

# Control de drones mediante Reinforcement Learning en plataformas reales

Miguel Fernández Cortizas

# Índice general

<b>1</b>	<b>Introducción</b>	<b>2</b>
<b>2</b>	<b>Estado del arte</b>	<b>3</b>
<b>3</b>	<b>Background</b>	<b>4</b>
<b>4</b>	<b>Reinforcement learning</b>	<b>5</b>
<b>5</b>	<b>Metodología</b>	<b>6</b>
<b>6</b>	<b>Hardware</b>	<b>7</b>
6.1	Frame* . . . . .	7
6.2	ESC* y motores . . . . .	7
6.3	Baterías . . . . .	7
6.4	Autopiloto . . . . .	7
6.4.1	Fase de Potencia . . . . .	7
6.4.2	El microcontrolador (ESP32) . . . . .	8
6.4.3	Sensores . . . . .	8
6.5	Otros (Receptora radio) . . . . .	9
<b>7</b>	<b>Experimentos</b>	<b>10</b>
<b>8</b>	<b>Discusión</b>	<b>11</b>
<b>9</b>	<b>Conclusiones y trabajo futuro</b>	<b>1</b>

## Resumen

# Introducción

gola

# Estado del arte

Las teoría clásica de control empleada para estabilizar un cuadricóptero requiere un fino ajuste de los parámetros del modelo. Los últimos avances en aprendizaje automático han permitido que se desarrollen nuevos algoritmos de control empleando técnicas de aprendizaje por refuerzo y redes neuronales.

En 2004, HJ Kim et al. [1] emplearon algoritmos de *reinforcement learning* para estabilizar (*hover*) un helicóptero y conseguir realizar maniobras acrobáticas. En 2006 Andrew Y. et al [2] siguieron con esta investigación consiguiendo que el helicóptero se estabilizara al revés (*inverted hover*). En 2010 Travis Dierks et al. [3] desarrollaron un controlador no lineal, basado en redes neuronales, para estabilizar un cuadricóptero y seguir trayectorias.

Unos años después, en 2017 Jemin Hwangbo et al. [4] desarrollaron un método para controlar un quadricóptero con una red neuronal usando técnicas de *reinforcement learning*. En 2018 William Koch et al. [5] desarrollaron un entorno de simulación, GYMFC, para el desarrollo de sistemas de control empleando RL.

# Background

# Reinforcement learning

El aprendizaje por refuerzo o *Reinforcement learning* es un área del aprendizaje automático o *Machine Learning* en el que un agente interactúa con un medio, el cual le da recompensas en función de sus acciones.

Elementos del Reinforcement Learning [6]

- **Policy.** Define el comportamiento que va
- **Policy.** Define el comportamiento que va
- **Policy.** Define el comportamiento que va
- **Policy.** Define el comportamiento que va

# Metodología



# Hardware

Enumerar las partes de las que consta un dron

## 6.1. Frame\*

## 6.2. ESC\* y motores

## 6.3. Baterias

## 6.4. Autopiloto

En los drones, el sistema que se encarga de estabilizar al cuadricóptero y hacerlo pilotable se denomina la controladora de vuelo o el Autopiloto. Existe una gran variedad de controladoras en el mercado, pero para este trabajo se ha diseñado una controladora propia con el fin de poder tener acceso a todos los sensores y a implementar el algoritmo de control de forma óptima. El autopiloto consta de 3 partes diferenciadas: la electrónica de potencia, el microcontrolador y los sensores. A continuación [se tratará](#) sobre estas partes con más detalle.

[Estaría bien un par de imágenes de la PCB \(anverso y reverso\)](#)

### 6.4.1. Fase de Potencia

Con el fin de poder gestionar la potencia entregada por las baterias a la placa y a los motores se ha diseñado una etapa de potencia en la que se debe mencionar dos partes: el interruptor de potencia y el regulador a 3.3 Voltios.

#### Interruptor de potencia

Los motores del dron pueden llegar a consumir 12 Amperios cada uno, lo que los cuatro motores pueden llegar a consumir 48 Amperios. Un interruptor con tamaño reducido no puede manejar tanta corriente, por ello se ha empleado un transistor MOSFET de canal P por el que pueden circular hasta 100 Amperios, con el fin de abrir o cerrar el paso de corriente desde las baterías al resto de la placa. El MOSFET se controla con un interruptor de poca potencia entre drenador y puerta.

Cuando se cierra el interruptor se alimenta directamente al ESC y al regulador de tensión.

### Regulador a 3.3V

La electrónica digital de la PCB se alimenta y emplea lógica a 3.3 Voltios, por lo que no la podemos conectar a las baterías de 11.1 Voltios. Para adecuar la tensión se ha escogido un regulador Step-down de tipo Buck ([¿explico como funciona un convertidor Buck?](#)). El circuito integrado que se encarga de conmutar la fuente es el chip AP3211.

### 6.4.2. El microcontrolador (ESP32)

El microcontrolador por el que se ha optado para este Autopiloto es el ESP32, un microcontrolador de doble núcleo con dos CPUs XTensaL6 con arquitectura Harvard [7]. El ESP32 tiene una frecuencia de reloj de hasta 240MHz ,y cuenta con una antena WiFi a 2,4 GHz y conexión Bluetooth 4.2 BLE [8]. Los motivos por los que se ha decidido emplear este microcontrolador son:

- Elevada frecuencia de procesamiento y dos núcleos de procesamiento.
- Antena WiFi incorporada.
- Bajo consumo de potencia.

Para poder programar el microcontrolador se utiliza un convertidor USB (Bus Serie Universal) a UART (Transmisor-Receptor Asíncrono Universal) que permite conectar por USB el microcontrolador para poder programarlo y hacer depuración utilizando comunicaciones Serial. El chip que realiza esta función es el CP2104.

### 6.4.3. Sensores

La principal fuente de información procedente del exterior que recibe una controladora de vuelo se la proporcionan las unidades de medición inercial (IMU). Las IMUs son dispositivos electrónicos que son capaces de medir aceleraciones, velocidades y detectar la orientación de un sistema. El principal problema de estos sensores es que suelen sufrir error acumulativo. [¿profundizo en los sensores MEMS \(imus electrónicas \)?](#)

Otros sensores utilizados frecuentemente en los autopilotos son brújulas (se encuentran integrados en la IMU para corregir errores de orientación) y barómetros (para estimar la altitud a la que se encuentra el dron).

Nuestro autopiloto cuenta con dos IMUs de 9 Grados de Libertad y un barómetro para conseguir una mejor estimación del estado del cuadricóptero:

1. **BNO 055 (BOSCH)**: El circuito integrado de Bosch es un sensor "inteligente" que incluye los sensores y la fusión de las lecturas de los distintos sensores en un único componente. Este sensor nos proporciona estimaciones del estado con muy poca deriva. [argumentar un poco mejor](#)

2. **MPU 9250 (TDK InvenSense)**: El sensor inercial de TDK tiene una mejor respuesta dinámica, aunque la fusión de las lecturas del sensor se realiza externamente en el microcontrolador del dron.
3. **BMP388 (BOSCH)**:

## 6.5. Otros (Receptora radio)

# Experimentos

# Discusión

# Conclusiones y trabajo futuro

# Bibliografía

- [1] H. J. Kim, M. I. Jordan, S. Sastry, and A. Y. Ng, “Autonomous helicopter flight via reinforcement learning,” in *Advances in neural information processing systems*, 2004, pp. 799–806.
- [2] A. Y. Ng, A. Coates, M. Diel, V. Ganapathi, J. Schulte, B. Tse, E. Berger, and E. Liang, “Autonomous inverted helicopter flight via reinforcement learning,” in *Experimental robotics IX*. Springer, 2006, pp. 363–372.
- [3] T. Dierks and S. Jagannathan, “Output feedback control of a quadrotor uav using neural networks,” *IEEE transactions on neural networks*, vol. 21, no. 1, pp. 50–66, 2010.
- [4] J. Hwangbo, I. Sa, R. Siegwart, and M. Hutter, “Control of a quadrotor with reinforcement learning,” *IEEE Robotics and Automation Letters*, vol. 2, no. 4, pp. 2096–2103, 2017.
- [5] W. Koch, R. Mancuso, R. West, and A. Bestavros, “Reinforcement learning for uav attitude control,” *ACM Transactions on Cyber-Physical Systems*, vol. 3, no. 2, p. 22, 2019.
- [6] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [7] “Esp32 technical reference manual,” Espressif Systems, 2018. [Online]. Available: [https://www.espressif.com/sites/default/files/documentation/esp32\\_technical\\_reference\\_manual\\_en.pdf](https://www.espressif.com/sites/default/files/documentation/esp32_technical_reference_manual_en.pdf)
- [8] “Esp32 datasheet,” Espressif Systems, 2019. [Online]. Available: [https://www.espressif.com/sites/default/files/documentation/esp32\\_datasheet\\_en.pdf](https://www.espressif.com/sites/default/files/documentation/esp32_datasheet_en.pdf)