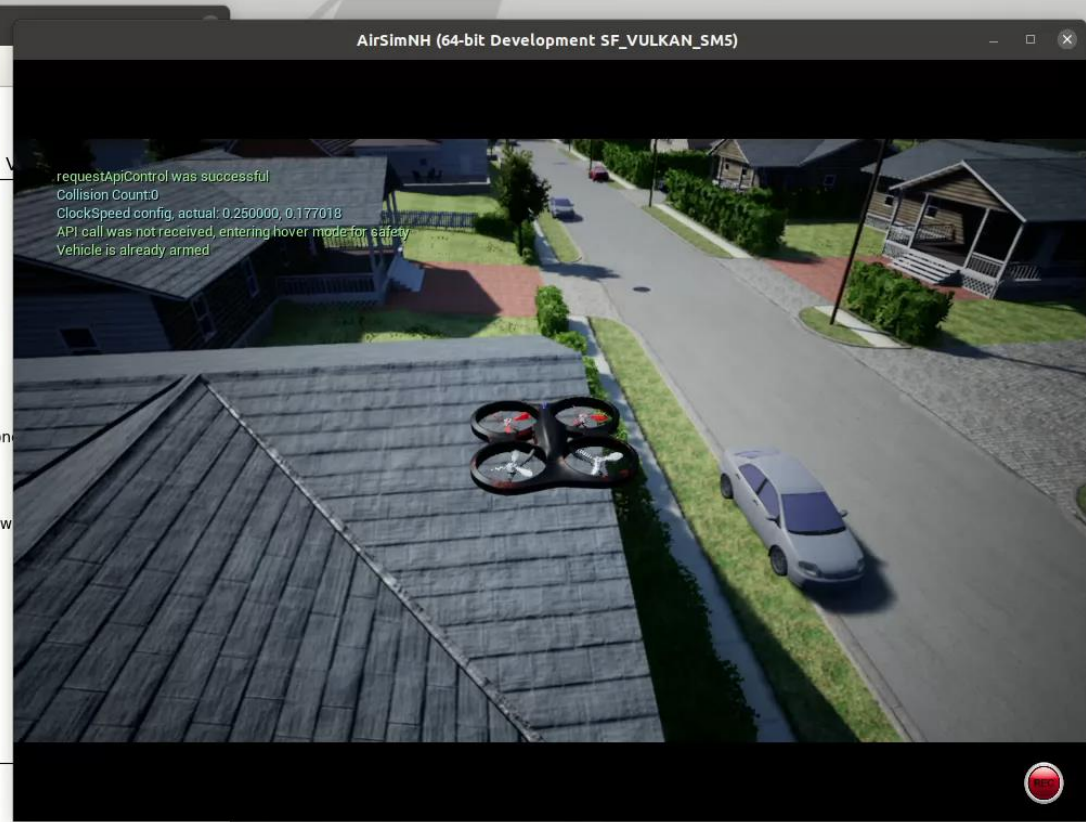
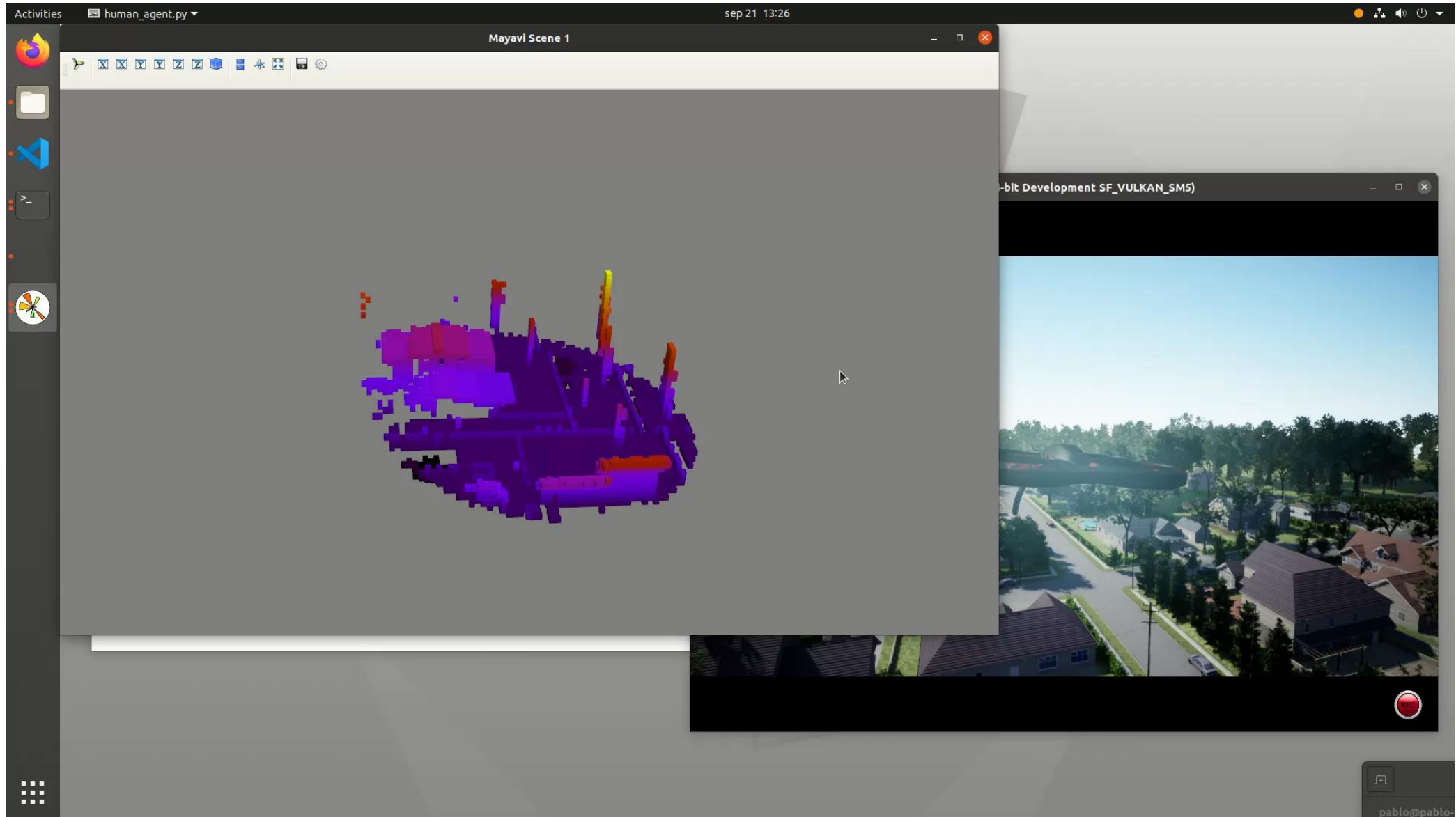


# Trabajo de Fin de Grado

Aprendizaje por refuerzo para la navegación autónoma de drones para  
el mapeo tridimensional de entornos

Pablo Cañadas Miquel





# OBJETIVOS DEL TRABAJO

- Reunir información
- Prueba de concepto

# PRIMERA PARTE

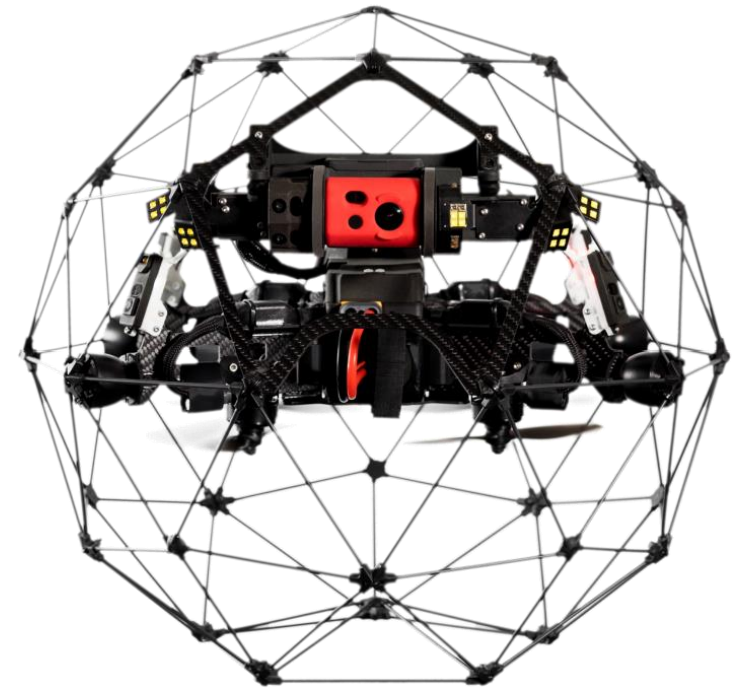
## INVESTIGACIÓN

# Mapeo tridimensional de entornos



[google.es/maps](https://google.es/maps)

Flyability, Elios 2



# SLAM

Simultaneous Localization and Mapping

Localización y Mapeo Simultáneos



# TECNOLOGÍAS



VS



Sensor IMU



Chip GPS



UAV/Dron





# NAVEGACIÓN AUTÓNOMA DE DRONES

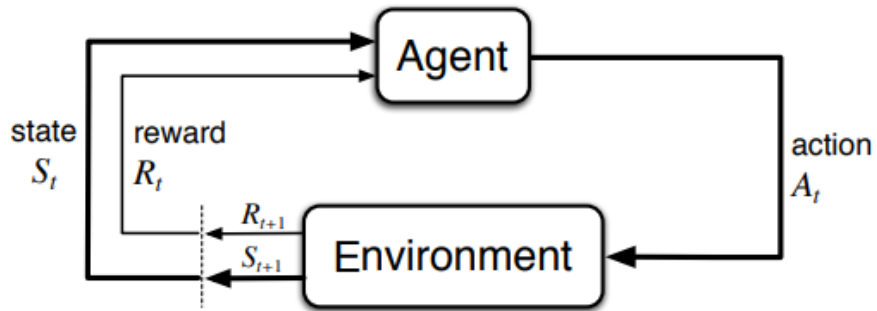
Navegación

Exploración

Búsqueda de caminos

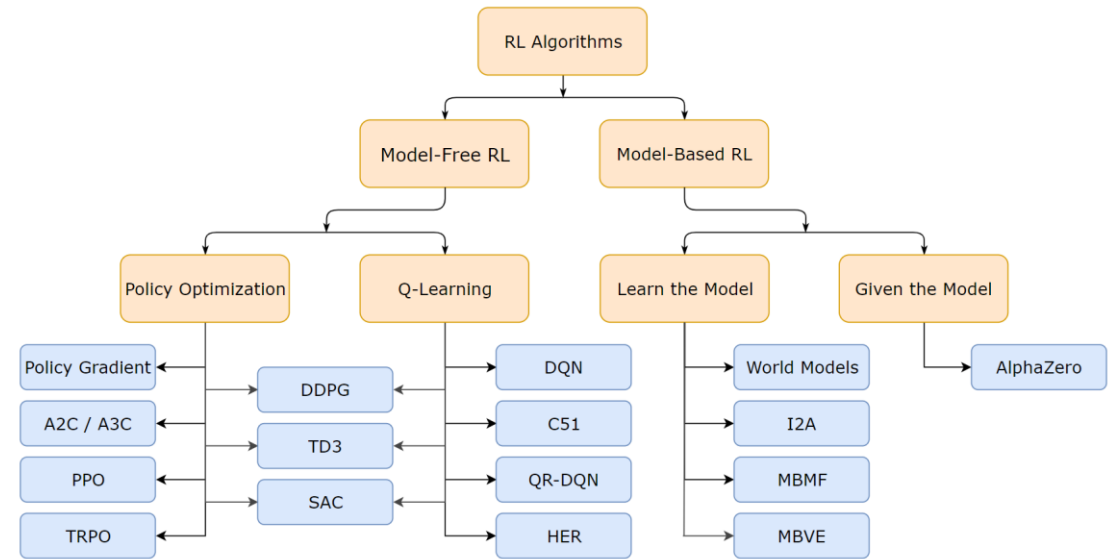
¿Aprendizaje por Refuerzo?

# APRENDIZAJE POR REFUERZO



Esquema general definido por MDP

## Taxonomía Aprendizaje por Refuerzo



# SEGUNDA PARTE

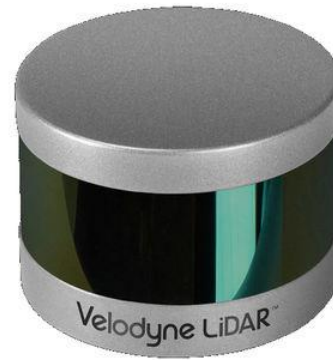
DISEÑO Y DESARROLLO

# VEHÍCULO



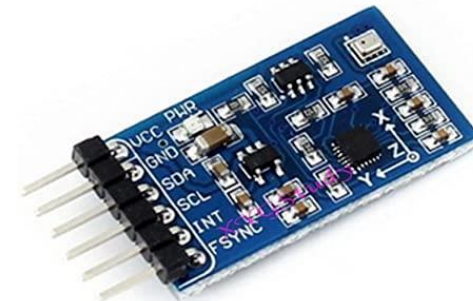
Dron

+



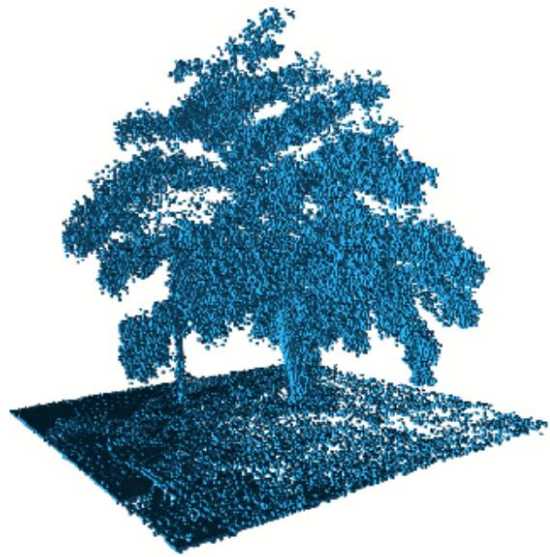
LiDAR

+



GNSS/IMU

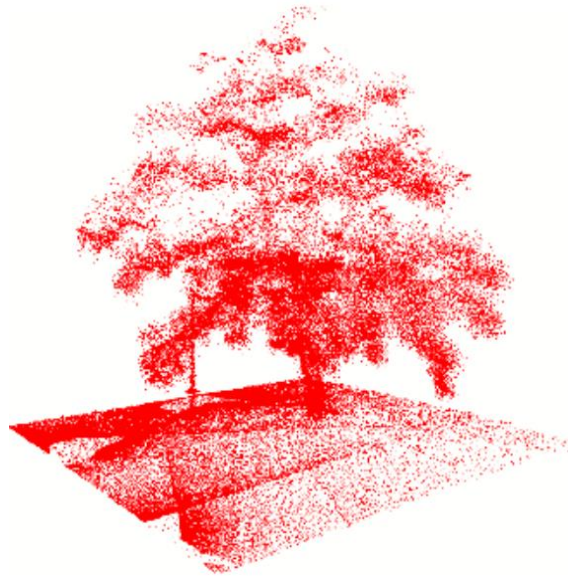
# HERRAMIENTA DE MAPEO



OctoMap (vóxels)

Eficiencia computacional

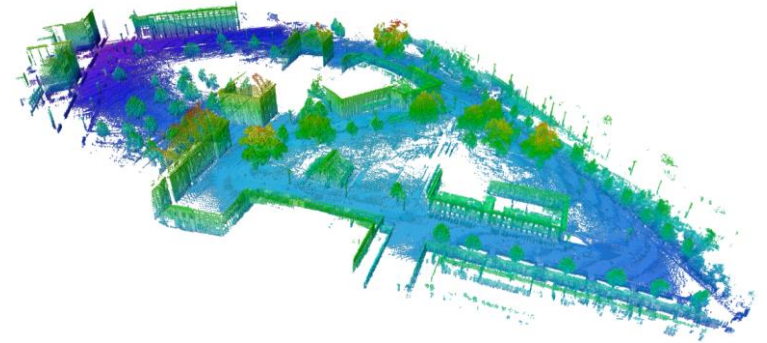
+



Nube de puntos

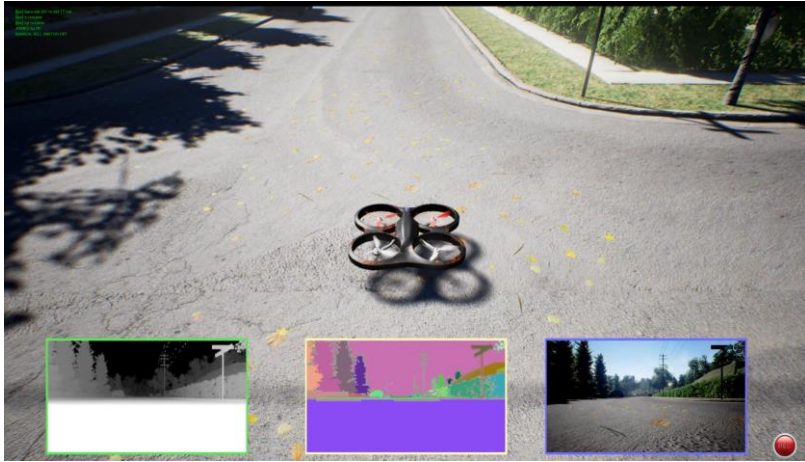
Alto detalle

Mapas indefinidamente grandes



# ENTORNO DE RL

SIMULADOR | AirSim



Percepción



Función de recompensa

$$r_t^+ = \max(4, (voxels_t - voxels_{t-1}) \cdot C); C \in (0, 1)$$

Condición de reinicio

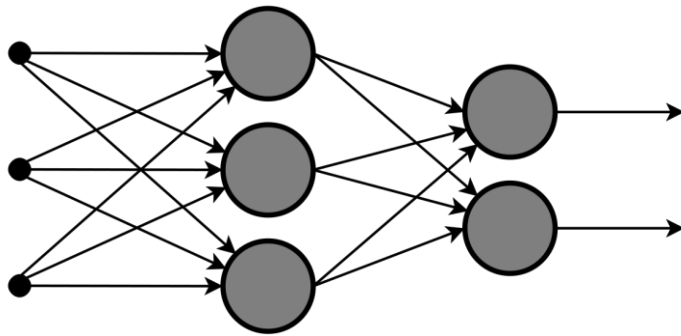


ACCIONES



# AGENTE AUTÓNOMO

DEEP REINFORCEMENT LEARNING



Variantes

- 'Feed-forward'
- Recurrente

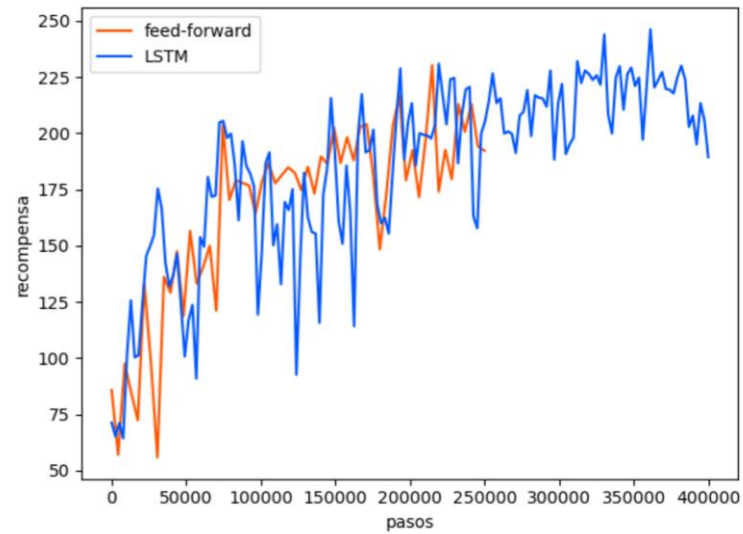
+

ALGORITMO 'ACTOR-CRITIC'

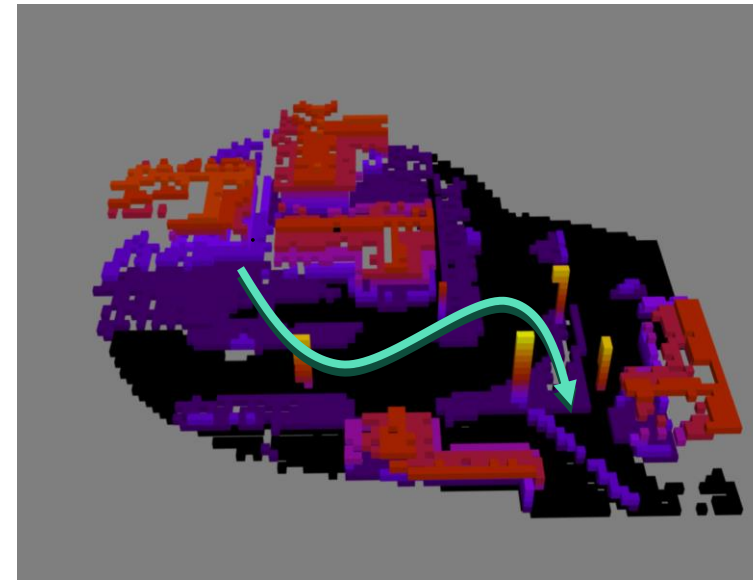
PROXIMAL POLICY OPTIMIZATION  
(PPO)

# AGENTE AUTÓNOMO: RESULTADOS

Recompensa media durante el entrenamiento:



(Mejores instancias)



Malos resultados

# CONCLUSIONES