Finding and Navigating to Humans in Complex Environments for Assistive Tasks

Asfand Yaar[1][a], Antonino Furnari[1][b], Marco Rosano[1][c], Aki Härmä[2][d] and Giovanni Maria Farinella[1][e]

[1]*Department of Mathematics and Computer Science, University of Catania, Italy*

[2]*Department of Advanced Computing Sciences, Maastricht University, The Netherlands*

{*antonino.furnari, marco.rosano, giovanni.farinella*}*@unict.it*

Keywords: Assistive AI, Autonomous Exploration, Human-Robot Interaction, Embodied Navigation, Habitat Simulator.

Abstract: Finding and reaching humans in unseen environments is a major challenge for intelligent agents and social robots. Effective exploration and navigation strategies are necessary to locate the human performing various activities. In this paper, we propose a problem formulation in which the robot is required to locate and reach humans in unseen environments. To tackle this task, we design an approach that makes use of state-of-the-art components to allow the agent to explore the environment, identify the human's location on the map, and approach them while maintaining a safe distance. To include human models, we utilized Blender to modify the scenes of the Gibson dataset. We conducted experiments using the Habitat simulator, where the proposed approach achieves promising results. The success of our approach is measured by the distance and orientation difference between the robot and the human at the end of the episode. We will release the source code and 3D human models for researchers to benchmark their assistive systems.

1 INTRODUCTION

Autonomous robots able to navigate and interact with humans could be helpful in many assistive scenarios. Consider for instance a robot assisting an elderly in their home to carry out daily activities. The robot could provide instructions on how to successfully prepare a recipe, remind them to take the medicines at a given hour or recommend not to sit too much in front of the TV and go out for a walk once in a while. In order to achieve such a varied range of assistive tasks in the home, and in particular to initiate any form of visual or vocal interaction with the human, robots should be able to locate the human and reach them appropriately. For example, in a scenario where a human instructs a robot, *come to me* as shown in Fig 2.

The robot needs to explore the environment and locate the human while keeping track of its progress to avoid redundant searches. Once the robot has reached the area in which the human is located, it can calculate the human's position on the map and approach them from the right angle to initiate a conversation. This task requires complex exploration strategies, including a combination of implicit objectives such as exploration, efficient navigation, and interaction. While the ability to locate humans and navigate to them is a fundamental building block for assistive robotic applications, there is still a need for a more systematic investigation of the ability of current algorithms to tackle this task in different environments.

To fill this gap, in this paper, we focus on evaluating robot performance in locating and navigating to

Figure 1: 3D models of humans involved in different activities such as watching TV, eating, being on a call, and cooking used in our experiments.

[a] https://orcid.org/0009-0006-6329-3500
[b] https://orcid.org/0000-0001-6911-0302
[c] https://orcid.org/0009-0003-8680-1246
[d] https://orcid.org/0000-0002-2966-3305
[e] https://orcid.org/0000-0002-6034-0432

Yaar, A., Furnari, A., Rosano, M., Härmä, A. and Farinella, G.
Finding and Navigating to Humans in Complex Environments for Assistive Tasks.
DOI: 10.5220/0012271700003660

In *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2024) - Volume 4: VISAPP*, pages 245-252
ISBN: 978-989-758-679-8; ISSN: 2184-4321

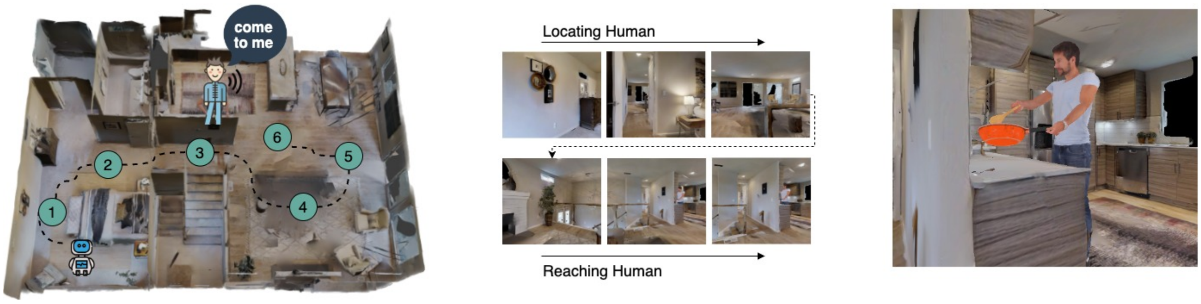

Figure 2: (Left) Robot responding to a human call by exploring the environment with the help of multiple global goals (1-6) to locate and reach the human. (center) Robot's observations upon reaching each global goal (1-6) during its exploration of the environment. (right) Robot's final observation upon successfully reaching the human at an appropriate angle, depending on the human's activity, to initiate a conversation.

the human engaged in various activities such as cooking, eating, talking, watching TV, etc. To complete this task, the robot explores the environment to locate and reach the human at a safe distance. We provide a problem formulation and design a set of baselines based on human detection and point-goal navigation to tackle the task. We validate the feasibility of the task and the effectiveness of the considered baselines using the Habitat simulator (Savva et al., 2019) and the Gibson (Xia et al., 2018) validation dataset which consists of five complex 3D environments. As this dataset lacks human models, we used Blender[1] to create four human models, in poses coherent with the execution of four different activities (eating, cooking, watching TV, and on a call) as illustrated in Fig. 1. Subsequently, we modified the Gibson environments in Blender to incorporate these human models at multiple locations, such as the kitchen, TV lounge, bedroom, and other relevant areas. Results show that considered baselines achieve promising performance in locating and reaching the human in an unseen environment. However, further research is still needed in this area, and we believe that our proposed approach can serve as a starting point for future works on how assistive robots can be used to provide support to users. The main contributions of this work are listed below:

- We propose a novel pipeline for efficiently locating and reaching humans in complex environments, with a focus on assistive tasks and human-robot interaction.
- Our approach utilizes global and local goal policies to generate objectives and precisely reach the human.
- We made modifications to the Gibson environments by integrating 3D human models into various locations, aligning their poses with different activities in areas such as the kitchen, TV lounge, and other relevant areas that were previously absent in the original dataset.
- We show that considered baselines achieve promising performance in locating and reaching human in complex 3D environments.

[1]https://www.blender.org/

2 RELATED WORK

Our work is related to previous research on embodied navigation and environment exploration. The embodied visual navigation problem involves an agent using visual sensing to navigate an environment avoiding obstacles to reach a given destination (Anderson et al., 2018a; Anderson et al., 2018b; Batra et al., 2020; Savva et al., 2019). Over the last decade, the field has made substantial progress due to the availability of large photorealistic 3D scene datasets (Chang et al., 2017; Ramakrishnan et al., 2021; Xia et al., 2018) and fast navigation simulators (Savva et al., 2019; Xia et al., 2018; Kolve et al., 2017). Current literature on embodied visual navigation can be divided into classic navigation, approaches based on reinforcement learning, and exploration.

Classic Navigation. Traditional navigation approaches involve building a map of the environment, localizing the agent in the map, and planning paths to guide the agent to desired locations. Mapping, localization, and path-planning have been extensively studied in this context (Hartley and Zisserman, 2003; Thrun, 2002; LaValle, 2006). However, most of this research relies on human-operated traversal of the environment and is classified as passive SLAM. Active SLAM, which focuses on automatically navigating a new environment to build spatial representations, has received less attention. We refer the reader to (Cadena et al., 2016) for a comprehensive review of active SLAM literature.

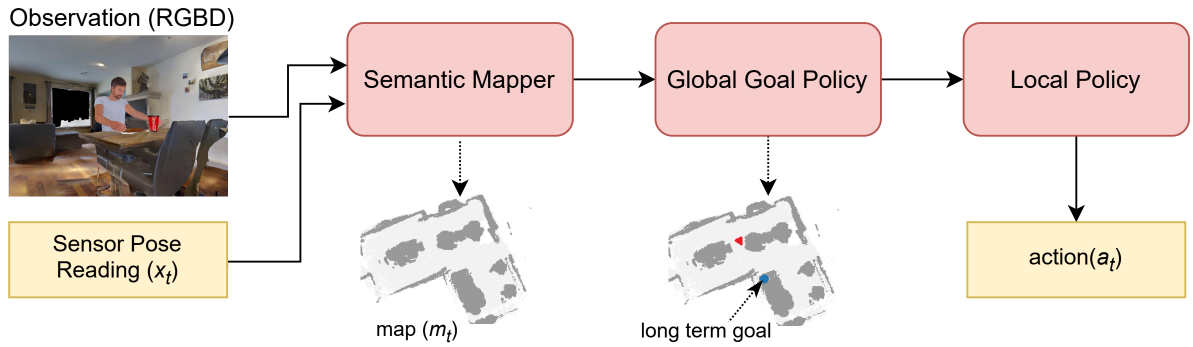

Figure 3: The Semantic Mapper leverages RGB-D and sensor pose reading x_t to construct a map of the environment (m_t). The Global Goal Policy utilizes this map to generate a long-term goal on the map. Finally, the Local Policy generates low-level actions a_t to guide the agent toward this long-term goal.

Reinforcement Learning. Different previous works have formulated navigation as a reinforcement learning problem (Zhu et al., 2017; Gupta et al., 2017; Mirowski et al., 2016; Savinov et al., 2018) in which the robot is an agent interacting with a simulated environment in order to learn how to navigate it. By training in several environments, the agents eventually learn how to extract semantic cues from the input images and generalize them to unseen spaces. Past works have investigated methods including feed-forward networks (Zhu et al., 2017), vanilla neural network memory (Mirowski et al., 2016), spatial memory and planning modules (Gupta et al., 2017), semi-parametric topological memory (Savinov et al., 2018), and imitation learning from an optimal expert (Gupta et al., 2017). In addition, learning-based approaches have been used to develop low-level collision avoidance policies (Dhiraj et al., 2017; Sadeghi and Levine, 2016). However, these approaches do not consider task context and only focus on moving toward open space. Other works (Zhang et al., 2017) use a differentiable map structure to mimic SLAM techniques.

Exploration. Navigation algorithms are generally shaped around two main objectives: point-goal navigation and environment exploration. The first class of methods aims to navigate in order to reach a given destination provided to the agent in the form of coordinates relative to the current location. Exploration approaches aim instead to navigate the unknown environment without an explicit target location in mind but with the goal to “uncover” all the available space, which can be useful for mapping the environment or searching for specific objects. Environment exploration as an Active Neural SLAM (ANS) to gather information for downstream tasks has been a popular topic in the past, with many works investigating it in the context of reinforcement learning (Schmidhuber, 1991; Stadie et al., 2015; Pathak et al., 2017; Fu et al., 2017; Lopes et al., 2012; Chentanez et al., 2004). These works design intrinsic reward functions that capture the novelty of states or state-action transitions, which are then used to optimize exploration policies using reinforcement learning.

Other related works have proposed alternative exploration methods, such as generating smooth movement paths for high-quality camera scans (Xu et al., 2017), information-theoretic exploration method using Gaussian Process regression (Bai et al., 2016), or assuming access to the ground-truth map at training time to learn an optimized trajectory that maximizes the accuracy of the SLAM-derived map (Kollar and Roy, 2008). Recently, (Chen et al., 2019) used human navigation trajectories to learn task-independent exploration through imitation learning. To improve exploration specifically in object goal navigation tasks, SemExp (Chaplot et al., 2020b) made use of a modular policy for semantic mapping and path planning that directly predicts which action the agent should take next and estimates the map on the fly.

Overall, there has been a growing interest in developing robots that can perform diverse tasks in a variety of environments (Lim et al., 2021). Exploration and navigation are critical components of such systems, and recent work has made significant progress in learning exploration policies and developing modular architectures for navigation. Our work is akin to these methods, but we investigate the navigation problem in an assisting care robot scenario by proposing a method that relies on Habitat simulator and comparing different baselines. The four activities we selected come from that use case.

3 PROBLEM DEFINITION

We aim to assess the performance of a robot in locating and navigating to a human involved in various activities. We perform our evaluation in an episode-based fashion, following works on navigation with reinforcement learning (Chaplot et al., 2020b; Chaplot et al., 2020a; Ramakrishnan et al., 2022). At the beginning of an episode, the agent is initialized at a random location in the environment and receives a visual observation o (an RGB-D image) and sensor position reading x_t (i.e. x and y coordinates of the agent and its orientation at time t). The agent then takes a navigation action a_t following a learned policy to achieve the goal of locating and navigating to the human. At each time step, the robot can choose among the following actions: *move_forward*, *turn_right*, *turn_left*, and *stop*.

To successfully complete the task, the *stop* action should be called when the agent is confident that the human has been reached. An episode ends when the agent calls the *stop* action or when it reaches the limit of 500 steps. Note that, since the human may not be visible from the initial location, the agent should first explore the environment, then navigate to the human when they are detected from the visual observation. This makes this task different from classic point goal navigation (Anderson et al., 2018a) or environment exploration works (Zhang et al., 2017; Chaplot et al., 2020a), effectively requiring a mix of both objectives. We consider two versions of this problem:

- **V1**: The first version considers an episode successful if the robot reaches the human at a safe distance ($1m$) at the end of the episode.
- **V2**: The second version considers an episode successful if the robot reaches the human at a safe distance ($1m$) and the difference in orientation between the robot and the human θ is below a given threshold.

Evaluations are performed by computing the Success weighted by Path Length (SPL) and the Success Rate (SR) for both versions.

4 PROPOSED METHOD

The proposed approach relies on three key components: a Semantic Mapper, a Global Goal Policy, and a Local Policy. Fig. 3 illustrates the proposed approach.

Semantic Mapper. The Semantic Mapper is responsible for creating an allocentric semantic map m_t of the world by aggregating semantic information obtained from individual RGB-D observations acquired from time *0* to t. This is done using a state-of-the-art semantic exploration method (Chaplot et al., 2020b), which creates a point cloud from depth observations. Each point in the point cloud is then classified as either a person or a background class using the semantic segmentation model. The point cloud is then projected into the top-down map space using differentiable geometric operations (Henriques and Vedaldi, 2018), resulting in the $3 \times M \times M$ semantic map m_t, with channels 1 and 2 representing obstacles and explored areas, and the last channel representing the person class. In our setup, we considered $M = 240$, while each element of m_t corresponds to a 25 cm^2 (5cm×5cm) cell in the physical world and indicates whether the location contains an obstacle, has been explored, or contains a person. The spatial map is initialized to all zeros at the beginning of an episode and refined during the navigation process.

Global Goal Policy. The Global Goal Policy network consists of 5 convolutional layers followed by 3 fully connected layers. It is responsible for determining the long-term goal in order to reach the human by using the current map m_t. If the human is not detected, the global goal policy aims to explore the environment and hence predicts a long-term goal using the map and the agent's current and previous positions. To reduce the exploration complexity, the long-term goal is predicted once every 25 steps as described in (Chaplot et al., 2020a). If the human is detected, the global goal policy selects a point close to the human as a long-term goal. It is worth mentioning that both versions of the task, i.e. V1 and V2, employ the same global goal policy, with the only distinction being the evaluation procedure when the robot reaches the human.

Local Policy. The Local Policy is used to navigate continuously to the long-term goal defined by the Global Goal Policy by calculating the shortest path from the current position to the target one using the Fast Marching Method (Sethian, 1999). The obstacle channel from the semantic map is used to determine the optimal path while avoiding obstacles. The local policy then uses deterministic actions to navigate the agent along this shortest path. At each time step, the map is updated and the path to the long-term goal is re-computed.

5 EXPERIMENTS AND RESULTS

Gibson (Xia et al., 2018) and Matterport3D (MP3D) datasets (Chang et al., 2017) were employed in the Habitat simulator (Savva et al., 2019) for training purposes. These datasets contain 3D reconstructions of

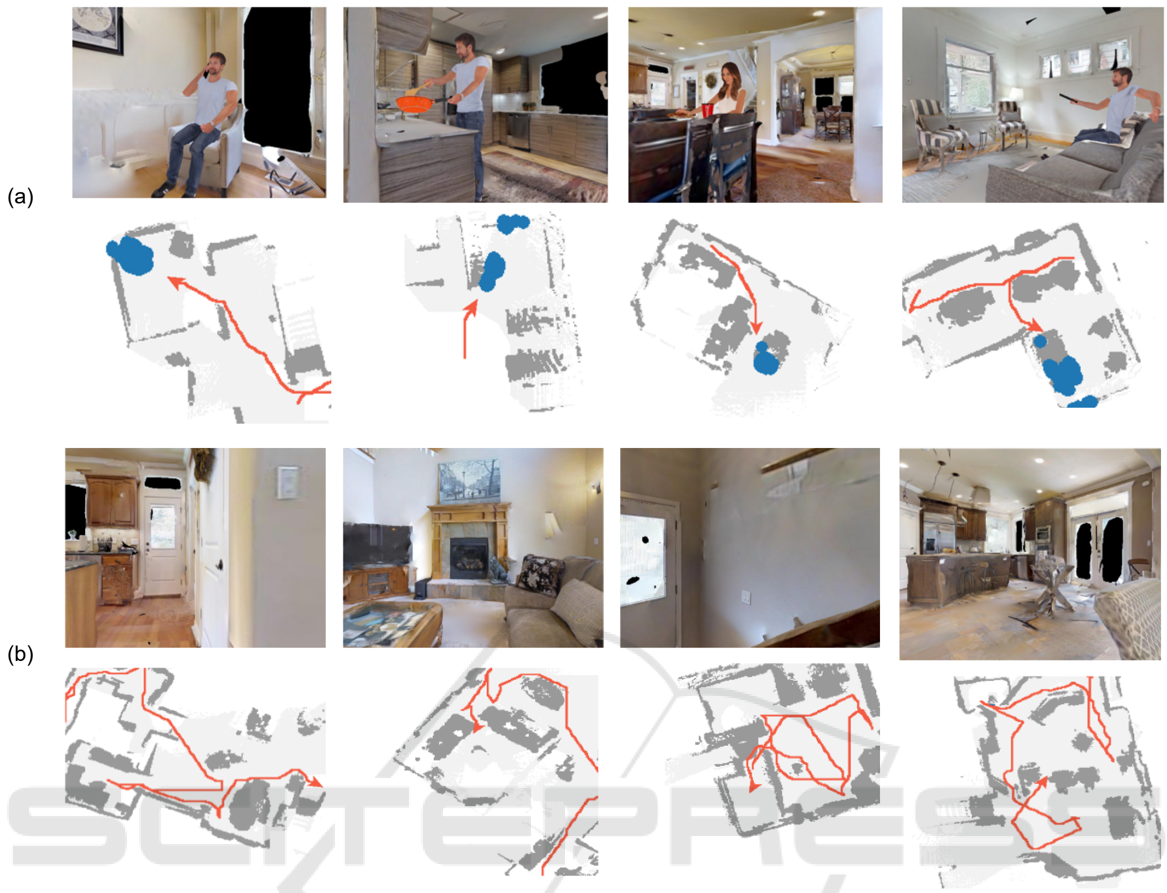

Figure 4: (a) Final observation of robot while successfully locating and reaching a human engaged in various activities, on call, cooking, eating, and watching TV respectively. (b) Some examples where the robot was unable to locate the human within 500 steps.

real-world environments. The training set includes a total of 86 scenes, consisting of 25 scenes from the Gibson tiny set and 61 scenes from the MP3D dataset. Since the human models are not included in these datasets, we employed Blender to create multiple human models. These models were posed to align with the execution of various activities, such as eating, cooking, watching TV, and being on a call (see Fig 1), etc. We then edited the Gibson environments using Blender to integrate these human models at multiple locations, including the kitchen, TV lounge, bedroom, and other relevant areas. Note that, due to the computation-intensive nature of the manual integration process, our current implementation includes four human models.

The observation space consists of RGBD images with a size of *4×640×480*, while the action space includes four possible actions: *move_forward* (0.25 cm), *turn_right* (10 degrees), *turn_left* (10 degrees), and *stop*. The success threshold is set to *1m*. For person detection and segmentation, we use a Mask-RCNN semantic segmentation model (He et al., 2017) with a ResNet50 (He et al., 2016) backbone, pretrained on MS-COCO (Lin et al., 2014). We use Success weighted by Path Length (SPL) and Ratio of successful episodes (SR) to measure the efficiency of locating and reaching the human. We evaluate the proposed approach on the 20 modified environments of the Gibson dataset that were not seen during the training of the different components of our approach. This allowed us to examine how well the learned policies generalize to previously unseen environments.

We run 2000 evaluation episodes, with each scene containing 100 episodes. We consider the two variants of the task: V1 aims to reach the human from any angle, whereas in V2 episode success depends on the orientation difference between the robot and the human at the end of the episode. Table 1 provides quantitative results for V1, where the SPL and SR values for each activity in different environments are presented separately. The performance of the robot is observed to vary across environments, with larger

Table 1: SPL and SR for the V1 task on the Gibson validation dataset for each activity.

	Gibson Environments									
	Collierville		Corozal		Darden		Markleeville		Wiconisco	
Activity	SPL	SR	SPL	SR	SPL	SR	SPL	SR	SPL	SR
1. Eating	0.80	0.99	0.62	0.99	0.68	0.96	0.63	0.94	0.57	0.96
2. Cooking	0.59	0.95	0.41	0.89	0.54	0.88	0.69	0.97	0.41	0.85
3. Watching TV	0.70	0.99	0.31	0.77	0.40	0.94	0.55	0.90	0.69	0.99
4. On a call	0.74	0.99	0.67	0.96	0.56	0.98	0.53	0.93	0.40	0.84

Table 2: Average SPL and SR on the Gibson validation dataset.

Task	SPL (↑)	SR (↑)
V1 (any angle)	0.57	0.93
V2 ($\theta \leq 60°$)	0.25	0.44
V2 ($\theta \leq 30°$)	0.14	0.26

environments posing greater challenges for the robot. Notably, the robot has a limit of 500 steps to locate the human, and as a result, there are instances where the robot fails to locate the human within the specified time frame. Fig 4b provides visual examples of such instances of failure. Our proposed approach achieves a 93% SR and a 57% SPL under $V1$. However, in $V2(\theta = 30°)$, our approach only achieved a 26% SR and a 14% SPL. This suggests that V2 of the task is much more challenging and more research is still needed.

To illustrate the effect of evaluating models with different orientation thresholds, we plot the SPL and SR for multiple variants of $V2$ approach with varying orientation thresholds $(0° - 180°)$ in Fig 5. The plot shows that the SPL and SR increase as we raise the orientation tolerance threshold. Interestingly, even when setting a threshold of 90°, results are not satisfactory, with a SPL of about 0.35 and a SR of about 0.6. This suggests that the considered task is challenging and there is a lot of space for improvement. Fig. 4 finally shows some success (a) and failure (b) qualitative navigation episodes along with the final visual observation of the agent. As can be seen, the approach can reach the human from the right angle in some of the cases. Table 2 presents the overall results of the proposed approach.

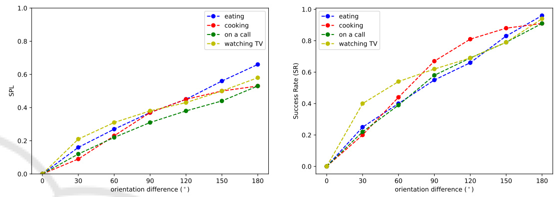

Figure 5: SPL (left) and Success Rate (right) of the navigation tasks, considering different thresholds on the angle between the robot and the human. The episode is considered successful if the robot reaches the human at a safe distance and with a robot-human orientation difference lower than the given threshold.

6 CONCLUSION

In this paper, we proposed a navigation problem formulation as a first essential step towards building autonomous task-oriented assistive robots for home use cases. In the considered setup, an agent has to find humans and reach them at a safe distance, in order to provide assistance. The experiments performed on the Gibson dataset comprising 3D human models show that this is a promising direction for the development of a flexible framework for assistive robots. In future research, we plan to extend the proposed framework with more intelligent task-oriented robot behaviors sensitive to the situational and social conventions of natural home environments.

ACKNOWLEDGEMENT

This research has been supported by the PhilHumans project[2] supported by Marie Skłodowska-Curie Innovative Training Networks - European Industrial Doctorates under grant agreement No. 812882.

REFERENCES

Anderson, P., Chang, A., Chaplot, D. S., Dosovitskiy, A., Gupta, S., Koltun, V., Kosecka, J., Malik, J., Mottaghi, R., Savva, M., et al. (2018a). On evaluation of embodied navigation agents. *arXiv preprint arXiv:1807.06757*.

[2]https://www.philhumans.eu/

Anderson, P., Wu, Q., Teney, D., Bruce, J., Johnson, M., Sünderhauf, N., Reid, I., Gould, S., and Van Den Hengel, A. (2018b). Vision-and-language navigation: Interpreting visually-grounded navigation instructions in real environments. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3674–3683.

Bai, S., Wang, J., Chen, F., and Englot, B. (2016). Information-theoretic exploration with bayesian optimization. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1816–1822. IEEE.

Batra, D., Gokaslan, A., Kembhavi, A., Maksymets, O., Mottaghi, R., Savva, M., Toshev, A., and Wijmans, E. (2020). Objectnav revisited: On evaluation of embodied agents navigating to objects. *arXiv preprint arXiv:2006.13171*.

Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., and Leonard, J. J. (2016). Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Transactions on robotics*, 32(6):1309–1332.

Chang, A., Dai, A., Funkhouser, T., Halber, M., Niessner, M., Savva, M., Song, S., Zeng, A., and Zhang, Y. (2017). Matterport3d: Learning from rgb-d data in indoor environments. *arXiv preprint arXiv:1709.06158*.

Chaplot, D. S., Gandhi, D., Gupta, S., Gupta, A., and Salakhutdinov, R. (2020a). Learning to explore using active neural slam. *arXiv preprint arXiv:2004.05155*.

Chaplot, D. S., Gandhi, D. P., Gupta, A., and Salakhutdinov, R. R. (2020b). Object goal navigation using goal-oriented semantic exploration. *Advances in Neural Information Processing Systems*, 33:4247–4258.

Chen, T., Gupta, S., and Gupta, A. (2019). Learning exploration policies for navigation. *arXiv preprint arXiv:1903.01959*.

Chentanez, N., Barto, A., and Singh, S. (2004). Intrinsically motivated reinforcement learning. *Advances in neural information processing systems*, 17.

Dhiraj, G., Pinto, L., and Gupta, A. (2017). Learning to fly by crashing. In *IROS*.

Fu, J., Co-Reyes, J., and Levine, S. (2017). Ex2: Exploration with exemplar models for deep reinforcement learning. *Advances in neural information processing systems*, 30.

Gupta, S., Davidson, J., Levine, S., Sukthankar, R., and Malik, J. (2017). Cognitive mapping and planning for visual navigation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2616–2625.

Hartley, R. and Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge university press.

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778.

Henriques, J. F. and Vedaldi, A. (2018). Mapnet: An allocentric spatial memory for mapping environments. In *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8476–8484.

Kollar, T. and Roy, N. (2008). Trajectory optimization using reinforcement learning for map exploration. *The International Journal of Robotics Research*, 27(2):175–196.

Kolve, E., Mottaghi, R., Han, W., VanderBilt, E., Weihs, L., Herrasti, A., Deitke, M., Ehsani, K., Gordon, D., Zhu, Y., et al. (2017). Ai2-thor: An interactive 3d environment for visual ai. *arXiv preprint arXiv:1712.05474*.

LaValle, S. M. (2006). *Planning algorithms*. Cambridge university press.

Lim, V., Rooksby, M., and Cross, E. S. (2021). Social robots on a global stage: establishing a role for culture during human–robot interaction. *International Journal of Social Robotics*, 13(6):1307–1333.

Lin, T., Maire, M., Belongie, S. J., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L. (2014). Microsoft COCO: common objects in context. In Fleet, D. J., Pajdla, T., Schiele, B., and Tuytelaars, T., editors, *Computer Vision - ECCV 2014 - 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V*, volume 8693 of *Lecture Notes in Computer Science*, pages 740–755. Springer.

Lopes, M., Lang, T., Toussaint, M., and Oudeyer, P.-Y. (2012). Exploration in model-based reinforcement learning by empirically estimating learning progress. *Advances in neural information processing systems*, 25.

Mirowski, P., Pascanu, R., Viola, F., Soyer, H., Ballard, A. J., Banino, A., Denil, M., Goroshin, R., Sifre, L., Kavukcuoglu, K., et al. (2016). Learning to navigate in complex environments. *arXiv preprint arXiv:1611.03673*.

Pathak, D., Agrawal, P., Efros, A. A., and Darrell, T. (2017). Curiosity-driven exploration by self-supervised prediction. In *International conference on machine learning*, pages 2778–2787. PMLR.

Ramakrishnan, S. K., Chaplot, D. S., Al-Halah, Z., Malik, J., and Grauman, K. (2022). Poni: Potential functions for objectgoal navigation with interaction-free learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18890–18900.

Ramakrishnan, S. K., Gokaslan, A., Wijmans, E., Maksymets, O., Clegg, A., Turner, J., Undersander, E., Galuba, W., Westbury, A., Chang, A. X., et al. (2021). Habitat-matterport 3d dataset (hm3d): 1000 large-scale 3d environments for embodied ai. *arXiv preprint arXiv:2109.08238*.

Sadeghi, F. and Levine, S. (2016). Cad2rl: Real single-image flight without a single real image. *arXiv preprint arXiv:1611.04201*.

Savinov, N., Dosovitskiy, A., and Koltun, V. (2018). Semi-parametric topological memory for navigation. *arXiv preprint arXiv:1803.00653*.

Savva, M., Kadian, A., Maksymets, O., Zhao, Y., Wijmans, E., Jain, B., Straub, J., Liu, J., Koltun, V., Malik, J., et al. (2019). Habitat: A platform for embodied ai research. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9339–9347.

Schmidhuber, J. (1991). A possibility for implementing curiosity and boredom in model-building neural controllers. In *Proc. of the international conference on simulation of adaptive behavior: From animals to animats*, pages 222–227.

Sethian, J. A. (1999). Fast marching methods. *SIAM review*, 41(2):199–235.

Stadie, B. C., Levine, S., and Abbeel, P. (2015). Incentivizing exploration in reinforcement learning with deep predictive models. *arXiv preprint arXiv:1507.00814*.

Thrun, S. (2002). Probabilistic robotics. *Communications of the ACM*, 45(3):52–57.

Xia, F., Zamir, A. R., He, Z., Sax, A., Malik, J., and Savarese, S. (2018). Gibson env: Real-world perception for embodied agents. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 9068–9079.

Xu, K., Zheng, L., Yan, Z., Yan, G., Zhang, E., Niessner, M., Deussen, O., Cohen-Or, D., and Huang, H. (2017). Autonomous reconstruction of unknown indoor scenes guided by time-varying tensor fields. *ACM Transactions on Graphics (TOG)*, 36(6):1–15.

Zhang, J., Tai, L., Liu, M., Boedecker, J., and Burgard, W. (2017). Neural slam: Learning to explore with external memory. *arXiv preprint arXiv:1706.09520*.

Zhu, Y., Mottaghi, R., Kolve, E., Lim, J. J., Gupta, A., Fei-Fei, L., and Farhadi, A. (2017). Target-driven visual navigation in indoor scenes using deep reinforcement learning. In *2017 IEEE international conference on robotics and automation (ICRA)*, pages 3357–3364. IEEE.