



Universidad Politécnica
de Madrid

**Escuela Técnica Superior de
Ingenieros Informáticos**



Máster Universitario en Inteligencia Artificial

Trabajo Fin de Máster

**Navegación Reactiva Aplicada a Agentes
Físicos en Entornos Domésticos usando
Habitat Sim**

Autor(a): Luna Jiménez Fernández
Tutor(a): Martín Molina Gómez

Madrid, Septiembre - 2021

Este Trabajo Fin de Máster se ha depositado en la ETSI Informáticos de la Universidad Politécnica de Madrid para su defensa.

Trabajo Fin de Máster
Máster Universitario en Inteligencia Artificial

Título: Navegación Reactiva Aplicada a Agentes Físicos en Entornos Domésticos
usando Habitat Sim

Septiembre - 2021

Autor(a): Luna Jiménez Fernández
Tutor(a): Martín Molina Gómez
Computer Vision and Aerial Robotics (CVAR)
ETSI Informáticos
Universidad Politécnica de Madrid

Resumen

«Aquí va el resumen del TFM. Extensión máxima 2 páginas.»

Abstract

«Abstract of the Master Project. Maximum length: 2 pages.»

Tabla de contenidos

1. Introducción	1
1.1. Introducción	1
1.2. Motivación	2
1.3. Estructura	2
2. Descripción del problema	5
2.1. Definición del problema	5
2.2. Antecedentes	5
2.3. Objetivos	7
3. Revisión de técnicas	9
3.1. <i>Deep Learning</i>	9
3.2. Aprendizaje por refuerzo	9
3.2.1. Algoritmos de aprendizaje por refuerzo clásicos	9
3.2.2. Algoritmos de aprendizaje por refuerzo profundos	9
3.3. Algoritmos de navegación automática	9
4. Simulador: <i>Habitat Sim</i> y <i>Habitat Lab</i>	11
4.1. <i>Habitat Sim</i> y <i>Habitat Lab</i>	11
4.1.1. <i>Habitat Sim</i>	11
4.1.2. <i>Habitat Lab</i>	11
4.1.3. Principales conceptos de <i>Habitat Lab</i>	11
4.1.3.1. Entornos	11
4.1.3.2. Tareas	11
4.1.3.3. Conjuntos de datos	11
4.1.3.4. Episodios	12
4.1.3.5. Sensores	12
4.1.3.6. Ficheros de configuración	12
4.1.3.7. Entrenadores	12
4.1.3.8. Agentes	12
4.1.3.9. <i>Benchmarks</i>	12
4.2. Instalación del simulador	12
4.2.1. Requisitos y versiones	12
4.2.2. Proceso de instalación	12
5. Diseño del agente	13
5.1. Caracterización del conocimiento	13
5.1.1. Descripción del problema	13

5.1.2. Estado	13
5.1.3. Acciones	13
5.1.4. Recompensas	13
5.2. Arquitectura del agente	13
5.2.1. Propuesta 1: Red convolucional (CNN)	13
5.2.2. Propuesta 2: Red híbrida (CNN + MLP)	13
5.3. Actuación del agente	13
5.4. Entrenamiento del agente	13
5.4.1. <i>Experience Replay</i>	13
5.4.2. Memorización de experiencias	13
5.4.3. Aprendizaje a partir de las experiencias	13
5.5. Otros agentes propuestos	13
5.5.1. Agentes heurísticos	14
5.5.2. Agente PPO	14
6. Experimentación	15
6.1. Experimentos realizados y parametros utilizados	15
6.2. Resultados durante el entrenamiento	15
6.3. Resultados durante la evaluación	15
7. Conclusiones	17
7.1. Conclusiones	17
7.2. Trabajo futuro	17
7.3. Agradecimientos	17
Bibliografía	17
Anexos	20

Capítulo 1

Introducción

En este capítulo se realizará una breve introducción a los contenidos que serán expuestos posteriormente a lo largo de la memoria. Tras ésta presentación, se expondrá la motivación que ha propiciado el desarrollo de este trabajo. Finalmente, se describirá la estructura seguida por la memoria.

1.1. Introducción

La **navegación autónoma** de robots en entornos desconocidos y complejos es un problema de gran interés en la actualidad para el que se ha propuesto una amplia gama de soluciones, buscando que éstas sean a la vez eficientes durante su entrenamiento y capaces de navegar entornos de forma exitosa. Una de las familias de algoritmos más relevantes para este propósito son los **algoritmos de aprendizaje por refuerzo**, capaz de aprender de forma autónoma a navegar entornos desconocidos a partir de experiencia previa, con gran éxito.

Además, la **simulación virtual** tanto de estos robots como de otros problemas es un campo en crecimiento, especialmente durante la pandemia del CoVID-19, al verse limitadas las capacidades de experimentación en entornos físicos. Por tanto el objetivo de este trabajo es aunar el **desarrollo de un algoritmo híbrido eficiente** para la navegación en entornos complejos (como el interior de un domicilio) con el **estudio y uso de *Habitat Sim***, un simulador novedoso para el entrenamiento y evaluación de agentes robóticos físicos.

Esta memoria comienza con una revisión de las principales técnicas usadas actualmente tanto en *Deep Learning* como en aprendizaje por refuerzo y en navegación autónoma de robots (centrándose en los algoritmos de *Artificial Potential Field (APF)*). Tras esto, se realiza un estudio en detalle del simulador *Habitat Sim* y su principal librería / API para *Python*, *Habitat Lab*; centrándose en los principales componentes de la librería, su funcionamiento y su uso. Posteriormente, se describe el diseño e implementación que se ha realizado en el trabajo, haciendo hincapié en la representación del conocimiento (problema a resolver y definiciones de estado, acción y recompensa), las arquitecturas del agente propuestas y su funcionamiento tanto durante el entrenamiento como la actuación.

Se procede después a la explicación de los experimentos realizados (tanto los parámetros usados como los experimentos a realizar), analizando el rendimiento de las

variantes propuestas del agente durante el entrenamiento y en una evaluación posterior, analizando éstos resultados y comparándolos con otros agentes usados como *benchmarks*. Finalmente, se interpretarán estos resultados, extrayendo unas conclusiones y ofreciendo futuras líneas de trabajo a partir del conocimiento adquirido.

1.2. Motivación

Este trabajo se puede entender como una continuación del trabajo realizado por C. Sampedro *et al.* en 2018 [1], en el que se desarrolla con buenos resultados un sistema de navegación autónomo para drones aéreos usando aprendizaje por refuerzo profundo con campos de potenciales artificiales y láseres para percibir el entorno. Una de las metas de este trabajo es estudiar si la implementación de un algoritmo de características similares pero aplicado a robots terrestres usando cámaras de profundidad en interiores (domicilios, fábricas...) sería igualmente efectivo.

Además, la situación de pandemia actual ha dejado en evidencia la necesidad del uso de simuladores, especialmente para algoritmos que necesiten un entrenamiento largo y que puedan necesitar equipamiento especializado para ello (como robots, drones, instalaciones especializadas...). Por eso, otra de las principales metas del trabajo es el estudio de la herramienta *Habitat Sim* [2] [3], viendo su viabilidad de cara a futuros trabajos.

Para acabar, otra razón no despreciable para la elección de esta temática de trabajo es el propio interés de la alumna por el campo del aprendizaje por refuerzo profundo. Ya se realizó un trabajo previo estudiando la aplicación de estas técnicas a juegos reales como Tetris [4], y este trabajo sirve para ampliar más el conocimiento y aplicarlo a tecnologías modernas y a otros campos de interés.

1.3. Estructura

Esta memoria está dividida en un total de 7 capítulos, que serán descritos brevemente a continuación.

- **Capítulo 1:** En este capítulo se introduce el trabajo desarrollado, la motivación que ha llevado a éste y la estructura general de la memoria.
- **Capítulo 2:** En este capítulo se describe en profundidad el problema a resolver, presentando los antecedentes previos al trabajo realizado y detallando los objetivos que se esperan cumplir.
- **Capítulo 3:** En este capítulo se realiza una revisión de las principales técnicas en los campos relacionados con el trabajo: *deep learning* y redes neuronales convolucionales, aprendizaje por refuerzo (estudiando tanto las técnicas clásicas como las técnicas de aprendizaje por refuerzo profundo) y algunos de los principales algoritmos de navegación automática.
- **Capítulo 4:** En este capítulo se presentan tanto *Habitat Sim* como *Habitat Lab*, las principales herramientas usadas durante el desarrollo del trabajo. Tras esto, se exponen los principales componentes de Habitat Lab, explicando su funcionamiento y uso. Finalmente, se habla sobre la instalación y las dependencias necesarias del simulador.

- **Capítulo 5:** En este capítulo se detalla el diseño del agente de navegación reactiva propuesto. Se describe tanto la representación del conocimiento (estado, acciones y recompensas) como la arquitectura, el método de actuación y el entrenamiento llevado a cabo por el agente. Finalmente, se realiza una breve explicación del funcionamiento y la arquitectura del resto de agentes usados como *benchmarks* y comparativas ofrecidos por *Habitat Lab*.
- **Capítulo 6:** En este capítulo se detalla la experimentación realizada, indicando los parámetros utilizados. Además, se presentan los resultados y el rendimiento obtenido por los agentes tanto durante el entrenamiento como durante la evaluación posterior.
- **Capítulo 7:** Finalmente, en este capítulo se presentan las conclusiones alcanzadas tras el desarrollo del trabajo, proponiendo posibles líneas de trabajo futuro para continuarlo.

Además, al final de la memoria se incluye una bibliografía en la que se encuentra la lista de fuentes y referencias usadas a lo largo de ésta.

Capítulo 2

Descripción del problema

En este capítulo se planteará en detalle el problema a resolver. Tras esto, se comentarán soluciones previas propuestas al problema, y se describirán los objetivos que se esperan alcanzar con el desarrollo del trabajo.

2.1. Definición del problema

El problema a resolver es el diseño de un agente físico (un robot) capaz de navegar desde una posición inicial hasta una posición final (conocidas sus coordenadas) en un entorno de interior del que no se tiene conocimiento previo (como puede ser el interior de una casa) de forma eficiente. Este problema se encuadra en el campo de la navegación automática, concretamente en el de la **navegación a meta** (*Point Goal Navigation*) **sin exploración previa** del entorno [5].

Para lograr esto, es necesario entrenar al agente en un entorno controlado para que aprenda a navegar de forma autónoma en interiores desconocidos, usando alguna técnica de aprendizaje por refuerzo. Además, se busca que el agente sea capaz de navegar **sin exploración previa** (como se ha mencionado anteriormente), buscando un sistema reactivo frente a uno basado en exploración-navegación.

El agente físico cuenta con las siguientes características a tener en cuenta:

- **Movimiento:** El agente es terrestre (se desplaza con ruedas sobre el suelo), pudiendo moverse hacia adelante y rotar sobre si mismo. El agente no es capaz de realizar movimientos ortogonales.
- **Sensores:** El agente cuenta con dos sensores para recibir información del entorno: una **cámara de profundidad** para observar el espacio frente al robot y un **GPS**, indicando la distancia y ángulo hasta la meta.

2.2. Antecedentes

La navegación autónoma de robots en interiores es un campo de gran interés tanto para la investigación como para la industria, existiendo gran cantidad de grupos dedicados a la propuesta y desarrollo de agentes capaces de resolver estos problemas de forma eficaz.

Algunos ejemplos de este interés *Habitat Challenge*, un desafío anual organizado por *Facebook AI Research* (el grupo de investigación de inteligencia artificial de Facebook) parte de la *Conference on Computer Vision and Pattern Recognition (CVPR)* en el que se buscan los mejores algoritmos para resolver problemas de **navegación autónoma a metas** [6] (aunque también incluye problemas de **navegación autónoma a objetos** [7]) aplicados a agentes físicos en entornos de interiores. Algunas de las mejores propuestas de estos últimos años han sido:

- **Devendra Singh Chaplot et al. (2019)** [8]: Se propone un nuevo algoritmo, *Active Neural SLAM*, que combina las capacidades de planificadores de ruta clásicos como SLAM con la capacidad del aprendizaje por refuerzo profundo para generar políticas de acciones locales y globales. Esta propuesta obtuvo el primer puesto del *Habitat Challenge 2019*.
- **Santhosh K. Ramakrishnan et al. (2020)** [9]: Se propone un sistema de navegación entrenado con aprendizaje por refuerzo que, a partir de sus observaciones a través de una cámara RGB, es capaz de inferir la posición de los objetos más allá de su ángulo de visión para mejorar su rendimiento a la hora de generar un mapa de su entorno. Esta propuesta alcanzó el primer puesto del *Habitat Challenge 2020*.
- **Samyak Datta et al. (2020)** [10]: Se propone un sistema de navegación entrenado con aprendizaje por refuerzo (*PPO*) que tiene en cuenta el ruido existente en entornos reales (problemas durante la actuación, desviaciones entre la posición real y estimada...) haciendo esfuerzo para corregirlo. Esta propuesta alcanzó el segundo puesto del *Habitat Challenge 2020*.
- **Ruslan Partsey (2021)** [11]: Se desarrolla un agente que usa técnicas de odometría (usar datos de sensores de movimiento para estimar la posición real) visual con aprendizaje por refuerzo para conseguir una navegación eficiente en entornos con ruido. Esta propuesta alcanzó el primer puesto del *Habitat Challenge 2021*.

Si bien todas las propuestas anteriores utilizan técnicas de aprendizaje por refuerzo, la gran mayoría de éstas utilizan enfoques basados en el mapeado y navegación de los entornos, frente a una propuesta puramente reactiva (sin mapa) como la de éste trabajo.

El antecedente más directo al trabajo descrito en esta memoria es la propuesta realizada por **Carlos Sampedro et al. (2018)** [1], siendo éste trabajo una continuación directa. El agente propuesto utiliza un método basado en campos de potenciales, entrenado con aprendizaje por refuerzo profundo para navegar un dron aéreo a través de entornos de interior, usando un conjunto de láseres (para percibir el entorno a su alrededor) y la posición relativa del propio dron respecto a la meta, con buenos resultados. Esta arquitectura se puede observar en la Figura 2.1.

Ahora bien, existen diferencias destacables entre la propuesta original y el trabajo descrito en esta memoria:

- **Movimientos del agente:** La propuesta original utiliza como agente a un dron multirrotor, capaz de navegar por el aire (aunque se mantiene a una altura constante) y de realizar movimiento omnidireccional. En cambio, el trabajo desarrollado utiliza un robot terrestre (susceptible a obstáculos en el suelo) incapaz

Descripción del problema

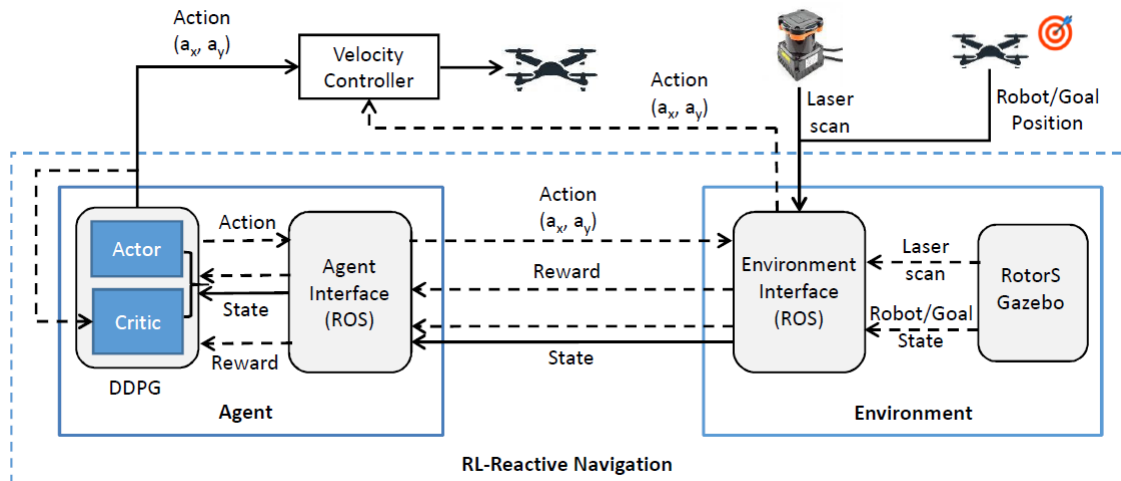


Figura 2.1: Arquitectura propuesta para el sistema de navegación reactivo [1].

de movimiento omnidireccional, siendo necesario el giro del agente para poder esquivar obstáculos y desplazarse.

- **Sensores del agente:** Mientras que la propuesta original utiliza un conjunto de láseres para percibir su entorno, el trabajo desarrollado propone un agente con una única cámara de profundidad frontal. Si bien el conjunto de láseres es más caro que una cámara, también ofrece un ángulo de visión mayor de los obstáculos del entorno.
- **Complejidad del entorno:** La propuesta original entrena al agente en entornos de interior simples (contando con espacios simples con obstáculos dispersos), frente a los entornos usados por el trabajo desarrollado (siendo recreaciones del interior de domicilios, incluyendo topografías más complejas con cuartos y una mayor cantidad de obstáculos y ruido).

2.3. Objetivos

El principal objetivo de este trabajo es el estudio y aplicación de técnicas de aprendizaje por refuerzo profundo y navegación autónoma basada en campos de potenciales para el desarrollo de un agente capaz de navegar entornos de interior, evaluando su viabilidad y eficacia.

Para cumplir este objetivo, es necesario a su vez cumplir una serie de objetivos parciales:

- Revisión de bibliografía para comprender plenamente las técnicas a usar durante el desarrollo.
- Búsqueda y evaluación de librerías y herramientas disponibles para el desarrollo del agente (incluyendo simuladores, entornos de trabajo...)
- Caracterización, formalización e implementación del agente y de posibles variaciones propuestas dentro del entorno elegido, para poder ser evaluado posteriormente.

- Realización de experimentos para estudiar el comportamiento del agente durante el entrenamiento y posteriormente al enfrentarse a problemas reales.
- Estudio y análisis de los resultados, realizando comparación con *benchmarks* para extraer observaciones y conclusiones que permitan valorar la viabilidad y eficacia del agente propuesto.

Este trabajo además aborda un segundo objetivo, el estudio y uso del simulador *Habitat Sim*, con el fin de evaluar su utilidad de cara a posteriores trabajos. Para esto, se plantean los siguientes objetivos parciales:

- Revisión y estudio de documentación oficial y ejemplos ofrecidos por el simulador.
- Desarrollo del agente descrito previamente en el marco del simulador, usando las herramientas ofrecidas.
- Creación de documentación sobre el uso adecuado del simulador para facilitar trabajos posteriores.
- Evaluación de la idoneidad del simulador para la resolución de problemas de navegación autónoma.

Capítulo 3

Revisión de técnicas

3.1. *Deep Learning*

[SOBRE DEEP LEARNING, REDES NEURONALES PROFUNDAS Y CNNs]

3.2. Aprendizaje por refuerzo

[APRENDIZAJE POR REFUERZO, Q LEARNING (y quizás otros clásicos?), DEEP Q LEARNING (y mejoras), AGENTE / CRITICO, PPO...]

3.2.1. Algoritmos de aprendizaje por refuerzo clásicos

3.2.2. Algoritmos de aprendizaje por refuerzo profundos

3.3. Algoritmos de navegación automática

[PROBABLEMENTE CENTRARSE MAS EN ALGORITMOS DE ATRACCION REPULSION]

Capítulo 4

Simulador: *Habitat Sim* y *Habitat Lab*

4.1. *Habitat Sim* y *Habitat Lab*

[UNA INTRODUCCION MAS BREVE AL SIMULADOR COMO TAL]

4.1.1. *Habitat Sim*

[HABLA POR SEPARADO DE HABITAT 1 Y HABITAT 2]

4.1.2. *Habitat Lab*

[HABLA DE HABITAT LAB Y HABITAT BASELINES]

4.1.3. Principales conceptos de *Habitat Lab*

[DE CADA CONCEPTO, PROBABLEMENTE EXPLICAR QUE ES, EXPLICAR QUE OFRECE POR DEFECTO HABITAT / BASELINES]

[NO SE SI SERÁ NECESARIO HABLAR DEL PROPIO AGENTE AQUÍ O EN UN ANEXO]

[EL ORDEN ES TENTATIVO]

4.1.3.1. Entornos

[ENV, RENV, NAVRENV Y LOS METODOS QUE HAY QUE IMPLEMENTAR]

4.1.3.2. Tareas

[PROBABLEMENTE MENCIONAR LAS PRINCIPALES TAREAS OFRECIDAS POR ENCIMA]

4.1.3.3. Conjuntos de datos

[FOTOS DE CADA CONJUNTO DE DATOS] [MENCIONAR LOS 4 DATASETS DISPONIBLES, Y COMO USARLOS (ESTRUCTURA)]

4.1.3.4. Episodios

4.1.3.5. Sensores

4.1.3.6. Ficheros de configuración

[PROBABLEMENTE MENCIONAR SECCIONES CLAVES Y ELEMENTOS CLAVES QUE SIGNIFICAN]

4.1.3.7. Entrenadores

[AQUI UN PSEUDOCODIGO DEL METODO DE ENTRENAR QUEDARÍA DE LUJO]

4.1.3.8. Agentes

4.1.3.9. Benchmarks

4.2. Instalación del simulador

4.2.1. Requisitos y versiones

[INDICAR LAS VERSIONES USADAS DE TODO, Y ESPECIFICAR QUE NO SE GARANTIZA QUE FUNCIONE CON TODO]

4.2.2. Proceso de instalación

[INDICAR TAMBIEN CUDA]

Capítulo 5

Diseño del agente

[SE HABLAN DE TODOS LOS AGENTES, PERO NOS CENTRAMOS EN EL PROPIO (que para eso esta)]

5.1. Caracterización del conocimiento

5.1.1. Descripción del problema

5.1.2. Estado

5.1.3. Acciones

5.1.4. Recompensas

5.2. Arquitectura del agente

5.2.1. Propuesta 1: Red convolucional (CNN)

5.2.2. Propuesta 2: Red híbrida (CNN + MLP)

5.3. Actuación del agente

[exploracion explotacion]

5.4. Entrenamiento del agente

5.4.1. *Experience Replay*

5.4.2. Memorización de experiencias

5.4.3. Aprendizaje a partir de las experiencias

5.5. Otros agentes propuestos

[AQUI DENTRO SE INDICAN LOS AGENTES PROPUESTOS Y BREVEMENTE SU ARQUITECTURA]

5.5.1. Agentes heurísticos

5.5.2. Agente PPO

Capítulo 6

Experimentación

[PONEMOS EL ANALISIS EN UN CAPITULO APARTE?]

6.1. Experimentos realizados y parametros utilizados

[EXPERIMENTOS A REALIZAR, PARAMETROS A USAR, ORDENADOR USADO, ETC]
[PARA REPRODUCIBILIDAD VAMOS]

6.2. Resultados durante el entrenamiento

[MEDIDAS TIPO TIEMPO PARA ENTRENAR, RECOMPENSA MEDIA POR EPOCH, ACCIONES POR EPOCH, TASA DE EXITO...]

6.3. Resultados durante la evaluación

[TASA DE EXITO, CUAL ES MEJOR, ETC]

Capítulo 7

Conclusiones

7.1. Conclusiones

7.2. Trabajo futuro

7.3. Agradecimientos

[YA QUE NO LO TIENEN ANTES...]

Bibliografía

- [1] C. Sampedro, H. Bavle, A. Rodriguez-Ramos, P. De La Puente y P. Campoy, "Laser-Based Reactive Navigation for Multirotor Aerial Robots using Deep Reinforcement Learning," *IEEE International Conference on Intelligent Robots and Systems*, págs. 1024-1031, 2018, ISSN: 21530866. DOI: 10.1109/IROS.2018.8593706.
- [2] M. Savva, A. Kadian, O. Maksymets y col., "Habitat: A Platform for Embodied AI Research," en *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.
- [3] A. Szot, A. Clegg, E. Undersander y col., "Habitat 2.0: Training Home Assistants to Rearrange their Habitat," *arXiv preprint arXiv:2106.14405*, 2021.
- [4] L. Jimenez, "Aplicación de Deep Reinforcement Learning a un juego real - Tetris," *Universidad de Castilla-La Mancha*, pág. 109, 2020. dirección: <https://github.com/MoonDollLuna/dqlearning-tetris>.
- [5] P. Anderson, A. X. Chang, D. S. Chaplot y col., "On Evaluation of Embodied Navigation Agents," *CoRR*, vol. abs/1807.06757, 2018. arXiv: 1807.06757. dirección: <http://arxiv.org/abs/1807.06757>.
- [6] Abhishek Kadian*, Joanne Truong*, A. Gokaslan y col., "Sim2Real Predictivity: Does Evaluation in Simulation Predict Real-World Performance?," 4, vol. 5, 2020, págs. 6670-6677.
- [7] D. Batra, A. Gokaslan, A. Kembhavi y col., "ObjectNav Revisited: On Evaluation of Embodied Agents Navigating to Objects," en *arXiv:2006.13171*, 2020.
- [8] D. S. Chaplot, D. Gandhi, S. Gupta, A. Gupta y R. Salakhutdinov, "Learning to Explore using Active Neural SLAM," *CoRR*, vol. abs/2004.05155, 2020. arXiv: 2004.05155. dirección: <https://arxiv.org/abs/2004.05155>.
- [9] S. K. Ramakrishnan, Z. Al-Halah y K. Grauman, "Occupancy Anticipation for Efficient Exploration and Navigation," *CoRR*, vol. abs/2008.09285, 2020. arXiv: 2008.09285. dirección: <https://arxiv.org/abs/2008.09285>.
- [10] S. Datta, O. Maksymets, J. Hoffman, S. Lee, D. Batra y D. Parikh, "Integrating Egocentric Localization for More Realistic Point-Goal Navigation Agents," *CoRR*, vol. abs/2009.03231, 2020. arXiv: 2009.03231. dirección: <https://arxiv.org/abs/2009.03231>.
- [11] R. Partsey, "Robust Visual Odometry for Realistic PointGoal Navigation," *Ukrainian Catholic University*, pág. 87, 2021. dirección: <https://er.ucu.edu.ua/handle/1/2703>.

Apéndice A

[NO TENGO MUY CLARO COMO SE ESCRIBIAN LOS APENDICES NO TE VOY A ENGAÑAR] [EN LATEX ME REFIERO]

[APENDICES PROBABLES:] [MANUAL DE USUARIO, COMO GENERO GRAFICAS, CONTENIDOS DEL CD]