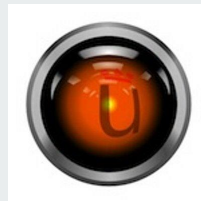




Conducción autónoma de un robot con visión mediante aprendizaje por refuerzo.

Ignacio Arranz Águeda • Octubre 2020

Tutor: José María Cañas
Cotutor: Eduardo Perdices



Universidad
Rey Juan Carlos



Índice

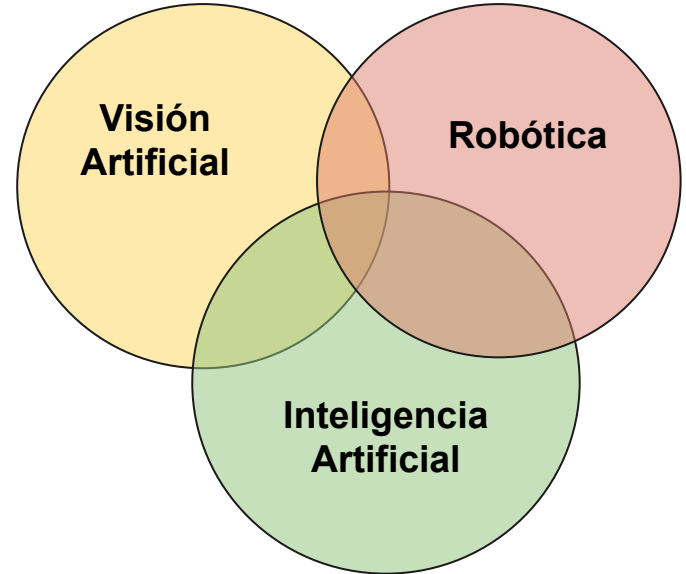
1. Introducción
2. Objetivos
3. Infraestructura
4. Aprendizaje por refuerzo
de un controlador visual
5. Validación experimental
6. Conclusiones

1. Introducción

Introducción

Motivación

- Creciente interés en los campos:
 - Inteligencia artificial.
 - Robótica.
 - Visión artificial.



Introducción

Robótica

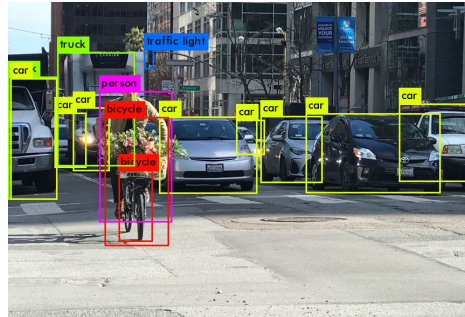
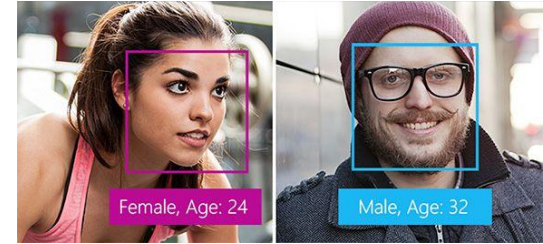
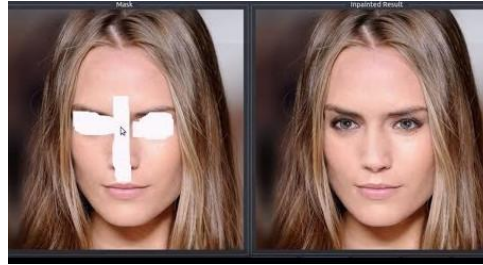
- Constante crecimiento.
- Actualización de las cadenas de montaje.
- Sector automovilístico demanda más robótica a nivel mundial.



Introducción

Visión Artificial

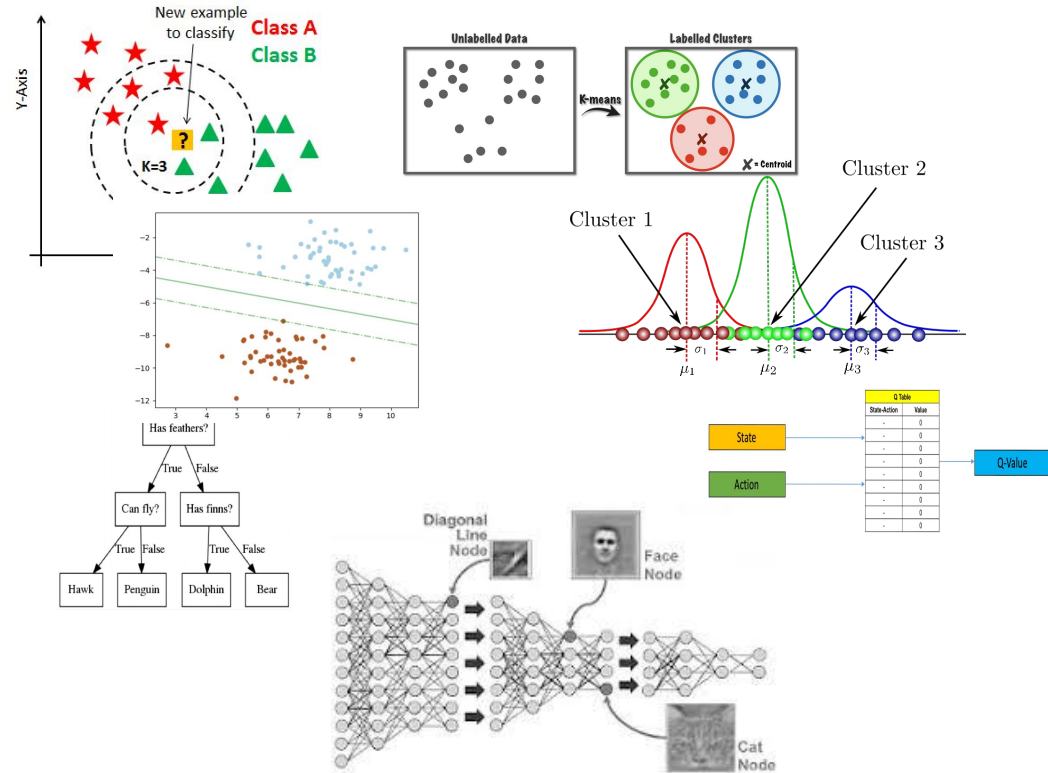
- Reconstrucción.
- Reconocimiento.
- Detección.
- Segmentación.



Introducción

Inteligencia artificial

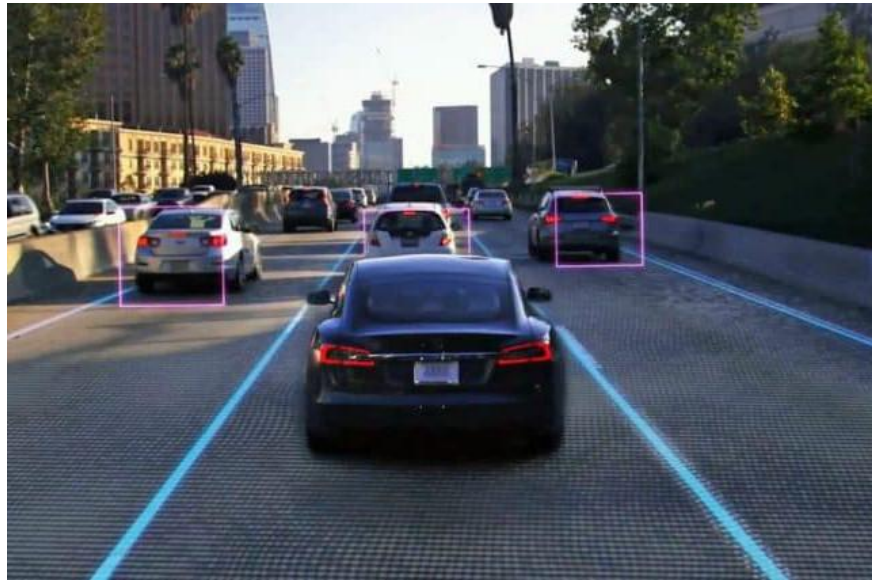
- Supervisado
- No Supervisado
- Aprendizaje por Refuerzo
- Aprendizaje Profundo



Introducción

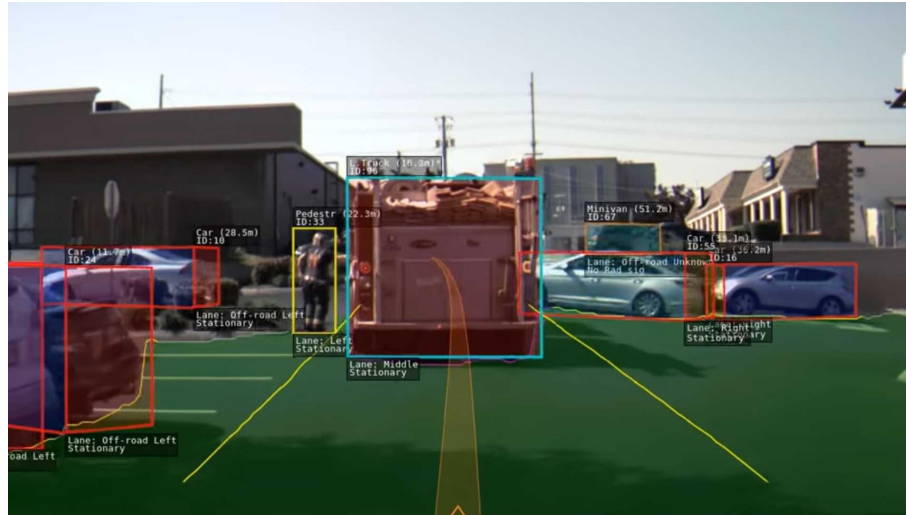
Conducción Autónoma

- Seguimiento del carril.
- Detección de vehículos.
- Detección de señales.
- Navegación.



Introducción

Ejemplo de aplicación



2. Objetivos

2. Objetivos

Son:

- Programar un entorno de aprendizaje por refuerzo con visión y robots.
- Entrenar un controlador visual para conducción autónoma siguiendo una línea.
- Validación experimental y análisis de las posibilidades del aprendizaje por refuerzo.

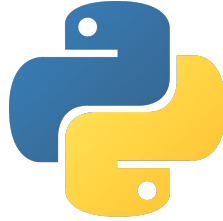
3. Infraestructura

3. Infraestructura

Herramientas y librerías:



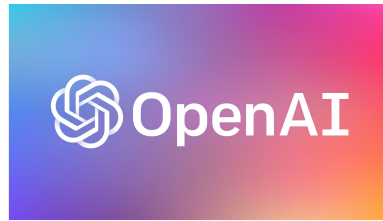
ROS



matplotlib



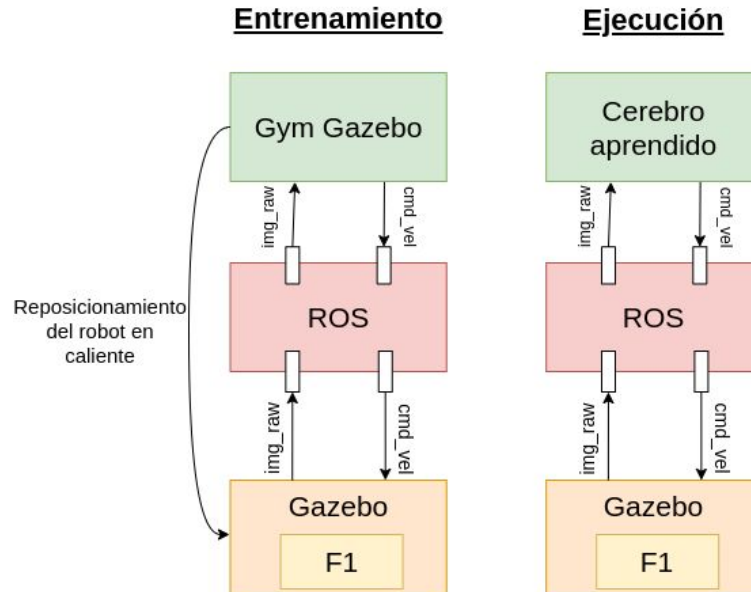
NumPy



4. Aprendizaje por refuerzo de un controlador visual

4. Aprendizaje por refuerzo de un controlador visual

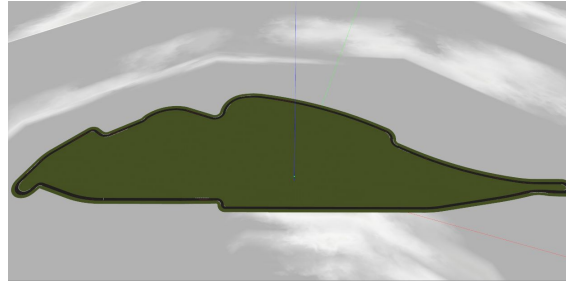
Diseño:



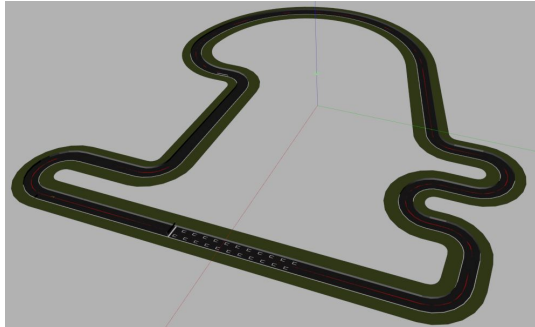
4. Aprendizaje por refuerzo de un controlador visual

Escenarios y modelos:

Montreal

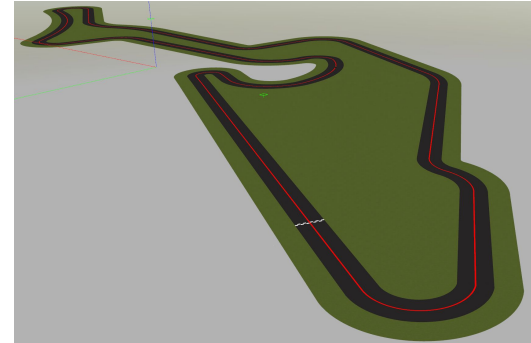


Circuito Simple



Modelo de Fórmula-1

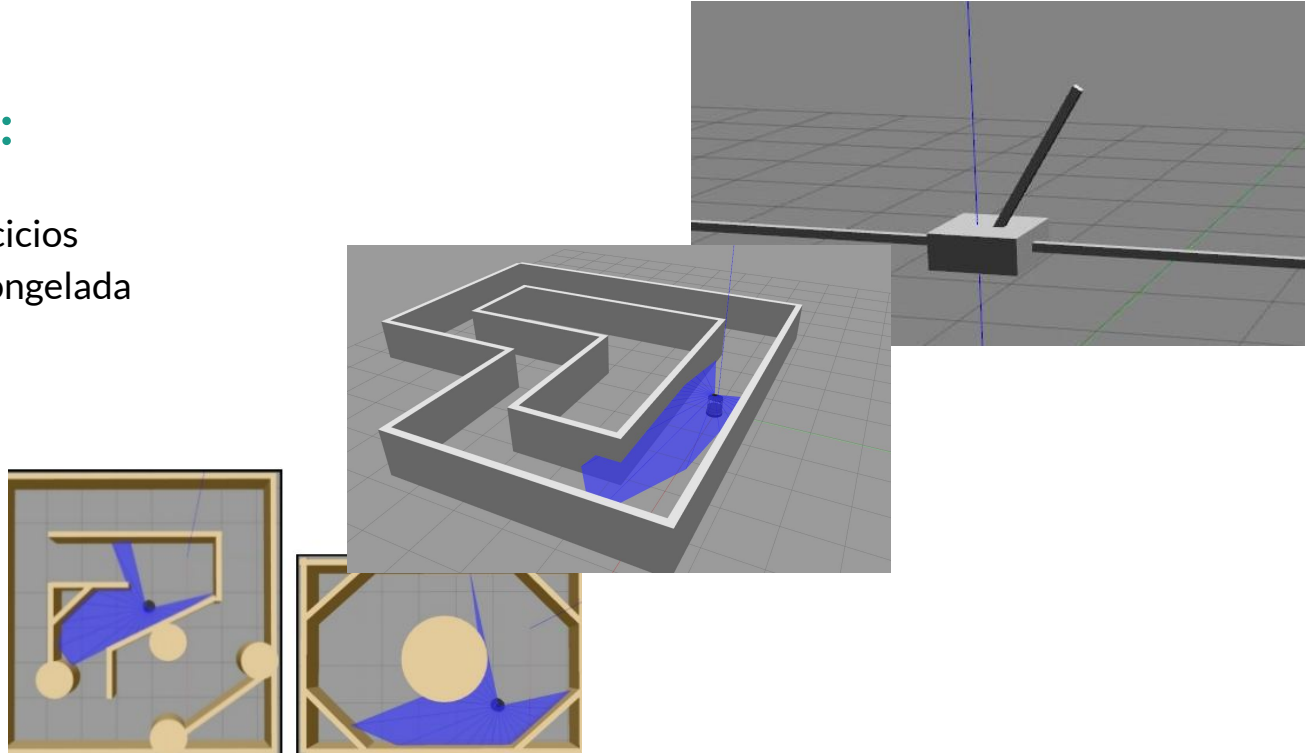
Nürburgring



4. Infraestructura

Gym-Gazebo:

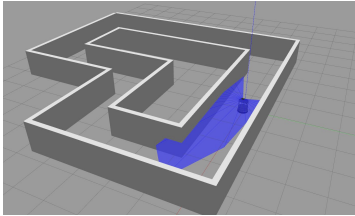
- Distintos ejercicios
- Repositorio congelada



4. Infraestructura

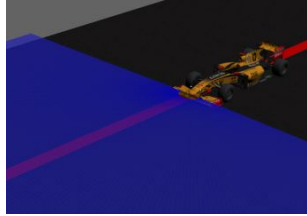
Fases:

Fase 1



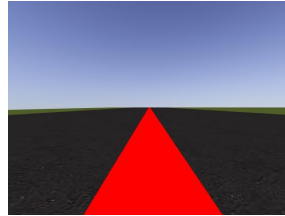
Seguir el carril con el Turtlebot

Fase 2



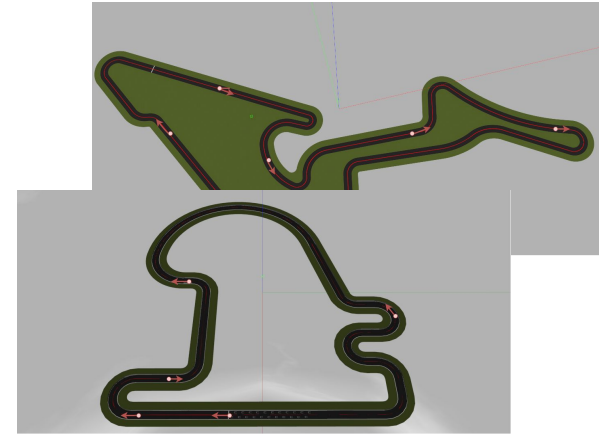
Nuevo ejercicio con el Fórmula-1 y el láser

Fase 3



Cambio de sensor. Uso de la cámara

Fase 4



Reinicios aleatorios

4. Aprendizaje por refuerzo de un controlador visual

Q-Learning:

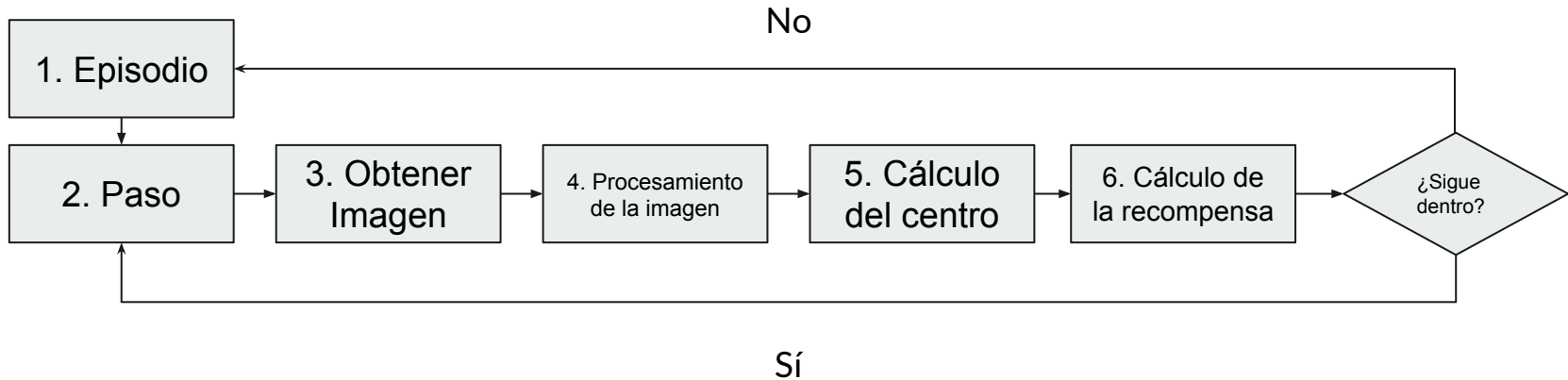
$$Q(s, a) \leftarrow (1 - \alpha)Q_{s,a} + \alpha(r + \gamma \cdot \max_{a'} Q_{s',a'})$$

1. Tabla-Q vacía
2. Se obtiene de un paso los valores:
 - a. Estado
 - b. Acción
 - c. Recompensa
 - d. Estado siguiente
3. Se actualiza la ecuación de Bellman
4. Se comprueban las condiciones de convergencia.

$$Q = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 0 & 0 & 0 & 0 & 80 & 0 \\ 0 & 0 & 0 & 64 & 0 & 100 \\ 0 & 0 & 0 & 64 & 0 & 0 \\ 0 & 80 & 51 & 0 & 80 & 0 \\ 64 & 0 & 0 & 64 & 0 & 100 \\ 0 & 80 & 0 & 0 & 80 & 100 \end{bmatrix} \end{matrix}$$

4. Aprendizaje por refuerzo de un controlador visual

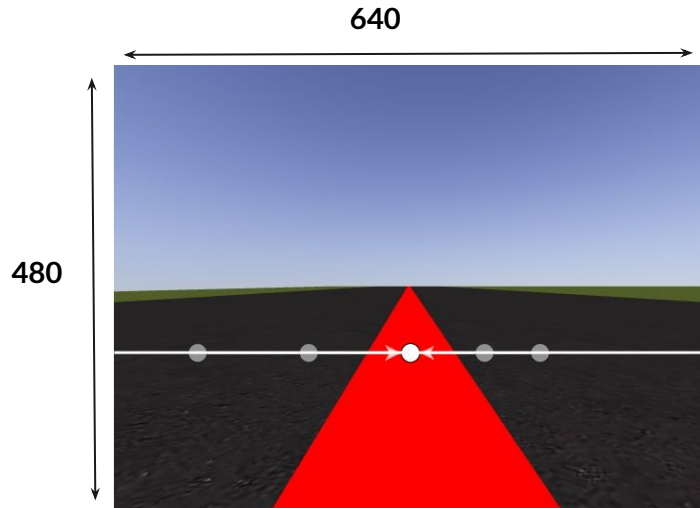
Flujo del programa de entrenamiento:



4. Aprendizaje por refuerzo de un controlador visual

Número de percepciones:

- 1, 2 y 3 puntos.



Número de acciones:

- Simple, medio y difícil

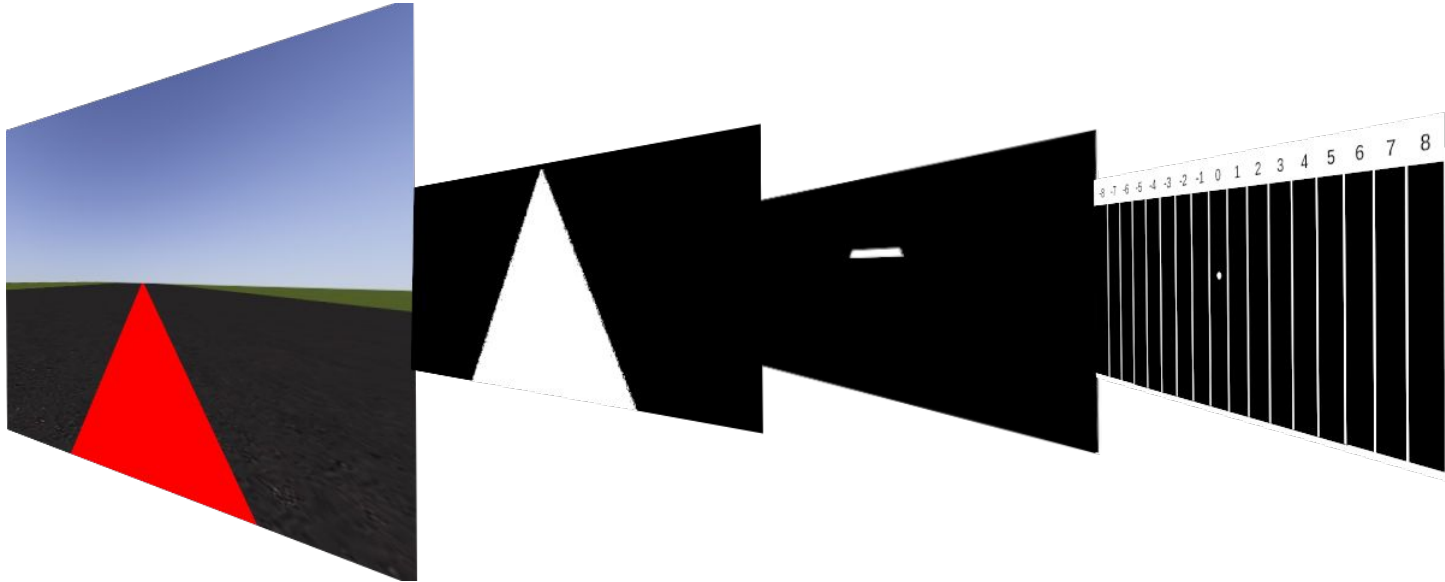
Acción	0	1	2
V. Lineal (m/s)	3	2	2
V. Angular (rad/s)	0	1	-1

Acción	0	1	2	3	4
V. Lineal (m/s)	3	2	2	1	1
V. Angular (rad/s)	0	1	-1	1.5	-1.5

Acción	0	1	2	3	4	5	6
V. Lineal (m/s)	3	2	2	1.5	1.5	1	1
V. Angular (rad/s)	0	1	-1	1	-1	1.5	-1.5

4. Aprendizaje por refuerzo de un controlador visual

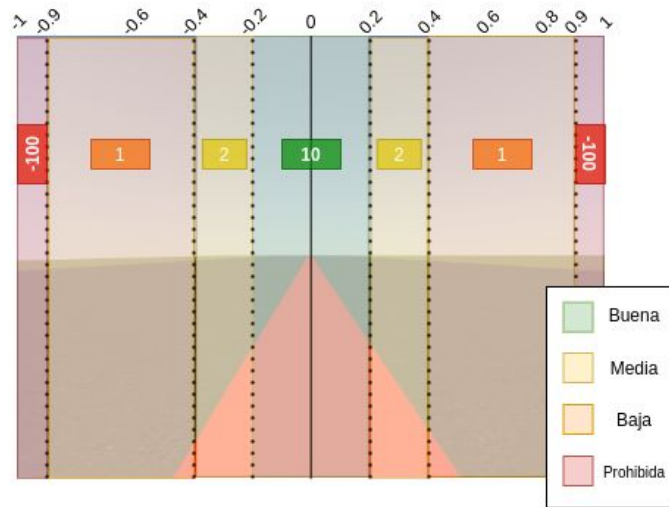
Flujo de procesamiento:



[0]

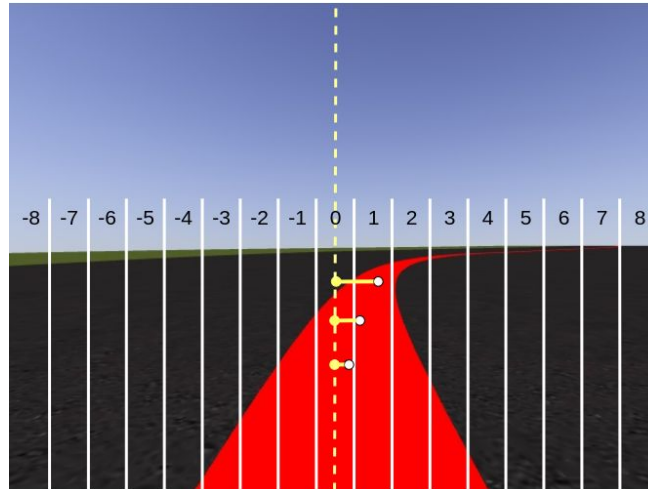
4. Aprendizaje por refuerzo de un controlador visual

Cálculo de la recompensa:



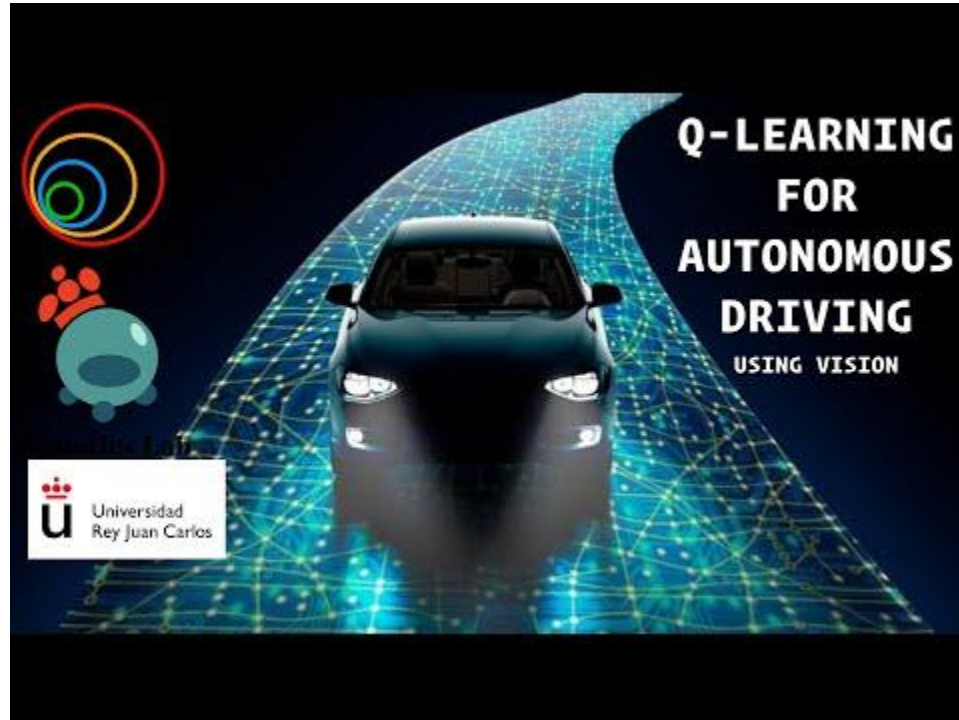
4. Aprendizaje por refuerzo de un controlador visual

Ejemplo para tres puntos de percepción:



5. Validación experimental

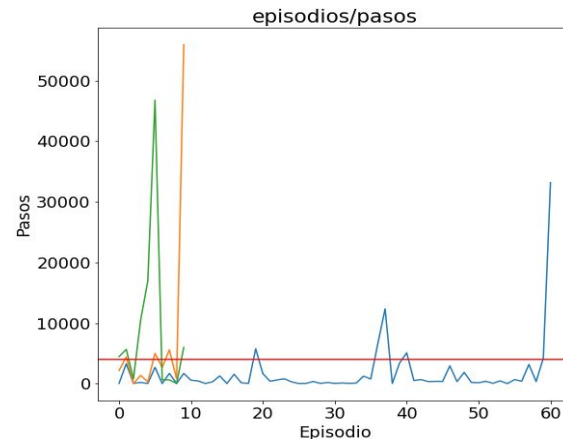
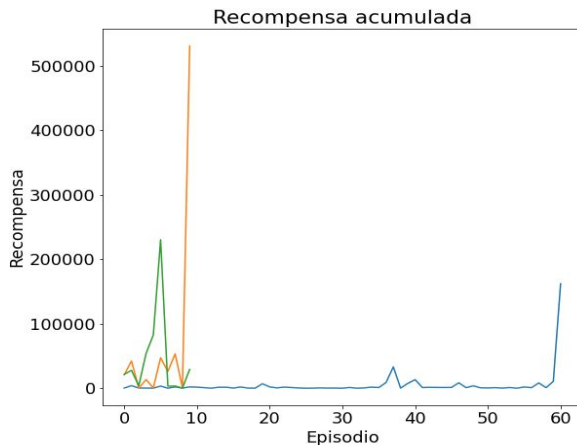
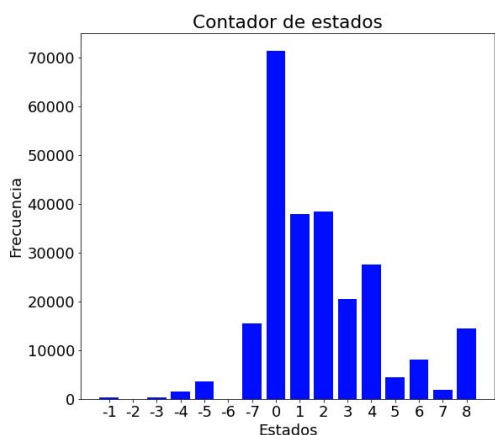
5. Validación experimental



5. Validación experimental

Análisis de los entrenamientos

- Entrenamiento en el Circuito Simple con un punto de percepción y conjunto de acciones simple.



5. Validación experimental

Análisis de los entrenamientos

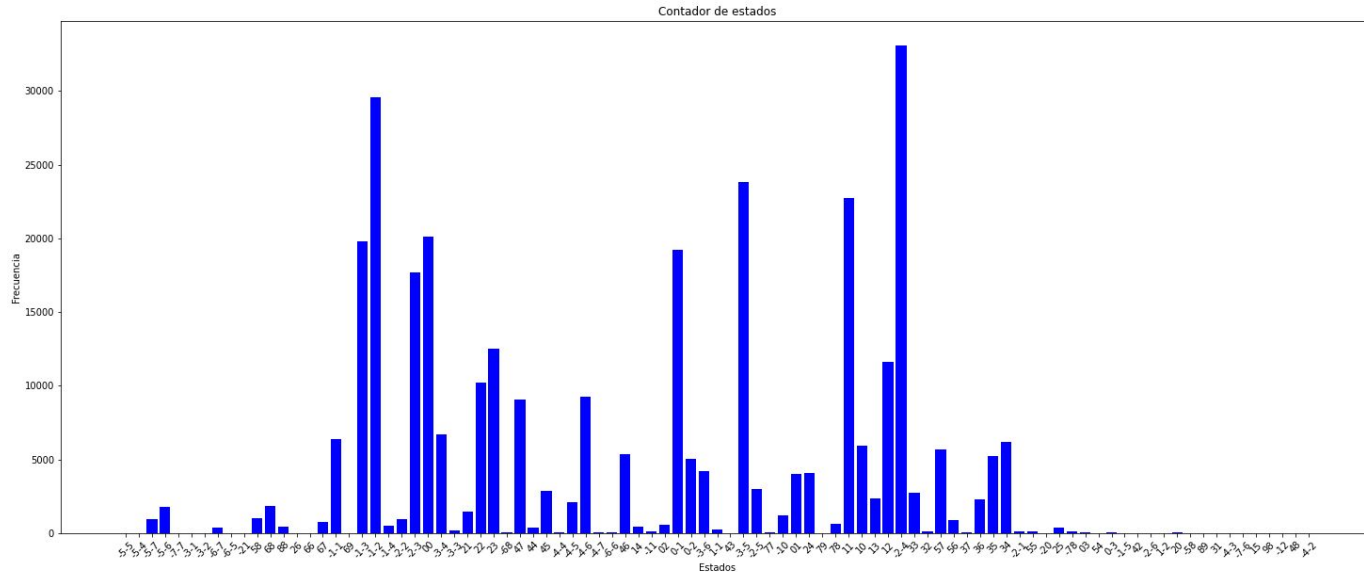
- Entrenamiento en el Circuito Simple con un punto de percepción y conjunto de acciones simple.

Entrenamiento	1	2	3
Vuelta completada	Sí	Sí	Sí
Tiempo	20:06 min	6:53 min	4:26 min
Épocas hasta completar	20	2	1
Tamaño de la Tabla-Q	40	23	22
Total épocas	61	10	10

5. Validación experimental

Análisis de los entrenamientos

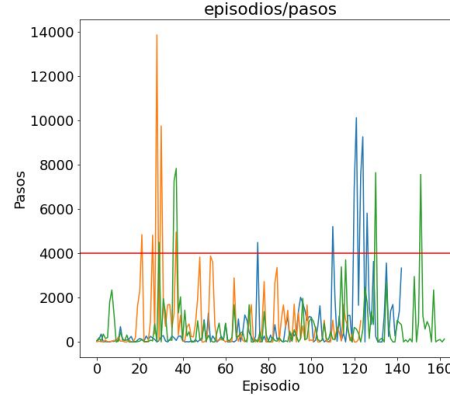
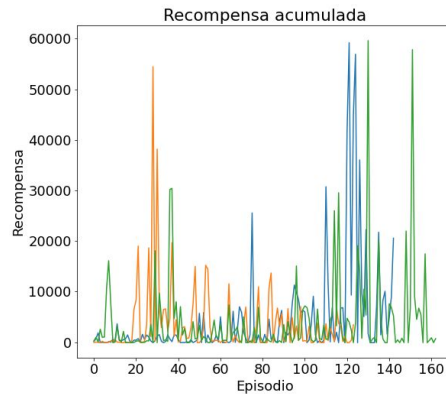
- Entrenamiento en Nürburgring con dos puntos de percepción y conjunto de acciones medio (contador de estados).



5. Validación experimental

Análisis de los entrenamientos

- Entrenamiento en Nürburgring con dos puntos de percepción y conjunto de acciones medio.



Entrenamiento	1	2	3
Vuelta completada	Si	Si	Si
Tiempo	19:32 min	9:11 min	13:12 min
Épocas hasta completar	75	21	29
Tamaño de la Tabla-Q	147	140	115
Total épocas	143	124	163

5. Validación experimental

Tabla de tiempos

Los mejores resultados se obtienen con la combinación de:

- Un punto de percepción
- Conjunto de acciones simple

Circuito	Piloto manual	Piloto RL	Diferencia
Circuito Simple	2.35 min	3.18 min	+43 seg
Nürburgring	3.19 min	4.13 min	+54 seg
Montreal	8.45 min	10.54 min	+2.09 min

6. Conclusiones



Conclusiones y líneas de trabajo futuras

Conclusiones

- Modificación, extensión y actualización del entorno Gym-Gazebo para el entrenamiento de algoritmos de aprendizaje por refuerzo con Visión.
- Entrenamientos con diferentes combinaciones que solucionan el problema. Distintos parámetros como: número de acciones o niveles de percepción.
- La mejor combinación de parámetros devuelve una ejecución con resultados satisfactorios.



Conclusiones y líneas de trabajo futuras

Líneas futuras

- Entrenamiento de los parámetros de v y w .
- Extensión del número de problemas robóticos entrenados con aprendizaje por refuerzo.
- Aprendizaje con refuerzo profundo usando DQN.



¡Muchas Gracias!

- Repositorio:

`https://github.com/RoboticsLabURJC/2019-tfm-ignacio-arranz`

- Cuaderno de bitácora:

`https://roboticslaburjc.github.io/2019-tfm-ignacio-arranz/`