavely和Y. Artzi, “触地得分：

- [63] W. 朱, H. 胡, J. 陈, Z. 邓, 五. 耆那教, E. Ie和F. Sha, "Babywalk: 通过小步进一步推进视觉和语言导航," ACL, 2020年, 页. 2539 - 2556.
- [64] D. J. 伯恩特和J. "使用动态时间扭曲来寻找模式在时间序列中," "在AAAIWS '94第三国际论文集知识发现和数据挖掘会议, 1994年, 第3页. 359 - 370.
- [65] H. Sakoe和. 《动态规划算法优化》  
对于语音识别, "IEEE交易的声学、语音和信号处理, 卷. 26日, 没有。1, pp. 43 - 49, 1978.
- [66] A. 瓦坎斯基, 我. 曼特, A. 爱尔兰和F. Janabi-Sharifi, "轨迹通过演示使用隐藏的马尔可夫模型和动态时间扭曲来学习机器人编程," "IEEE交易上的系统, 人, 和控制学, vol. 42岁, 没有。4, pp. 1039 - 1052, 2012.
- [67] E. J. 基奥和M. J. "扩大动态时间扭曲为数据挖掘应用程序," 在KDD, 2000, 页. 285 - 289.
- [68] G. Ilharco, V. Jain. Ku, E. Ie和J. 鲍德里奇, "一般评价"对于使用动态时间扭曲的指令条件导航, "ViGIL@NeurIPS, 2019".
- [69] V. Mnih. P. Badia, M. 米尔扎, A. 坟墓, T. 哈雷, T. P. Lillicrap, D. 银和K. "深度强化学习的异步方法", 《ICML》, 2016年, 页. 1928 - 1937.
- [70] A. TamarY. 吴, G. 托马斯. Levine和P. Abbeel, "价值迭代"网络", 发表在IJCAI, 2017, 页. 4949 - 4953.
- [71] L. 李, E. Parisotto, D. S. Chaplot, E. P. Xing和R. 萨拉库蒂诺夫《门控路径规划网络》, ICML, 2018年, 页. 2947 - 2955.
- [72] S. 透镜和M. "视觉声纳: 快速避障技术的使用"单眼视觉," 在IROS, 卷. 1, 2003, pp. 886 - 891.
- [73] E. 罗耶, J. Bom, M. DhomeB. Thuilot, M. LhuillierF. Marmoiton, "使用单眼视觉的户外自主导航", 发表在IROS, 2005年, 页. 1253 - 1258.
- [74] H. 哈达德, M. 卡蒂布, S. Lacroix和R. 查提拉, "反应性导航"在户外环境中使用潜在的领域," 在IROS, 卷. 2, 1998, pp. 1232 - 1237.
- [75] A. Remazeilles, F. 乔米特和P. "机器人的运动控制来自一种视觉记忆," 在ICRA, 卷. 5, 2004, pp. 4695 - 4700.
- [76] Y. 朱, D. 戈登, E. Kolve, D. 福克斯, L. 费菲, A. 古普塔, R. Mottaghi, 和A. 法哈迪, "使用深度后继表示的视觉语义规划", ICCV, 2017, 页. 483 - 492.
- [77] T. P. Lillicrap, J. J. 亨特. Pritzel, N. Heess, T. Erez, Y. 塔萨, D. 银和D. 维尔斯特拉, "深度强化学习的持续控制", 发表在ICLR, 2016年。
- [78] Y. 李和J. "学习视图和目标不变的视觉伺服法"关于导航," 在ICRA, 2020年, 页. 658 - 664.
- [79] T.-Y. 林, M. 梅尔, S. J. 贝隆吉, J. 海斯, P. Perona, D. Ramanan, P. 美元和C. 齐特尼克, "微软可可: 语境中的共同对象", ECCV, 2014, 页. 740 - 755.
- [80] A. Mousavian, H. 皮尔萨瓦什和J. "联合语义分割" - 深度卷积网络的深度估计, 在3DV, 2016, pp. 611 - 619.
- [81] W. 沈, D. 徐, Y. 朱, L. Fei-FeiL. Guibas和S. 萨瓦斯 "视觉导航的视觉表示的情境融合", ICCV, 2019, 页. 2881 - 2890.
- [82] Y. Lv, N. 谢, Y. 施, Z. 王和H. T. 沈, "提高目标"关注三维空间关系的驱动视觉导航," arXiv预印本arXiv: 2005.02153, 2020.
- [83] Q. 吴, K. 徐, J. 王, M. Xu, X. 龚和D. Manocha 基于强化学习的视觉导航与信息理论正则化, "在IEEE机器人和自动化字母, vol. 卷. 6, 没有。2, 2021, pp. 731 - 738.
- [84] A. KhoslaB. 安, J. J. Lim和A. 托拉尔巴, "看超越可见场景", 在CVPR, 2014, 页. 3710 - 3717.
- [85] S. 勃拉姆哈特和J. 海斯, "深度导航: 学习在大城市中导航," 在CVPR, 2017, pp. 3087 - 3096.
- [86] P. 米罗斯基, M. K. 格兰姆斯, M. 马林诺夫斯基. M. 赫尔曼, K. 安德尔儿子, D. Teplyashin, K. 西蒙尼安, K. 卡武库格鲁, A. 齐塞尔曼和R. 哈德塞尔, "学习在没有地图的城市中导航", 在NeurIPS, 卷. 31, 2018, pp. 2419 - 2430.
- [87] C. 陈, S. Majumder, Z. 哈拉, R. 高, S. K. 罗摩克里希南和K. 格劳曼, "学习为视听导航设置航路点", 发表在ICLR, 2021年.
- [88] C. 多尔施, A. Gupta和A. A. "无监督的视觉表征-通过上下文预测学习, ICCV, 2015, 页. 1422 - 1430.
- [89] M. 诺鲁齐和P. "视觉表征的无监督学习-《ECCV》, 2016年, 页. 69 - 84.
- [90] R. 张, P. Isola和A. A. Efros, "彩色图像彩色", 在ECCV, 2016, pp. 649 - 666.

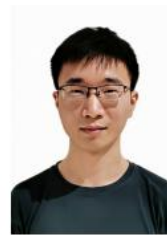
- [91] Y. 太阳, X. 王, Z. 刘, J. 米勒, A. Efros和M. “硬”测试时间在分布转移下进行泛化的自我监督训练”, 在ICML, 2020, 页. 9229 – 9248.
- [92] T. Schaul, D. 霍根, K. Gregor和D. 银器, “普遍价值函数” “近似者”, ICML出版社, 2015年, 页. 1312 – 1320.
- [93] A. 多索维茨基和V. 科尔顿, “通过预测未来来学会行动”, 在ICLR, 2016年.
- [94] M. 肯普卡, M. Wydmuch, G. Runc, J. Toczek和W. Jakowski, “末日: 基于末日的视觉强化学习研究平台”, 在CIG. IEEE, 2016, 页. 1 – 8.
- [95] J. 是的, D. 巴特拉, E. Wijmans和. “辅助任务加速” 学习点和目标导航.” *arXiv预印本arXiv: 2007.04561, 2020年*.
- [96] J. 是的, D. 巴特拉. Das和E. “辅助任务和探索” 启用对象. “*arXiv预印本arXiv: 2104.04112, 2021年*.”
- [97] X. 是的, Z. 林, JY. .-李, J. 张, S. 郑和Y. 杨, “Gaple: 室内环境中机器人物体搜索的广义接近策略学习”, “IEEE机器人自动化与自动化信件”, 卷. 4, 没有. 4, pp. 4003 – 4010, 2019.
- [98] GH. .-刘, A. Siravuru, S. P. 塞尔瓦拉杰, M. M. Veloso和G. Kan- “学习自主导航的端到端多模态传感器策略. CoRL, pp. 249 – 261, 2017.
- [99] S. 达塔, O. Maksymets, J. 霍夫曼, S. 李, D. Batra和D. 帕里克, “在 - 为更现实的点目标导航代理提供自我中心定位”, *arXiv预印本arXiv: 2009.03231, 2020*.
- [100] R. 比加齐, F. 兰迪, M. Cornia, S. Cascianelli, L. 巴拉迪和 R. “探索和解释: 自我监督的导航和叙述”, 在ICPR, 2021年, 页. 1152 – 1159.
- [101] V. 院长, S. 图尔西亚尼和A. 古普塔, “看, 听到, 探索: 好奇心通过视听联想”, 在NeurIPS, vol. 33, 2020.
- [102] H. Durrant-怀特和T. 贝利, “同时定位和 地图: 第一部分”, “IEEE《机器人与自动化》杂志, 第1卷. 13日, 没有. 2, pp. 99 – 110, 2006.
- [103] M. 菲利彭科和我. “各种大满贯系统的比较” 用于室内环境中的移动机器人”, 在IS, 2018, 页. 400 – 407.
- [104] S. Se, D. G. 劳和J. J. “移动机器人定位和 使用尺度确定的视觉地标进行不确定性映射, 国际机器人研究杂志, 卷. 21日. 8, pp. 735 – 758, 2002.
- [105] C. F. 奥尔森, L. H. Matthies, M. 直升机和M. W. 迈蒙, “罗孚” 使用立体自我运动的导航, “机器人和自主系统”, 卷. 43岁, 没有. 4, pp. 215 – 229, 2003.
- [106] Davison, “实时同步定位和映射与一个单一的 相机”, “ICCV, vol. 2, 2003, pp. 1403 – 1410.
- [107] J. Engel, V. Koltun和D. 创作者, “直接稀疏速速法”, TPAMI, 卷. 40岁, 没有. 3, pp. 611 – 625, 2017.
- [108] J. 恩格尔, T. 施普斯和D. “大满贯: 大规模直接” 单目大满贯, “在ECCV. 施普林格, 2014, 页. 834 – 849.
- [109] R. Mur-Artal和J. D. “球体大满贯2: 一个开源的大满贯 单眼、立体声和rgb-d相机系统”, “IEEE机器人交易, 卷. 33日, 没有. 5, pp. 1255 – 1262, 2017.
- [110] K. Tateno, F. 汤巴里, 我. Laina和N. Navab, “Cnn-slam: 实时 具有学习深度预测的密集单眼大满贯”, “在CVPR, 2017, pp. 6243 – 6252.
- [111] C. 王, J. 米格尔·布埃纳波萨达, R. 朱和S. Lucey, 学习 使用直接方法的单眼视频的深度”, 在CVPR, 2018, pp. 2022 – 2030.
- [112] N. 杨, L. v. Stumberg, R. 王和D. Cremers, “D3vo: 深度, 单眼视觉测程法的深度姿态和深度不确定性”, 在CVPR, 2020, pp. 1281 – 1292.
- [113] A. 盖格, P. Lenz和R. “我们准备好要自治了吗? 驾驶kitti视觉基准套件”, 在CVPR. IEEE, 2012, 页. 3354 – 3361.
- [114] M. Burri, J. Nikolicp. Gohl, T. 施耐德, J. 雷德, S. 奥马里, M. W. Achtelik和R. “欧洲微型飞行器数据集”, 国际机器人研究杂志, 第1 卷. 35岁, 没有. 10, pp. 1157 – 1163, 2016.
- [115] E. 副盐和R. “神经地图: 结构化记忆” 深度强化学习2017年, 在ICLR网站上.
- [116] J. 张, L. 泰, J. Boedecker, W. 伯加德和M. 刘, “神经” 《slam: 学习用外部记忆来探索》, “*arXiv预印本arXiv: 1706.09520, 2017*.”
- [117] A. Vaswani, N. Shazeer, N. 帕尔马, J. Uszkoreit, L. 琼斯. N. 戈麦斯 L. 凯泽和我. 《注意力是你所需要的》, 在NeurIPS, 卷. 30, 2017, pp. 5998 – 6008.
- [118] J. F. 亨利克斯和A. 维达尔迪, “Mapnet: 一种异中心的空间记忆” 对于映射环境”, 在CVPR, 2018, 页. 8476 – 8484.
- [119] T. JakschR. 奥特纳和P. 奥尔, “接近最优的遗憾界限” 强化学习”, JMLR, vol. 11日, 没有. 51, pp. 1563 – 1600, 2010.
- [120] C. 金, Z. Allen-Zhu. Bubeck和M. I. 这是可以证明的 效率高的在NeurIPS中, 卷. 31, 2018, pp. 4863 – 4873.
- [121] M. G. 阿扎尔, 我. 奥斯班德和R. “极大极小遗憾的界限” 强化学习”, 在ICML, 2017, 页. 263 – 272.
- [122] T. 陈, S. Gupta和A. “学习探索政策” 导航”, 在ICLR, 2019.
- [123] A. G. 巴托和S. “在等级制度方面的最新进展” 强化学习, “离散事件动态系统”, 第1卷. 13日, 没有. 1, pp. 41 – 77, 2003.
- [124] P. 达扬和G. E. 《封建强化学习》, 《神经学》, 1992年第5卷, 页. 271 – 278.
- [125] D. S. Chaplot, R. 萨拉库蒂诺夫. Gupta和S. 古普塔, “神经” 视觉导航的拓扑大满贯”, 在CVPR, 2020, 页. 12 875 – 12 884.
- [126] D. S. Chaplot, D. P. 甘地. Gupta和R. R. 萨拉库蒂诺夫, “对象” 使用目标导向语义探索的目标导航. 33, 2020.
- [127] I. 苏特斯克利, 哦. 乙烯基和Q. V. “序列到序列学习” 与神经网络, “NeurIPS, vol. 27, pp. 3104 – 3112, 2014.
- [128] X. 王, 问. 黄, A. Celikyilmaz, J. 高, D. 沈, YF. .-王 W. Y. 王和L. “视觉语言导航的强化跨模态匹配与自监督模仿学习. 在 CVPR, 2018, pp. 6629 – 6638.
- [129] CY. .-妈妈, J. 陆, Z. 吴, G. AlRegib, Z. 基拉, R. Socher和C. 熊, “通过辅助进度评估的自我监控导航代理”, 发表在ICLR, 2019年.
- [130] CY. .-马, Z. 吴, G. AlRegib, C. 熊和Z. 基拉, “遗憾的是 “代理: 通过进度估计的启发式辅助导航”, 在CVPR, 2019年, 页. 6732 – 6740.
- [131] L. Ke, X. 李, Y. Bisk, A. Holtzman, Z. 干, J. 刘, J. 高, Y. 崔, 和S. “战术倒带: 在视觉和语言导航中通过回溯的自我纠正”, CVPR, 2019年, 页. 6741 – 6749.
- [132] P. 安德森. 施里瓦斯塔瓦, D. 帕里克, D. Batra和S. 李, 《追逐》 幽灵: 贝叶斯状态跟踪指令”, NeurIPS, 2019, 页. 371 – 381.
- [133] R. 琼斯科夫斯基和O. 布洛克, “端到端可学习的直方图过滤器, ” 2016.
- [134] N. 斯里瓦斯塔瓦, G. 辛顿, A. 克里耶夫斯基, 我. 萨特斯克利和R. Salakhutdinov- “辍学: 防止神经网络过拟合的简单方法”, 第11卷. 15日, 没有. 1, pp. 1929 – 1958, 2014.
- [135] K. 西蒙尼安和A. “非常深的卷积网络为 “大规模图像识别”, 发表在ICLR, 2015年.
- [136] K. 他, X. 张, S. 任和J. 孙先生, “对图像的深度学习残差学习” 《CVPR》, 2016年, 页. 770 – 778.
- [137] H. 谭, L. Yu和M. “学会驾驭看不见的环境- 背景: 反翻译, NAACL, 2019, 页. 2610 – 2621.
- [138] Y. 朱, F. 朱, Z. 詹, B. 林, J. 焦, X. 张和X. 梁, “愿景” 通过探索跨模态记忆的对话框导航”, 在CVPR, 2020, pp. 10 730 – 10 739.
- [139] T. J. .-傅, X. E. 王, M. F. 彼得森. T. 格拉夫顿. P. 埃克斯坦 和W. Y. 王, “通过对抗性路径抽样的反事实视觉和语言导航”, 《EC CV》, 2019年, 页. 71 – 86.
- [140] J. Thomason, D. 戈登和Y. 移动基线: 单线 “视觉导航和qa上的模态表现”, 在NAACL, 2019, 页. 1977 – 1983.
- [141] H. 黄, 五. JainH. 梅赫塔, J. 鲍德里奇和E. 即 “多模态” 视觉和语言导航的判别模型”, 在RoboNLP, 2019, 页. 40 – 49.
- [142] H. 王, 问. 吴和C. 软专家奖励视觉学习 和语言导航, ECCV, 2020, 页. 126 – 141.
- [143] Q. 夏, X. 李, C. 李, Y. Bisk, Z. Sui, J. 高, Y. 崔和N. A. 铁匠 “面向视觉和语言导航的多视图学习.” *arXiv预印本arXiv: 2003.00857, 2020年*.
- [144] Y. 香港, C. 罗德里格斯. 齐, 问. 吴和S. 古尔德, “语言和 为代理导航的视觉实体关系图”, 在NeurIPS, vol. 33, 2020.
- [145] R. 胡, D. 炸了. 罗巴赫, D. 克莱因. 达雷尔和K. Saenoco 你在看吗? 视觉和语言导航的多种模式, ACL, 2019, 页. 6551 – 6557.
- [146] B. 施, X. 王, P. Lyu, C. Yao和X. 白, 《鲁棒的场景文本》 “识别与自动修正”, 在CVPR, 2016, 页. 4168 – 4176.



- [147] J. Devlin, 兆瓦. .-张, K. 李和K. N. 托塔诺瓦, “伯特: Pre-语言理解的深度双向转换器的培训, NAACL, 2018, 页. 4171 - 4186.
- [148] H. 谭和M. 莱克斯默特: 学习跨模态编码器《变形金刚的表示》, 在EMNLP, 2019, 页. 5099 - 5110.
- [149] X. 李, C. 李, 问. 夏, Y. Bisk, A. Celikyilmaz, J. 高, N. A. 史密斯和Y. 崔, “带有语言预训练和随机抽样的鲁棒导航”, 发表在EMNLP, 2019年, 页. 1494 - 1499.
- [150] A. Majumdar. 斯里瓦斯塔瓦, S. 李, P. 安德森, D. 帕里克和D. Batra, “利用网络上的图像-文本对改善视觉和语言导航”, 发表在ECCV, 2020年, 页. 259 - 274.
- [151] W. Hao, C. 李, X. 李, L. 卡林和J. 高, “通过预训练学习视觉和语言导航的通用代理”, CVPR, 2020年, 页. 13 137 - 13 146.
- [152] Y. 香港, Q. 吴, Y. 齐, C. 罗德里格斯-奥帕佐和S. 古尔德, “Vlnbert: A 导航的循环视觉和语言伯特”, CVPR, 2021, 页. 1643 - 1653.
- [153] A. 《递归神经网络的自适应计算时间》, *arXiv预印本*, *arXiv: 1603.08983, 2016年*.
- [154] A. 达斯, G. Gkioxari, S. 李, D. 帕里克和D. 巴特拉, “神经模块化” “控制具体化的问题回答”, 在CoRL, 2018年.
- [155] A. 阿南德. Belilovsky, K. 哈斯特纳. 拉洛谢尔和A. C. 考维尔 “具体化qa的盲法基线.” *arXiv: 计算机视觉与模式识别公司, 2018年*.
- [156] Y. 吴, L. 姜和Y. 杨: 一个简单的基线此外, “*IEEE图像处理事务*, 卷. 29, pp. 3984 - 3992, 2020.
- [157] E. Wijmans, S. 达塔, O. Maksymets. 达斯, G. Gkioxari, S. 避风处I. 埃萨, D. 帕里克和D. Batra, “在具有点云感知的逼真环境中体现回答问题”, 发表在CVPR, 2019年.
- [158] H. 罗, G. 林, Z. 刘, F. 刘, Z. 唐和Y. 姚, “Segeqa: 视频基于分割的视觉注意力回答, ICCV, 2019, pp. 9667 - 9676.
- [159] E. Ilg, N. 梅耶尔, T. Saikia, M. 库珀, A. 多索维茨基和T. Brox, “流网2.0: 深度网络光流估计的演变”, CVPR, 2017, 页. 2462 - 2470.
- [160] J. 李, S. 唐, F. 吴和Y. 庄, 《心灵一起行走》*图像增强, ACM多媒体, 2019, 页. 1211 - 1219*.
- [161] S. 棕褐色, W. 香, H. 刘, D. 郭和F. 阳光, “多代理体现”在互动环境下回答问题”, 在ECCV, 2020, 页. 663 - 678.
- [162] Y. 是的, Z. 赵, Y. 李, L. 陈, J. 肖和Y. 庄, “视频问题通过属性增强注意网络学习回答”, ACM, 2017, 页. 829 - 832.
- [163] C. Cangea, E. Belilovsky, P. 李和A. C. “视频卡: 弥合视觉上和具体化的问题回答之间的差距.” 在ViGIL@NeurIPS, 2019, p. 280.
- [164] Y. 邓, D. 郭, X. 郭, N. 张, H. 刘和F. Sun, Mqa: 通过机器人操作来回答这个问题.” 17, 2021.
- [165] D. 尼尔森, A. 皮林宁, E. Grtner和C. 斯明奇斯库, “体现”在语义分割中的视觉主动学习.” 在AAAI, 2020年, 页. 2373 - 2383.
- [166] H. R. 罗马, Y. Bisk, J. 托马森. 凯利基尔马兹和J. 高, “Rmm: 对话导航的递归心理模型”, 在EMNLP, 2020, 页. 1732 - 1745.
- [167] J. 托马森. 帕德马库马尔, J. Sinapov, N. 沃克, Y. 江, H. 耶多德, J. 哈特, P. 石头和R. J. 穆尼, “通过人机对话提高自然语言理解”, ICRA. IEEE, 2019, 页. 6934 - 6941.
- [168] S. Tellex, R. A. 膝盖, A. 李, D. Rus和N. 罗伊, “寻求帮助使用逆语义”, “在机器人: 科学与系统2014, 卷. 10, 2014.
- [169] K. 阮, D. Dey, C. 布罗克特和B. 多兰说, “基于视觉的导航” CVPR, 2019, 页. 12 527 - 12 537.
- [170] H. 德弗里斯, K. Shuster, D. 巴特拉, D. 帕里克, J. 韦斯顿和D. Kiela, “谈论行走: 通过扎根对话导航纽约市”, *arXiv: 人工智能, 2018年*.
- [171] M. 哈恩, J. Krantz, D. 巴特拉, D. 帕里克, J. M. Rehg, S. 李, 还有P. 安德森, “你在哪里? 本地化从具体化的对话框”, 在EMNLP, 2020, pp. 806 - 822.
- [172] U. Jain, L. 韦斯, E. Kolve, M. Rastegari, S. 拉泽布尼克. 法哈迪. G. 施温和A. , “两个身体问题: 协同视觉任务完成”, CVPR, 2019年, 页. 6689 - 6699.
- [173] A. 辛格. Jain和S. 苏赫巴托, “学习何时进行交流” “在多智能体合作和竞争任务中的规模”, 发表在ICLR, 2018年.
- [174] M. 玛吉, S. 诺加, C. J. 海耶斯, S. M. 卢金, J. Bloecker, E. 保持器, 和C. R. Voss, “多机器人与人类对话的研究平台”, 《计算语言学协会 (演示) 北美分会2019年会议论文集》, 2019年, 页. 132 - 137.
- [175] S. 班纳吉, J. 托马森和J. J. “机器人俚语的基准测试: 对话框引导的机器人定位和导航.” 《机器人学》, 2020年.
- [176] U. 穆勒, J. 本, E. Cosatto, B. Flepp和Y. L. 村, “越野障碍” “通过端到端学习而进行的回避.” 西特塞尔, 2006年, 页. 739 - 746.
- [177] R. 哈德塞尔, P. Sermanet, J. 本, A. Erkan, M. Scoffier, K. Kavukcuoglu, U. 穆勒和Y. 勒村, “学习自动越野驾驶的远程愿景”, 《野外机器人杂志》, 第1卷. 26日, 没有. 2, pp. 120 - 144, 2009.
- [178] W. 张, K. 刘, W. 张, Y. 张和J. “室内和室外环境中无线定位的深度神经网络”, *神经计算*, 卷. 194, pp. 279 - 287, 2016.
- [179] J. 张, J. T. Springenberg, J. 博德克和W. 伯加德, “深”使用跨类似环境导航的后续功能进行强化学习”, 在IROS, 2017, 页. 2371 - 2378.
- [180] V. Mnih, K. Kavukcuoglu, D. 银, A. A. Rusu, J. Veness, M. G. 贝莱马雷. 坟墓, M. 里德米勒, A. K. Fidjeland, G. Ostrovski, S. 彼得森, C. 贝蒂. Sadik, 我. Antonoglou, H. 国王, D. Kumaran, D. Wierstra, S. Legg和D. “通过深度强化学习的人类层面的控制”, 《自然》, 卷. 518, 没有. 7540, pp. 529 - 533, 2015.
- [181] G. 卡恩, A. Villafior, P. Abbeel和S. 莱文, “可作曲家的行动”条件预测: 机器人导航的灵活非策略学习, 在CoRL. PMLR, 2018, pp. 806 - 816.
- [182] S. 罗斯, N. 梅利克-巴克哈达罗夫, K. S. 山卡, A. 温德尔, D. Dey, J. A. 巴格内尔和M. 赫伯特, “在杂乱的自然环境中学习单眼反应性uav控制”, 在ICRA中. IEEE, 2013, 页. 1765 - 1772.
- [183] T. Manderson, J. C. G. 希格拉, S. 瓦普尼克, JF. .-特伦布雷, F. 什库尔蒂 D. Meger和G. Dudek, “基于视觉的水下导航的目标条件政策”, 在RSS, 卷. 16, 2020.
- [184] J. 博伦斯坦和Y. “面向快速移动设备的实时避障功能” “杂乱环境中的机器人.” IEEE, 1990年, 页. 572 - 577.
- [185] D. 哈和J. 《世界模型》, *arXiv预印本 arXiv:1803.10122, 2018*.
- [186] A. 弗朗西斯. 浮士德, HT. .-L. 蒋介石. Hsu, J. C. Kew, M. 提要器, 和TW. .-E. 李, “远程室内导航”, *IEEE《机器人学报》*, 第1卷. 36岁, 没有. 4, pp. 1115 - 1134, 2020.
- [187] D. 沙阿, B. 艾森巴赫, G. 卡恩, N. 莱茵哈特和S. 莱文, “Ving: 学习具有视觉目标的开放世界导航”, *arXiv预印本 arXiv: 2012.09812, 2020*.
- [188] S. “概率机器人”, 《ACM的通信》, vol. 45, 不3, pp. 52 - 57, 2002.
- [189] S. M. LaValle, *规划算法*. 剑桥大学出版社, 2006年.
- [190] A. J. 戴维森和D. W. “移动机器人本地化使用主动愿景”, 在ECCV. 施普林格, 1998年, 页. 809 - 825.
- [191] R. Sim和J. J. “基于自主视觉的探索和使用混合映射和线-黑线粒子滤波器进行映射”, 在IROS中. IEEE, 2006年, 页. 2082 - 2089.
- [192] G. Berseth, D. 耿, C. M. 德文, N. Rhinehart, C. 芬恩, D. 杰亚拉曼, 和S. Levine, “傻笑: 在不稳定的环境中意外地最小化强化学习”, 发表在ICLR, 2021年.
- [193] A. 纳加班迪, 我. Clavera, S. 刘, R. S. 担心, P. Abbeel, S. 莱文和C. 芬恩, “通过元强化学习学习在动态的、现实世界的环境中适应”, 发表在ICLR, 2018年.
- [194] K. 李, Y. Seo, S. 李, H. 李和J. “情境感知动态”在基于模型的强化学习中的泛化模型”, 在ICML, vol. 1, 2020, pp. 5757 - 5766.
- [195] E. 维曼斯. 卡迪亚, A. 莫斯科斯. 李, 我. 埃萨, D. 帕里克, M. 萨夫瓦, 和D. Batra, “Dd-ppo: 从25亿帧中学习近乎完美的点目标导航器”, 发表在第八届ICLR, 2020年.
- [196] H. 王, W. 王, T. 舒, W. 梁和J. 沈, “主动视觉”视觉语言导航的信息收集, ECCV, 2020, 页. 307 - 322.
- [197] C. 芬恩和S. 莱文, “规划机器人运动的深度视觉预见”, 在ICRA. IEEE, 2017, 页. 2786 - 2793.
- [198] N. 马龙和V. 设计一个情境判断在虚拟学习环境中评估11结构问题解决的测试”, 2012.

- [199] S. K. 罗摩克里希南, Z. Al-Halah和K. 格劳曼, “入住率”对高效探索和导航的预期”, ECCV, 2020, 页. 400 - 418.
- [200] K. Lobos-Tsunekawa, F. Leiva和J. “视觉导航”对于使用深度强化学习的两足类人机器人。3、没有。4, 2018, pp. 3247 - 3254.
- [201] J. 布鲁斯, N. Snderhauf, P. 米罗斯基, R. Hadsell和M. 米尔福德, “一个基于交互式回放的机器人导航的射击强化学习。” arXiv预印本, arXiv: 1711.10137, 2017年。
- [202] M. 法伊弗, S. 舒克拉, M. 图尔切塔, C. 卡迪纳, A. 克劳斯, R. Siegwart, 和J. I. 涅托, “强化模仿: 利用之前的演示进行无峰导航的高效深度强化学习样本”, 《IEEE机器人与自动化信件》, vol. 3、没有。4, 2018, pp. 4423 - 4430.
- [203] A. A. Rusu, S. G. 科尔梅纳雷霍, C. Gulcehre, G. 德斯贾丁斯, J. 柯克帕特里克, R. Pascanu, V. Mnih, K. 卡武库格魯和R. 哈德塞尔, “政策蒸馏”, 发表在ICLR, 2016年。
- [204] C. 芬恩, P. Abbeel和S. “模型不可知论的元学习” “深度网络的快速适应”, 在ICML, 2017, 页. 1126 - 1135. [205] W. 张, Y. 张和N. 刘, “无地图导航: 一个dr1-基于不同尺寸机器人的控制器,” arXiv预印本arXiv: 2002.06320, 2020.
- [206] J. 李, X. 王, S. 唐, H. 施, F. 吴, Y. 庄和W. Y. 王 “具体化导航的可转移元技能的无监督强化学习”, 发表在CVPR, 2020年, 页. 12 123 - 12 132.
- [207] D. S. Chaplot, L. 李, R. 萨拉库蒂诺夫, D. 帕里克和D. 巴特拉 “体现多模式多任务学习”, 在IJCAI, 卷. 3, 2020, pp. 2442 - 2448.
- [208] X. E. 王五世. 耆那教, E. Ie, W. Y. 王, Z. 科拉雷瓦和S. 拉维, “自然语言基础导航的环境不可知多任务学习”, ECCV, 2020年, 页. 413 - 430.
- [209] B. 刘, L. 王和M. 刘, “终身联邦强化学习”: 一种用于云机器人系统导航的学习架构。4、没有。4, 2019, pp. 4555 - 4562.
- [210] D. 戈登. 卡迪亚, D. 帕里克, J. 霍夫曼和D. Splitnet: Sim2sim和任务2任务转移”, 在ICCV, 2019, pp. 1022 - 1031.
- [211] J. 托宾. 方, A. 雷, J. 施耐德, W. Zaremba和P. 阿贝尔, “将深度神经网络从模拟转移到现实世界的域随机化”, 在IROS中说。IEEE, 2017, 页. 23 - 30.
- [212] A. A. Rusu, M. Veerk, T. Rothrl, N. Heess, R. 帕斯卡努和R. 哈德塞尔 “模拟到真实的机器人从像素学习与进步的网,” 在CoRL. PMLR, 2017, pp. 262 - 270.
- [213] F. 萨德吉和S. 莱文, “Cad2rl: 真正的单图像飞行没有a 单一的真实图像,” 在RSS, vol. 13, 2017.
- [214] W. 元, K. 悬挂, D. Kragic, M. Y. 王和J. A. 鹤, “具有深度强化学习和模拟与现实转移的端到端非理解重排”, 《机器人与自主系统》, 第1卷. 119, pp. 119 - 134, 2019.
- [215] L. 泰. 保罗和M. 刘, “虚拟到真实的深度强化学习: 连续控制移动机器人的无级导航,” 在IROS. IEEE, 2017, 页. 31 - 36.
- [216] E. 曾, J. 霍夫曼, K. Saenko和T. 达雷尔, “对抗性的描述-原生域适应”, CVPR, 2017, 页. 7167 - 7176.
- [217] F. 朱, L. 朱和Y. 杨, “模拟联合加固转移 3d室内导航”, CVPR, 2019, 页. 11 388 - 11 397.
- [218] W. 赵, J. P. 奎拉尔塔和T. 韦斯特隆德, “机器人技术深度强化学习中的模拟到真实的转移: 一个调查”, 发表在SSCI, 2020年, 页. 737 - 744.
- [219] B. 秦, Y. 高和Y. 白, “从模拟到现实: 六条腿的机器人控制与”深度强化学习和课程学习,” 在ICRAE, 2019年。
- [220] H. F. Bassani, R. A. 德尔加多, J. N. de O. 利马JuniorH. R. 梅代罗斯 P. H. M. 布拉加和A. “通过强化学习踢足球, 并在现实世界中应用模拟现实来踢足球,” CoRR, 2020年。
- [221] B. 沈, F. 夏, C. 李, R. Martn-Martn, L. 风扇, G. 王, S. 布赫, C. D' Arpino, S. 斯里瓦斯塔瓦, L. P. Tchapi, M. E. Tchapi, K. 维尼奥, L. Fei-Fei和S. “伊吉布森, 一个在大型现实场景中进行交互式任务的模拟环境。” arXiv预印本arXiv: 2012.02924, 2020年。
- [222] Y. 朱, Y. 翁, F. 朱, X. 梁, 问. 是的. 卢和J. 焦, “自我”为现实世界视觉对话导航的动机沟通代理,” 在ICCV, 2021年。
- [223] L. 杰尼, Z. IsteneP. KorondiH. 桥本, “等级强化” “利用智能空间概念的机器人导航学习”, 2007年第11届智能工程系统国际会议. IEEE, 2007, 页. 149 - 153.

- [224] Z. 杨, Z. 戴, Y. 杨, J. G. 卡博内尔, R. 萨拉库蒂诺夫和 Q. V. “Xlnet: 语言理解的广义自回归预训练”, 在NeurIPS, 第1卷. 32, 2019, pp. 5753 - 5763.
- [225] J. 李, W. Yoon, S. 金, D. 金, S. 金, C. H. 所以, J. 康, “BioBERT: 一种用于生物医学文本挖掘的预先训练过的生物医学语言表示模型。” 生物信息学, 第1卷. 36岁, 没有。4, pp. 1234 - 1240, 2019.
- [226] S. 莱文和P. “在引导下学习神经网络策略” 在未知动态下的政策搜索,” 在NeurIPS, vol. 27, 2014, pp. 1071 - 1079.



2017年获得位于中国北京的北京航空航天大学软件工程学院学士学位。他目前正在攻读博士学位。D. 莫纳什大学信息技术学院, 教授指导。张晓军。他的研究兴趣包括机器学习、深度学习和强化学习。

李朱获得了B. S.。2013年获得中国广州中山大学软件工程专业学位。自2015年起, 她一直是中国科学院大学电子、电气和通信工程学院计算机科学专业的博士生。她目前的研究兴趣包括物体识别、场景理解、弱监督学习和视觉推理。

文森特·李目前是澳大利亚莫纳什大学印度理工学院数据科学和人工智能系机器学习和深度学习学科的副教授。他是美国IEEE协会的高级成员。从1988年到1991年, 他获得了澳大利亚州联邦政府的奖学金, 在澳大利亚新南威尔士州的新堡大学攻读博士学位。1973年至1974年, 他获得了国防部颁发的联合研究奖学金

(新加坡)和国防部(英国)在英国皇家空军学院系统的研究生学习。2006年11月至2007年3月, 他在清华担任访问学者; 1994年7月至1995年6月, 他还在新加坡客座教授。

梁晓丹, 现任中山大学副教授。她是卡内基梅隆大学机器学习系的博士后研究员, 与教授合作。邢立星, 2016年至2018年。2016年获得中山大学博士学位, 由梁林指导。她发表了几个与人类相关的分析的前沿项目, 包括人类解析、行人检测和实例分割、2D/3D人体姿态估计和活动识别。

张晓军, 澳大利亚理工大学计算机技术学院副教授。在加入RMIT之前, 他曾是澳大利亚维克市克莱顿市莫纳什大学克莱顿校区信息技术学院的高级讲师。他还在莫纳什大学数据科学中心工作。他是美国宾夕法尼亚州匹兹堡市卡内基梅隆大学计算机学院的博士后研究助理。A. 豪普特曼他已经花了大部分的时间

视觉、声学和本地), 以在无约束或监控视频中进行自

资料寄存器Chang是2019年至2021年获得的ARC发现早期职业研究员奖(DE CRA)研究员。他在各种国际比赛中取得了最好的表现, 如特AVS, 特辛, 和特AVS AVS。