



UNIVERSIDAD DE TALCA
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA CIVIL EN COMPUTACIÓN

**Evasión de obstáculos con flujo óptico y redes
neuronales para vehículos no tripulados**

JORGE GÓMEZ VALDERRAMA

Profesor Guía: MATTHEW BARDEEN

Memoria para optar al título de
Ingeniero Civil en Computación

Curicó – Chile
mes, año



UNIVERSIDAD DE TALCA
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA CIVIL EN COMPUTACIÓN

**Evasión de obstáculos con flujo óptico y redes
neuronales para vehículos no tripulados**

JORGE GÓMEZ VALDERRAMA

Profesor Guía: MATTHEW BARDEEN

Profesor Informante: PROFESOR INFORMANTE 1

Profesor Informante: PROFESOR INFORMANTE 2

Memoria para optar al título de
Ingeniero Civil en Computación

El presente documento fue calificado con nota: _____

Curicó – Chile

mes, año

Dedicado a ...

AGRADECIMIENTOS

Agradecimientos a ...

TABLA DE CONTENIDOS

	página
Dedicatoria	I
Agradecimientos	II
Tabla de Contenidos	III
Índice de Figuras	V
Índice de Tablas	VI
Resumen	VII
1. Introducción	8
1.1. Descripción del contexto	8
1.2. Objetivos	8
1.2.1. Objetivo general	8
1.2.2. Objetivos específicos	8
1.3. Alcances	8
2. Marco teórico	9
2.1. Vehículo aéreo no tripulado	9
2.2. Mecanismos de percepción	10
2.3. Software	10
2.3.1. OpenCV	10
2.3.2. Visión por Computador	11
2.3.3. OpenNN	12
2.3.4. Flujo óptico	12
2.3.5. Redes neuronales	16
2.4. Hardware	21
2.4.1. Robot omnidireccional	21
2.4.1.1. Rueda omnidireccional	22
2.4.1.2. Motores reductores	22

2.4.1.3.	Servo motores	22
2.4.1.4.	Controlador para motores	22
2.4.2.	Microcomputador	25
2.4.3.	Microcontrolador	25
2.4.4.	Camara	25
2.4.5.	Sensor Óptico	25
3.	Implementación	27
3.1.	Flujo óptico	27
3.1.1.	Captura de flujo óptico	27
3.1.2.	Procesado de flujo óptico	27
3.2.	Red neuronal	27
3.2.1.	Diseño de capas	27
3.2.2.	Entrenamiento	27
3.3.	Robot omnidireccional	29
4.	Pruebas	38
5.	Resultados	41
	Glosario	42
	Bibliografía	43
	Anexos	
A:	Configuración impresora 3D	46
A.1.	La primera sección del primer anexo	50
A.2.	La segunda sección del primer anexo	50
A.2.1.	La primera subsección de la segunda sección del primer anexo	50
B:	El segundo Anexo	51
B.1.	La primera sección del segundo anexo	51

ÍNDICE DE FIGURAS

	página
2.1. Proyección de un punto P_0 con movimiento v_0 en la escena a un punto P_i con movimiento v_i sobre la imagen.	13
2.2. Ejemplo de captura del flujo optico, por medio de cuandos de imagen consecutivos.	14
2.3. Perceptrón con tres entradas: x_1 , x_2 y x_3 que genera una salida binaria, con valor 0 o 1.	17
2.4. Red neuronal conformada por 3 capas, la primera con 3 perceptrones, la segunda con 4 y la tercera y ultima con 1 perceptron.	18
2.5. Función que utiliza una neurona sigmoide, como se puede ver la salida de la neurona sera un valor continuo, entre 0 y 1.	20
2.6. Función escalonada que utiliza un perceptron, el que tiene una salida binaria, con valores entre 0 y 1.	20
2.7. Direcciones en las cuales puede moverse una rueda omnidireccional. .	22
2.8. Funcionamiento de un puente H. A estructura del puente H. B interruptores S1 y S4 cerrado. C interruptores S3 y S2 cerrados.	23
2.9. Control logico dentro L298 sobre uno de sus puente H.	24
3.1. Diseño de rueda omnidireccional.	30
3.2. Acople para unir las ruedas del robot con la caja reductora de los motores.	30
3.3. Extension para unir los motores.	31
3.4. Cuerpo principal donde se alojan todos los componentes del robot omnidireccional.	32
3.5. Compartimiento para colocar en serie las tres baterias de litio 18650. .	32
3.6. Soporte para la instalación del sensor óptico que registra las trayectorias realizadas por el robot.	33
3.7. Soporte sobre el cual va la cámara digital.	34
3.8. Ensamble robot omnidireccional.	35
3.9. Motor del robot omnidireccional.	36

ÍNDICE DE TABLAS

	página
2.1. Paquete de datos del movimiento del <i>mouse</i>	26

RESUMEN

Aquí va el resumen (en Castellano)...

1. Introducción

1.1. Descripción del contexto

1.2. Objetivos

1.2.1. Objetivo general

1.2.2. Objetivos específicos

1.3. Alcances

2. Marco teórico

2.1. Vehículo aéreo no tripulado

Un vehículo aéreo no tripulado, UAV (por sus siglas del inglés: *unmanned aerial vehicle*) o dron, es cualquier aeronave que no cuente con un piloto o tripulación a bordo, por lo que ésta puede volar de forma autónoma por medio de un computador a bordo o ser controlada remotamente por un humano [15].

Dentro de los vehículos aéreos no tripulados se encuentran los multirrotores o multihélices, estas son aeronaves que generan sustentación por medio del giro de alas o hélices sobre un eje, siendo este conjunto llamado rotor. La Organización de Aviación Civil Internacional (CIAO por sus siglas del inglés: *International Civil Aviation Organization*) los define como aeronaves que se mantienen en vuelo por las reacción del aire en uno o más rotores. En el caso de los multirrotores estos están compuestos por dos o más rotores, se diferencian de los helicópteros en la forma que logran el control y la estabilidad, estos tienen mecanismos que varían el paso de las hélices modificando su ángulo de ataque, en cambio los multirrotores lo hacen por medio del cambio de velocidad relativa de cada rotor, produciendo cambios en el empuje y torque.

Un tipo de multirrotor son los cuadricópteros los cuales tiene dos pares de rotores con hélices iguales y paso fijo, dos giran en sentido horario y los otros dos en sentido antihorario, esta configuración permite que el torque de cada rotor se anule con el rotor correspondiente que gira en sentido contrario, de esta forma el cuadricóptero se estabiliza y no gira sobre su propio eje [2].

Como se mencionó anteriormente los movimientos de los multirrotores se logran variando la velocidad de cada rotor independientemente, referencia a la imagen, con esto el quadricóptero puede moverse sobre 3 ejes perpendiculares, por medio de los movimientos de *pitch*, *roll* y *yaw*, además de desplazarse por el eje *Z* variando su altitud.

Con el movimiento *pitch* el quadricóptero rota sobre el eje *X* y se desplaza hacia adelante o atrás por el eje *Y*, cuando ejecuta el movimiento de *roll* rota sobre el eje *Y* y se desplaza de izquierda a derecha sobre el eje *X* y gira sobre el eje *Z* con el movimiento de *yaw*. Con esto se logra tener control sobre 4 grados de libertad.

2.2. Mecanismos de percepción

2.3. Software

2.3.1. OpenCV

OpenCV (Open Source Computer Vision Library) es una librería open source (código abierto) de *computer visión* (visión por computador), distribuida bajo la licencia BSD, la cual está escrita en C y C++, siendo además multiplataforma. También existen interfaces para Python, Ruby, Matlab y otros lenguajes.

OpenCV está diseñado para ser eficiente computacionalmente y enfocado en ser utilizado en aplicaciones de procesamiento en tiempo real. Para esto, fue escrito en C optimizado y toma ventaja de los procesadores multi núcleo.

El proyecto fue iniciado por *Intel Research Labs* en 1990 con el propósito de hacer la infraestructura de la visión por computador universalmente disponible y de esta forma ampliar a su vez OpenCV fue pasado a código abierto [4].

2.3.2. Visión por Computador

Computer Vision (visión por computador) es una disciplina que busca que los computadores puedan tener una comprensión de alto nivel sobre imágenes o videos digitales, por ejemplo, con el objetivo de automatizar tareas que el sistema visual humano ya puede hacer[3, 19, 18].

Dentro de las tareas de la visión por computador están: la adquisición, procesamiento, análisis, el entendimiento de imágenes digitales y la extracción de información multi dimensional del mundo real, para producir información numérica o simbólica [12, 16, 13, 9].

Dado que nosotros percibimos el mundo principalmente por medio de la vista, parecería sencillo traspasar nuestra experiencia a la visión por computador, pero en realidad es una tarea muy difícil, debido a la complejidad de cómo funciona nuestro cerebro, contamos con múltiples sistemas que reciben distinta información segmentada desde el sistema visual e incluso otros sistemas, pero en el caso de la visión por computación solo contamos con una imagen (en el caso de video una secuencia de imágenes) y esto es todo lo el computador puede "ver".

Cuando hablamos de una imagen nos estamos refiriendo a la información que nos es capaz de entregar la luz que es recibida desde el entorno o una escena, esta información varia dependiendo del contexto en que se hable, si nos referimos a un marco biológico la luz incidirá en un ojo y las células en la retina generaran las señales eléctricas que el cerebro interpretará como imágenes, si hablamos de una cámara la imagen se formará en una película fogatica o en un sensor digital [4].

Dado el ámbito de este documento, es preciso profundizar en las imágenes digitales, esta puede ser definida como la integración y muestreo de información analógica (luz incidiendo en un sensor fotográfico) en un dominio espacial. Esta consiste en un arreglo rectangular de pixeles (x, y, u) , cada uno contiene una ubicación $(x, y) \in \mathbb{Z}^2$ y un valor u , muestreado en una ubicación (x, y) . \in es conjunto de puntos (x, y) del arreglo rectangular [12]. Entonces podemos definir una imagen como:

$$I = \{(x, y) : 1 \leq x \leq N_{cols} \wedge 1 \leq y \leq N_{rows}\} \subset \mathbb{Z}^2 \quad (2.1)$$

2.3.3. OpenNN

2.3.4. Flujo óptico

El flujo óptico es definido como el cambio de la estructura de la luz en una imagen, por ejemplo, en la retina de un ojo o en el sensor de una cámara, debido a movimiento relativo del observador y escena. Cuando los objetos se mueven frente a una cámara o esta se mueve en un entorno fijo, existe un cambio correspondiente a los movimientos en la imagen, estos cambios pueden utilizarse para recuperar información relativa al movimiento de las formas y los objetos.

Podemos definir un campo de movimiento, en el cual asignamos un vector de velocidad a cada punto de la imagen. En un instante en particular, un punto P_i en la imagen corresponde con algún punto P_0 en la superficie de un objeto. Los dos puntos son conectados por la ecuación de proyección. Si consideramos una proyección de perspectiva, una línea se extiende de un punto en la imagen, pasa por el centro de la lente (en el caso de tratarse de una cámara), hasta un punto en la superficie de la escena.

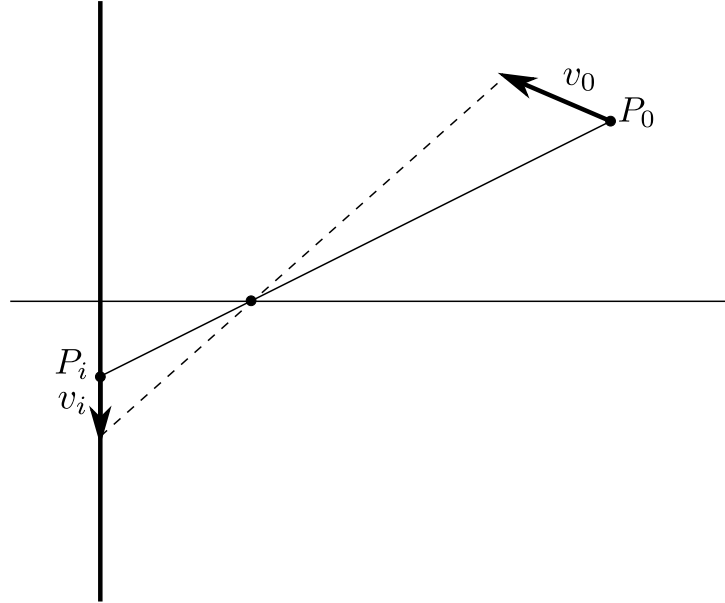


Figura 2.1: Proyección de un punto P_0 con movimiento v_0 en la escena a un punto P_i con movimiento v_i sobre la imagen.

El punto P_0 tiene una velocidad v_0 relativa a la cámara, esto induce un movimiento v_i en el punto P_i correspondiente en la imagen, como se puede ver en la figura 2.1[10].

Los vectores de velocidad producidos por el movimiento aparente son descritos de forma distinta, dependiendo del contexto en que se presenten, de un punto de vista biológico los cambios estructurados en los patrones de la luz en la retina de un ojo dejan la impresión de movimiento, en visión por computación los cambios en la escena son representados por serie de *image frames* (cuadros de imagen), la figura 2.2 muestra una secuencia de tres cuadros, mediante un muestreo espacial y temporal de la luz incidente en la imagen, es decir, como se desplazan los píxeles en la imagen a través del tiempo.

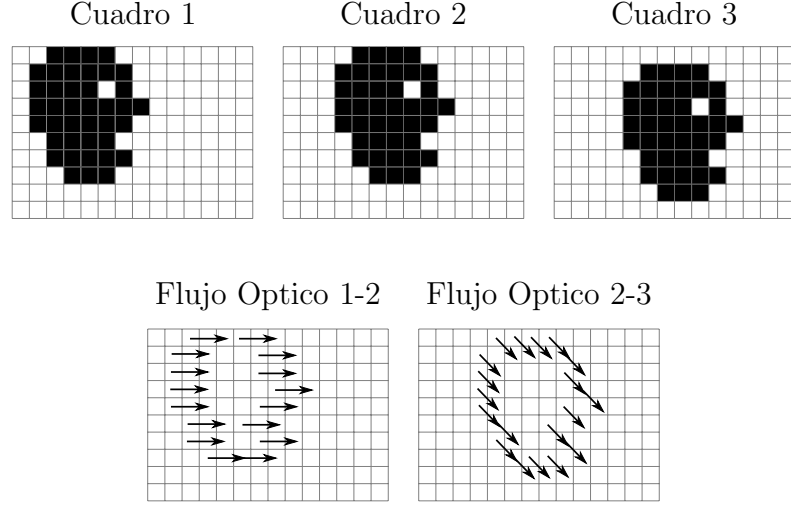


Figura 2.2: Ejemplo de captura del flujo optico, por medio de cuandos de imagen consecutivos.

Para calcular el flujo óptico por medio el análisis de imágenes se utiliza el método de *Lucas-Kanade*, el cual funciona bajo dos suposiciones: la intensidad de un pixel en un determinado objeto no cambia entre cuadros consecutivos, refiriéndose a que un objeto no cambia en la escena, siempre la imagen proyectada de éste en la cámara es la misma. La segunda suposición dice que los vecinos cercanos a un pixel tiene un movimiento similar a este, esto permite evaluar una zona en la imagen.

Consideramos un pixel $I(x, y, t)$ en la primera imagen (x e y coordenadas dentro de la imagen y t un instante de tiempo) el cual se mueve una distancia de (dx, dy) en el siguiente cuadro después de un instante de tiempo dt , bajo las suposiciones mencionadas, podemos decir que:

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (2.2)$$

Usando series de taylor aproximamos el lado derecho de la ecuación, removiendo términos semejantes y dividiendo por dt obtenemos:

$$f_x u + f_y v + f_t = 0 \quad (2.3)$$

donde:

$$f_x = \frac{\partial f}{\partial x} ; f_y = \frac{\partial f}{\partial y} \quad (2.4)$$

Esta es la ecuación del flujo óptico, donde f_x y f_y son gradientes de la imagen, como f_t es gradiente del tiempo. Pero (u, v) son desconocidas, por lo que no podemos resolver la ecuación con dos incógnitas. Aquí es donde el método *Lucas-Kanade* toma grupos de píxeles de 3×3 , considerando las suposiciones mencionadas anteriormente, donde estos tienen el mismo movimiento. Ahora es posible resolver (f_x, f_y, f_t) ya que contamos con nueve ecuaciones y las mismas dos incógnitas, resultando con la siguiente ecuación:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_i f_{x_i}^2 & \sum_i f_{x_i} f_{y_i} \\ \sum_i f_{x_i} f_{y_i} & \sum_i f_{y_i}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i f_{x_i} f_{t_i} \\ -\sum_i f_{y_i} f_{t_i} \end{bmatrix} \quad (2.5)$$

De esta forma se toma un grupo de píxeles y calculamos el flujo óptico sobre ellos, pero el método descrito funciona solo con pequeños movimientos, fallando en condiciones cuando los píxeles en la imagen sufren grandes movimientos, para solucionar este problema utilizamos el algoritmo de pyramid, el cual escala la imagen removiendo los pequeños movimientos y reduciendo los grandes movimientos a pequeños [1].

El algoritmo de *pyramid* se encarga de realizar *downsampled*, este término se refiere a reducir la tasa de muestreo de alguna señal, a una imagen sucesivamente hasta algún límite dado, creando una colección de imágenes llamada pirámide. Existen dos tipos variantes del algoritmo de *pyramid*: la gaussiana y la laplaciana, en el método de *Lucas-Kanade* se utiliza la variante gaussiana, por lo que esta será descrita en detalle.

Para agregar a la pirámide una nueva capa G_{i+1} , debemos aplicar a la capa G_i un filtro gaussiano y luego eliminar las columnas y filas pares, esto nos genera una nueva imagen con un cuarto del área de la imagen en la capa G_i , iterando este proceso desde la imagen original G_0 producimos la pirámide entera [4].

Cuando se realiza un proceso de muestreo en una señal, en este caso de una ima-

gen, puede producirse *aliasing*, el cual consiste en la distorsión de la señal orinal una vez que esta es muestreada, para evitar esto el teorema de muestreo *Nyquist-Shannon* dice que es necesario muestrear una señal al doble de su frecuencia [8].

Es por esta razón que *pyramid* aplica un filtro gaussiano para la capa G_i de la pirámide antes de producir la capa G_{i+1} , esto reduce la frecuencia de la imagen en la capa G_i dando como resultado una imagen con menor distorsión en la capa G_{i+1} , permitiendo tener una imagen de menor tamaño en una capa G_j , al final de la iteración, la cual tiene una baja distorsión frente a la imagen original en la capa G_0 .

2.3.5. Redes neuronales

Una red neuronal tiene la finalidad de ser un modelo de una red neuronal biológica, que recibe un número determinado de entradas, las cuales son procesadas dentro de la red y producen una o más salidas. Una red neuronal biológica está conformada por la interconexión de neuronas, las cuales son células nerviosas que reciben estímulos y conducen el impulso nervioso, por medio de un potencial de acción, a otras neuronas [5]. De la misma forma una red neuronal está formada por neuronas, a las que nos referimos por ahora como perceptrón, este funciona tomando varias entradas x_1, x_2, \dots , para producir una sola salida binaria, como se muestra en la figura 2.3.

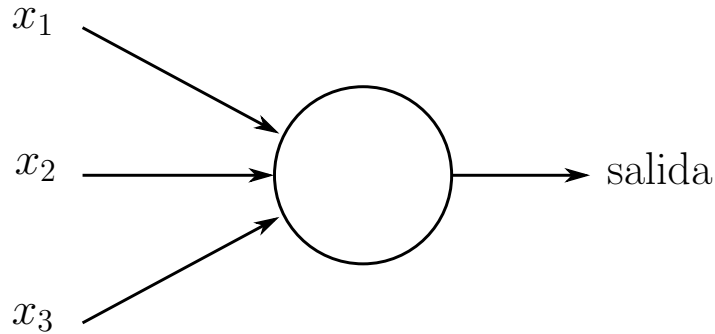


Figura 2.3: Perceptrón con tres entradas: x_1 , x_2 y x_3 que genera una salida binaria, con valor 0 o 1.

En este caso se utilizan tres entradas, para computar la salida se introducen *weights* (pesos) w_1, w_2, w_3 , números reales que representan la importancia de cada entrada con respecto a la salida. Finalmente, la salida de la neurona es calculada como la sumatoria $\sum_j w_j x_j$, la salida binaria es 1 o 0 dependiendo de si el resultado de la sumatoria es menor o mayor al *threshold value* (valor de activación), el cual es un número real, parámetro de la neurona.

$$\text{salida} = \begin{cases} 0 & \text{si } \sum_j w_j x_j \leq \text{activación} \\ 1 & \text{si } \sum_j w_j x_j > \text{activación} \end{cases} \quad (2.6)$$

La red neuronal es conformada por la interconexión de perceptrones, los que se agrupan en *layers* (capas), de tal modo que las salidas de los perceptrones en la primera capa son a su vez las entradas para los perceptrones de la segunda capa, así sucesivamente por la totalidad de las capas que constituyan la red neuronal.

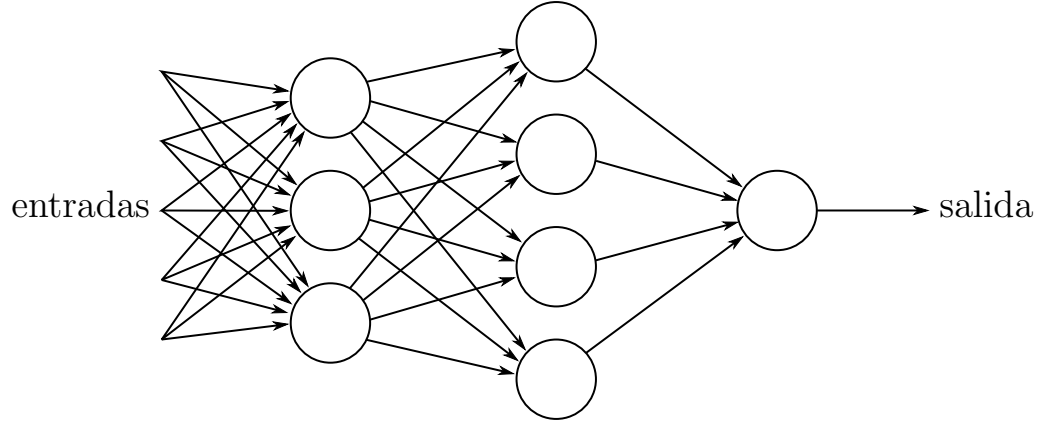


Figura 2.4: Red neuronal conformada por 3 capas, la primera con 3 perceptrones, la segunda con 4 y la tercera y ultima con 1 perceptron.

En la definición de perceptrón se menciona que este solo cuenta con una salida, en la figura 2.4 se muestra que un perceptrón tiene múltiples salidas, pero de hecho es la misma única salida que se comparte como entrada para los perceptrones de la siguiente capa.

A la hora de implementar la red neuronal es posible realizar algunas simplificaciones, redefiniendo el comportamiento de los perceptrones, podemos suponer que las entradas y los pesos de estos están definidos como dos vectores, con lo que ahora para producir la salida se calcula el producto punto entre la entrada y los pesos, $w \cdot x \equiv \sum_j w_j x_j$, también se mueve el valor de activación al otro lado de la inecuación y remplazándolo por el termino de *bias* (sesgo del perceptrón), $b \equiv -\text{activación}$, ahora el comportamiento del perceptrón se define como:

$$\text{salida} = \begin{cases} 0 & \text{si } w \cdot x + b \leq 0 \\ 1 & \text{si } w \cdot x + b > 0 \end{cases} \quad (2.7)$$

El sesgo puede definirse como que tan fácil el perceptrón puede producir una

salida de valor 1, mientras más grande el sesgo más fácil producir una salida que sea 1.

Una de las principales características de las redes neuronales es su capacidad de aprender, para lograr esto la red debe variar tanto el valor de los pesos tanto como de los sesgos, para que de esta forma podamos tener la salida esperada a partir de una entrada determinada.

Para lograr este entrenamiento pequeños cambios en los parámetros de la red, pesos y sesgos, deben a su vez generar pequeños cambios en la salida, pero esto no sucede en las redes conformadas por perceptrones, cuando se aplica un pequeño cambio en los parámetros de un perceptrón produce que este cambie completamente su salida de 0 a 1, por ejemplo [14].

Este problema se soluciona remplazando los perceptrones por *sigmoid neuron* (neurona sigmoide), en este tipo de neurona se pueden aplicar pequeños cambios en sus parámetros y esos se verán reflejados también en pequeños cambios en la salida. Una neurona sigmoide funciona de forma similar a un perceptrón donde la salida es calculada por $\sigma(w \cdot x + b)$, σ es una función sigmoide, definida cómo:

$$\sigma(z) \equiv \frac{1}{1 + e^{-z}} \quad (2.8)$$

Remplazando el parámetro de la función por los de la neurona:

$$\frac{1}{1 + \exp(-\sum_j w_j x_j - b)}. \quad (2.9)$$

Al ser sigma una función continua, ahora la salida de la red no es binaria, si no que valor real entre 0 y 1, como se muestra en la figura 2.5.

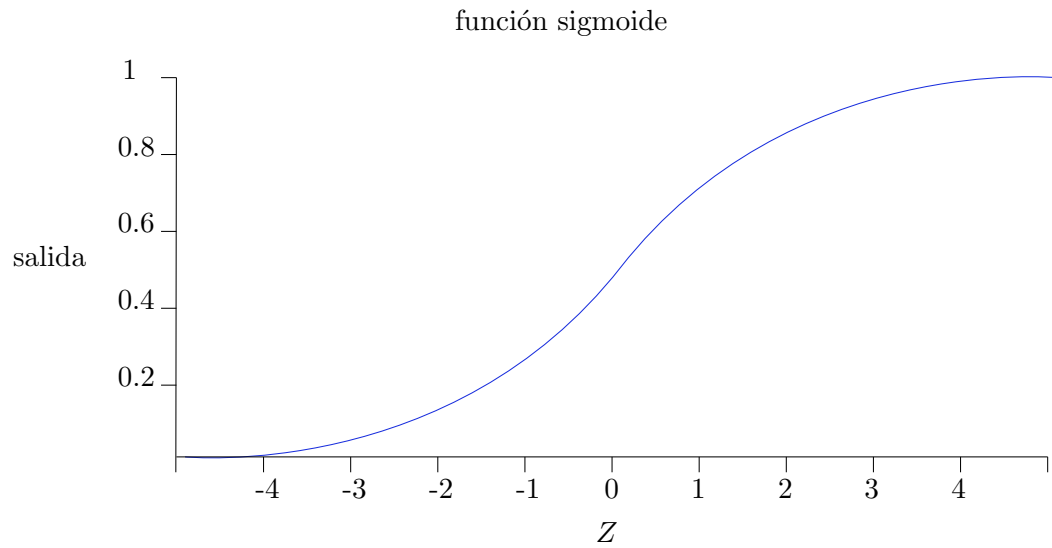


Figura 2.5: Función que utiliza una neurona sigmoide, como se puede ver la salida de la neurona será un valor continuo, entre 0 y 1.

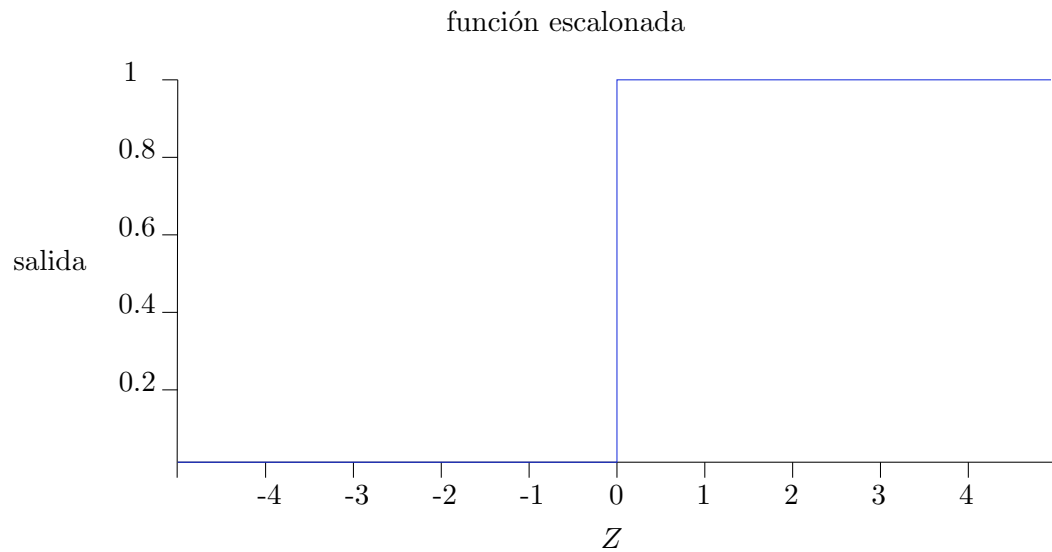


Figura 2.6: Función escalonada que utiliza un perceptron, el que tiene una salida binaria, con valores entre 0 y 1.

Ya se mencionó anteriormente que la forma en que la red neuronal aprende es ajustando los valores de los pesos y sesgos de las neuronas que la conforman, para

conseguir este objetivo se utiliza el algoritmo de *backpropagation*, primero se le entrega un vector de entrada a la red neuronal, se conoce de antemano la salida esperada que debe generar la red con estos valores, la red computa la salida y esta se compara con la salida esperada con algún método de optimización, en este caso se utiliza el descenso por gradiente.

El método de descenso por gradiente es un mecanismo de minimización, lo que busca minimizar es la función de error cuadrático medio entre la salida esperada y la salida que produce la red neuronal, una vez que se obtiene el gradiente del error este es propagado hacia atrás por la red modificando los pesos y los sesgos de las neuronas.

2.4. Hardware

2.4.1. Robot omnidireccional

Un robot omnidireccional (referido en este documento inherentemente como robot o robot omnidireccional) es un robot que tiene la capacidad de moverse en cualquier dirección sin tener que rotar antes de moverse, esto es posible a las ruedas omnidireccionales con que está construido el robot, esta tiene la capacidad de moverse libremente en dos direcciones, pueden rodar como una rueda normal o lateralmente con las ruedas ubicadas perpendicularmente a lo largo de la circunferencia, como se muestra en la figura 2.7.

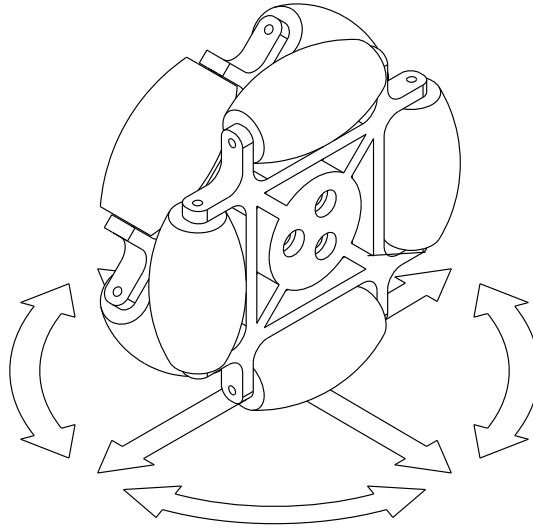


Figura 2.7: Direcciones en las cuales puede moverse una rueda omnidireccional.

Esto permite que el robot puede trazar un camino en línea recta, en cualquier dirección, sin tener que hacer giros, además de poder combinar el trayecto con algún tipo de rotación para cambiar la dirección a la que apunta el robot, si es necesario, esto es definido como un robot holonómico, en comparación con un robot normal no-holonómico el cual tiene 1,5 grados de libertad que puede desplazarse a través de los ejes coordenados X e Y , pero requiere movimientos complejos para poder moverse en el eje X . Un robot holonómico tiene 2 grados de libertad por lo que puede moverse libremente por ambos ejes X e Y [11].

Para lograr que el robot se desplace en cualquier dirección es necesario calcular la velocidad y dirección de giro para cada rueda de forma independiente,

2.4.1.1. Rueda omnidireccional

2.4.1.2. Motores reductores

2.4.1.3. Servo motores

2.4.1.4. Controlador para motores

Un controlador para motores es un dispositivo que maneja el desempeño de un motor eléctrico, este incluye mecanismos para arrancar y detener el motor, seleccionar la dirección de rotación, si es hacia adelante o atrás y regulando la velocidad

de este [17]. La mayoría de los controladores utilizan un circuito eléctrico llamado puente H (también referido como *full-bridge*, este permite aplicar corriente sobre una carga en cualquier dirección, con esto un motor de corriente continua girar en ambas direcciones [20].

El nombre es derivado de su representación gráfica, como se muestra en la figura 2.8.A, el cual cuenta con cuatro interruptores, los que son accionados de apares, según la figura 2.8.B S1 y S4 pueden estar en estado cerrado, mientras S2 y S3 en abierto, en esta configuración la corriente circula en una dirección en el motor, si en cambio S2 y S3 están cerrados y S1 y S4 abiertos como en la figura 2.8.C la corriente sobre el motor M circulara en la dirección opuesta. Los interruptores S1 y S2, al igual que S3 y S4 nunca son activados al mismo tiempo, ya que esto genera un corto circuito sobre la fuente de voltaje.

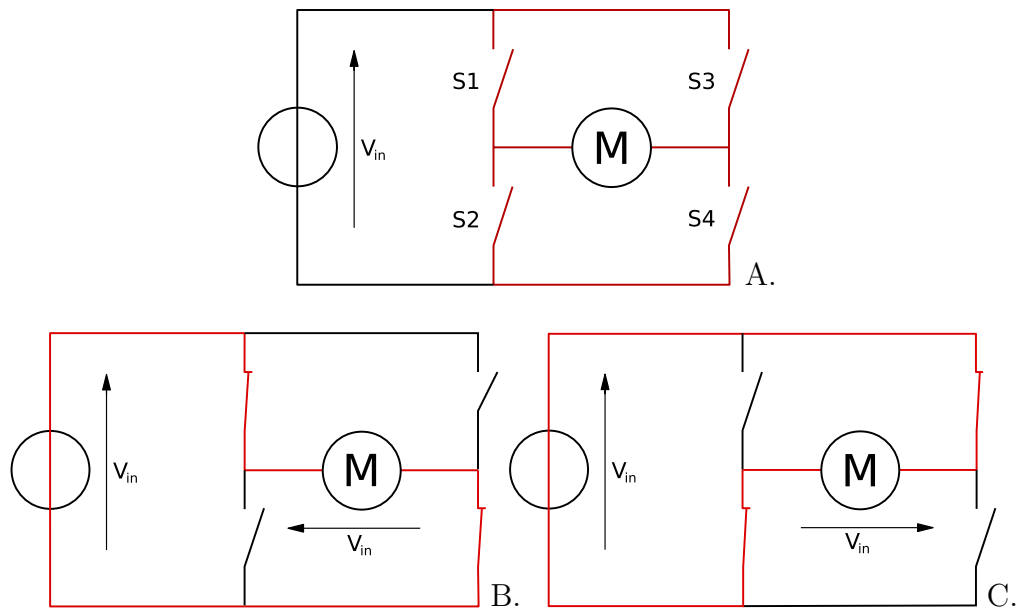


Figura 2.8: Funcionamiento de un puente H. A estructura del puente H. B interruptores S1 y S4 cerrado. C interruptores S3 y S2 cerrados.

El controlador utilizado es el L298, cual contiene dos puentes H compatibles con nivel lógico estándar TTL y puede manejar cargas inductivas, como motores de co-

riente directa. Cuenta con dos entradas de control para cada motor además de una para la habilitación del puente H. En la descripción de un puente H se mencionó el control sobre 4 interruptores, pero el controlador solo cuenta dos entradas de control por cada puente H, esto se debe a que dos interruptores nunca son activados al mismo tiempo, por ejemplo: S1 y S2, por lo que el control sobre estos puede ser compartida negando la señal de entrada a un interruptor, es decir, que cuando un interruptor está activado el otro siempre esta desactivado, esta misma situación se da para los interruptores S3 y S4. En la figura 2.9 se muestra el control lógico que realiza el L298 sobre el puente H y la utilización de transistores de canal-N como interruptores.

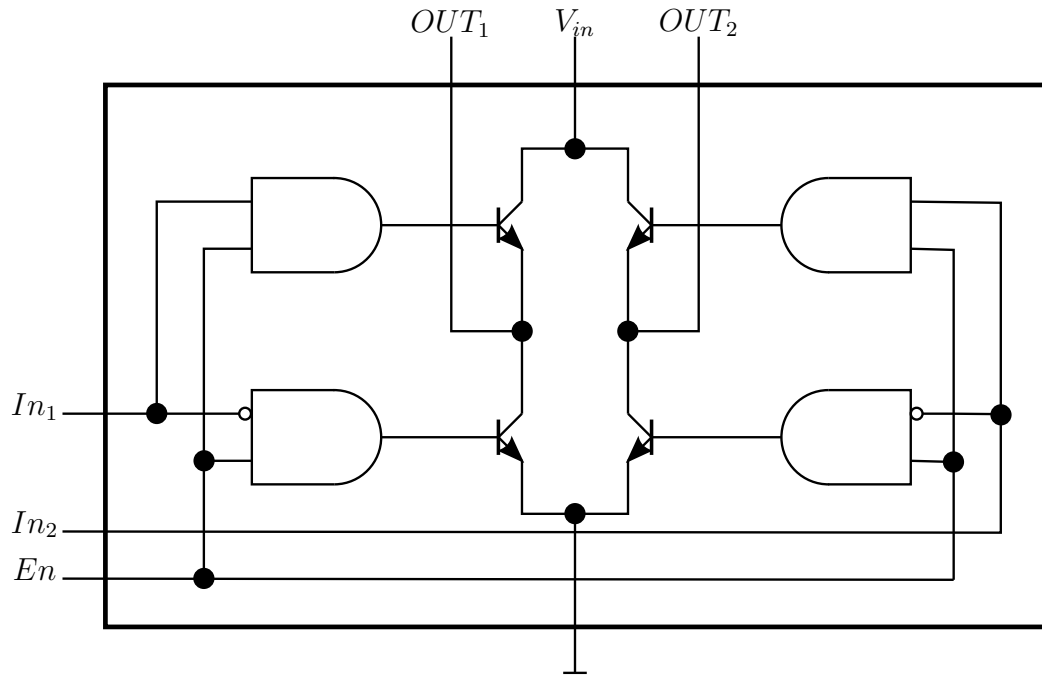


Figura 2.9: Control logico dentro L298 sobre uno de sus puente H.

2.4.2. Microcomputador

2.4.3. Microcontrolador

2.4.4. Camara

2.4.5. Sensor Óptico

El sensor óptico de movimiento acoplado al robot que tiene la función registrar las trayectorias que este sigue, proviene de un *mouse* óptico (referido en este documento indistintamente como *mouse* o *mouse* optico. Este utiliza una fuente de luz, en este caso un LED rojo y detector de luz como una matriz de fotodiodos, los que forman una pequeña cámara digital, para detectar el movimiento relativo a una superficie.

El sensor del *mouse* forma una imagen capturando las texturas sobre la superficie que se encuentra, estas texturas son resaltadas por el ángulo de incidencia de la luz proveniente del LED. Se toma una continua sucesión de imágenes de la superficie, las que son comparadas unas con otras para determinar cuánto en *mouse* se ha movido.

Para determinar esto el *mouse* evalúa dos imágenes consecutivas y busca coincidencias de la primera imagen en la segunda, cuantificando cuanto se desplazado la segunda imagen sobre la primera, así se puede cuantificar el movimiento del *mouse*.

El *mouse* toma cientos de imágenes por segundo, dependiendo de qué tan rápido se esté moviendo cada imagen va a estar desplazada con la imagen anterior por una fracción de pixel o tanto como varios pixeles. El *mouse* procesa matemáticamente las imágenes usando correlación cruzada para calcular cuánto cada imagen, de la sucesión de imágenes, se ha desplazado desde la imagen anterior.

El protocolo de comunicación utilizado es PS/2 (*IBM Personal System 2*), el cual es serial, síncrono y bireccional. Utiliza dos líneas de comunicación DATA por la que se envían los datos y CLOCK que proporciona la señal de reloj, el dispositivo que en este caso es el *mouse* siempre es el que genera la señal de reloj, cuando las dos líneas están en alto el dispositivo puede enviar datos, pero el *host* tiene el control sobre la línea de datos, por lo que puede inhibir la comunicación enviando la señal de comunicación a low [7].

	Bit 7	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0
Byte 1	Y overflow	X overflow	Y sign	X sign	always 1	Btn central	Btn derecho	Btn izquierdo
Byte 2	Movimiento X							
Byte 3	Movimiento Y							

Cuadro 2.1: Paquete de datos del movimiento del *mouse*

El *mouse* envía la información de estado de los botones, derecho, central e izquierdo, además del desplazamiento en el eje X e Y en un paquete de tres *byte*, que se muestra en la tabla 2.1. El movimiento es un entero de 9 *bite* a complemento 2, el *bit* más significativo es Y sign y X sign que aparece en el primer *byte*. Este valor representa la diferencia relativa a la posición del *mouse* desde la última vez que se envió el paquete de datos. El rango va de -255 a +255, si el rango es excedido el correspondiente *bit* de *overflow* es activado [6].

3. Implementación

3.1. Flujo óptico

3.1.1. Captura de flujo óptico

3.1.2. Procesado de flujo óptico

3.2. Red neuronal

3.2.1. Diseño de capas

3.2.2. Entrenamiento

Para entrenar la red neuronal se utiliza como entrada el flujo óptico obtenido de la captura de video de la pista de pruebas y como salida esperada la dirección en la cual se mueve el robot [agregar marco de referencia]. Para modificar los pesos al interior de la red, que se traduce como el aprendizaje de la red, se utiliza la técnica de *backpropagation*.

La generación de set de datos del cual la red aprende se hace por medio de un esquema [mejor traducción] *on-the-fly*, el cual tiene como objetivo que la red neuronal imite el comportamiento de un humano al realizar la misma tarea, la que cual es desplazar el robot por la pista de pruebas. Por lo que el flujo óptico y la dirección del robot se obtiene de la conducción manual del robot.

La obtención del flujo óptico se hace desde la captura de video del recorrido del robot por la pista de pruebas. Ya que la respuesta de la red neuronal sobre el robot es refleja solo sobre un eje de movimiento, solo importa rescatar desde las imágenes del

video la posición de la información de forma horizontal sin importar que tan arriba o abajo se encuentre dentro de la imagen, ya que sin importar la variación de este último parámetro la respuesta de red neuronal debería ser la misma. Siguiendo con este tópico y los alcances de las pruebas, no se necesita la información del flujo óptico que se registre de forma vertical, solo tiene importancia para este caso el flujo que se registre en forma horizontal. Esto permite simplificar los datos obtenidos desde las imágenes reduciendo la dimensión de los vectores obtenidos, de dos dimensiones a solo una.

Otra consideración en la obtención del flujo óptico es sobre la pista de pruebas, la cual entregara una imagen que es monótona de forma vertical [mejor descripción?], es decir, si se analiza la imagen de forma vertical los datos que se obtendrán serán muy similares, lo que conlleva a poder reducir los datos que se necesitan obtener sin reducir la información que generaliza a toda la imagen, por lo que solo basta con tomar una franja horizontal de la imagen para tener la información necesaria para la entrada de la red neuronal.

Originalmente la representación del flujo óptico consta de un vector bidimensional por cada punto que se examine de la imagen, esta información se podría representar por dos vectores de dimensión n , donde n es la cantidad de puntos observados en la imagen, y un vector representaría la información del flujo óptico horizontal y otro vertical. Dada las consideraciones mencionadas anteriormente solo se necesita utilizar como vector de entrada para la red neuronal el que contenga la información del flujo óptico en el eje horizontal de la imagen, además la dimensionalidad de este se verá disminuida por la reducción de puntos a tomar en cuenta en la imagen.

La salida esperada de la red neuronal, que representa la dirección en la que se mueve el robot en forma horizontal y que será utilizada como parámetro para el entrenamiento por backpropagation, será obtenida por el seguimiento de los movimientos que realice el robot por medio de un sensor óptico de movimiento. La arquitectura de la red neuronal tendrá una sola salida la que estará normalizada [rango de la salida], por lo que la dirección del robot será representada por un solo valor.

Para lograr que la red neuronal aprenda se utiliza un entrenamiento del tipo *on-*

the-fly, el cual consiste en que la red logre imitar el comportamiento que tendría un humano al realizar la misma tarea, se debe entregar a la red un set de datos generados, en este caso, de la conducción manual del robot por la pista de pruebas, este set de datos consta del flujo óptico obtenido de la captura de video, que se utilizara como entrada de red neuronal, y de la dirección en la cual se movió el robot [agregar marco de referencia]

3.3. Robot omnidireccional

El diseño del robot está basado en uno del repositorio de diseños 3D thingiverse (<https://www.thingiverse.com/thing:167923>), de este se tomó el concepto de las ruedas omnidireccionales y se rediseñaron, la construcción consta de dos ruedas paralelas rotadas sobre el eje de giro 45° una de la otra, unidas por medio de tres tornillos. Cada una de estas ruedas cuenta a su vez con cuatro ruedas o rodillos en su circunferencia colocadas perpendicularmente separadas cada una por 90° , esto permite tener un solapamiento con los rodillos de la otra rueda ubicada perpendicular a esta. Los rodillos van fijos a un eje de 2mm que los atraviesa y puede girar libremente en el soporte de la rueda. Cada rueda tiene un diámetro de 80mm y un ancho de 40mm, en la figura 3.1 se puede ver la construcción de esta.

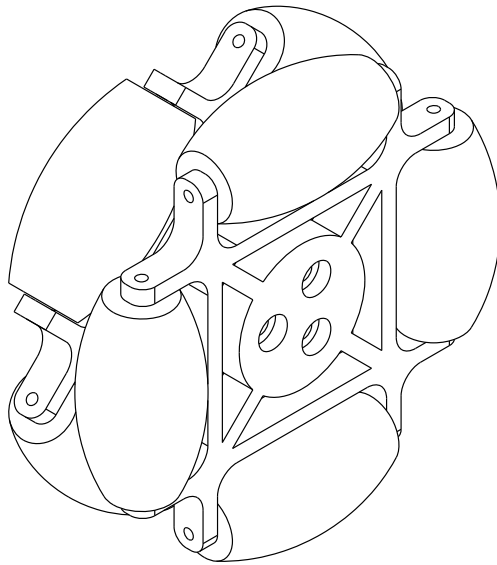


Figura 3.1: Diseño de rueda omnidireccional.

Cada rueda cuenta con un acople, que en un extremo va atornillado a la rueda por tres tornillos y en el otro va sujeto a presión al eje de la caja reductora del motor.

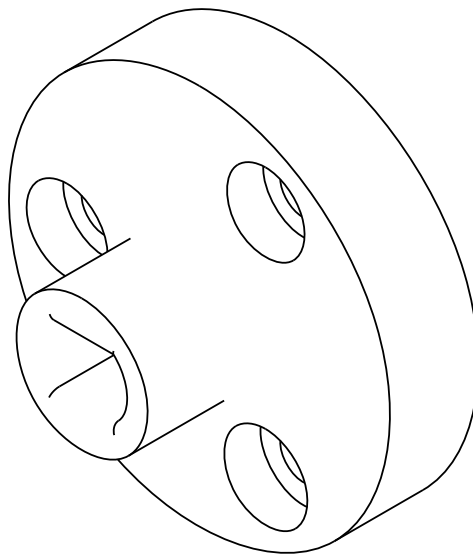


Figura 3.2: Acople para unir las ruedas del robot con la caja reductora de los motores.

A su vez los motores y su respectiva caja reductora están unidos por una ex-

tensión que permite unirlos al cuerpo principal, de esta forma se puede modificar el diámetro total del robot sin tener que modificar también el cuerpo principal, 3.3

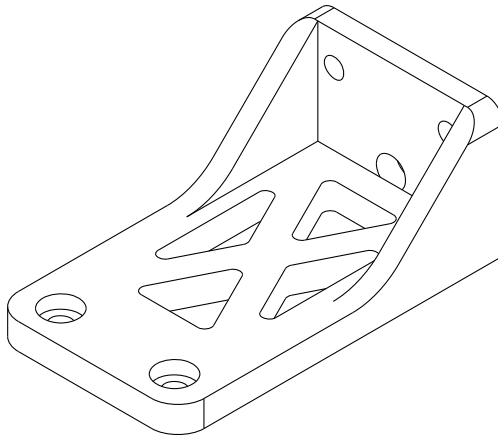


Figura 3.3: Extension para unir los motores.

En el cuerpo principal se posicionan el resto de los componentes necesarios para el funcionamiento, estos son el microcontrolador, los *driver* para los motores, las baterías, el receptor de radio, la cámara digital y sensor de movimiento. El cuerpo fue diseñado de acuerdo a las dimensiones de cada componente, de esta manera se pueden colocar de manera precisa, además de usar las fijaciones para las cuales fueron diseñados, que se muestra en la figura 3.4.

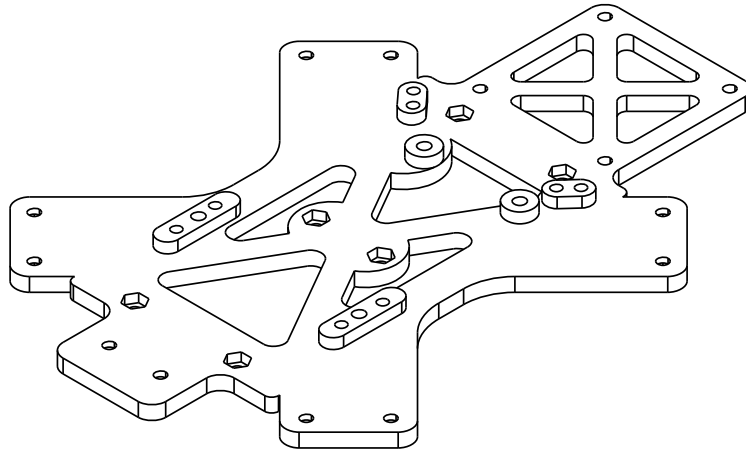


Figura 3.4: Cuerpo principal donde se alojan todos los componentes del robot omnidireccional.

El compartimiento para las baterías es ubicado debajo de cuerpo principal, este permite la fijación de baterías de litio estándar 18650 colocadas en serie, permite colocar un total de 3 baterías. El diseño se puede ver en la figura 3.5.

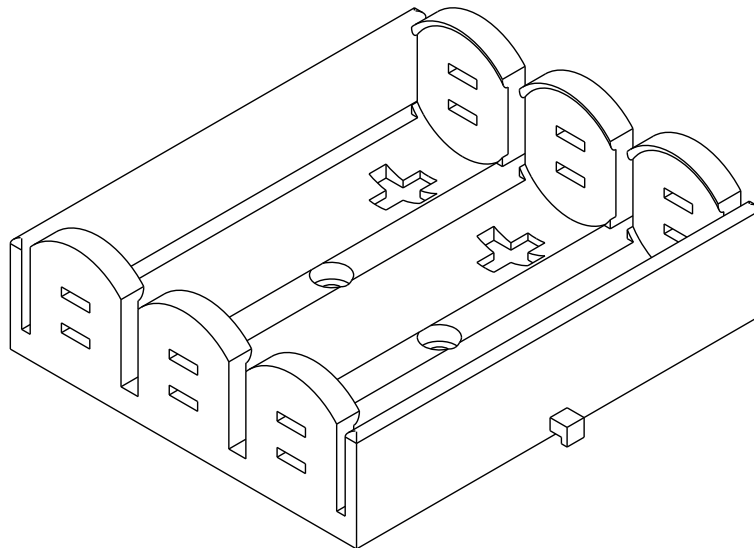


Figura 3.5: Compartimiento para colocar en serie las tres baterías de litio 18650.

Debajo de este, en contacto con la superficie donde se desplaza el robot, va el soporte del sensor óptico de movimiento, en la figura 3.6 se puede ver que el soporte

cuenta con perforaciones por las que se atornilla al cuerpo principal por medio de separadores, para ajustar la altura del soporte para que quede posicionado de forma precisa sobre la superficie.

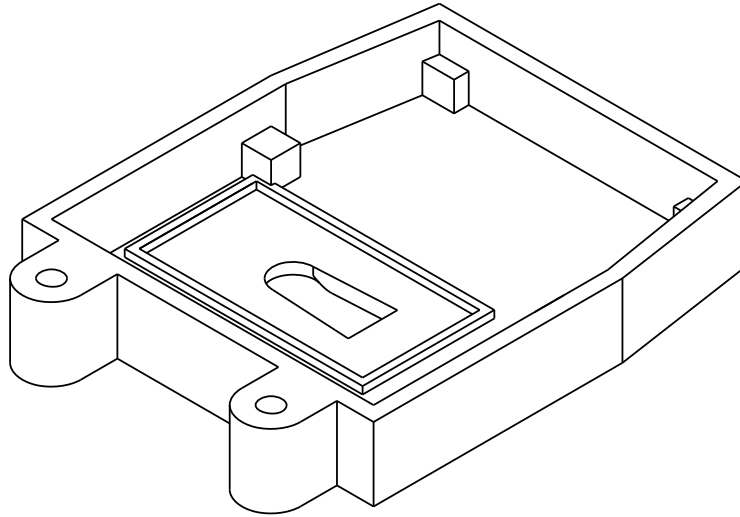


Figura 3.6: Soporte para la instalación del sensor óptico que registra las trayectorias realizadas por el robot.

Al frente del robot va instalada la cámara digital, esta va sobre un soporte que se ajusta a sus dimensiones, el que a su vez va unido al cuerpo principal, en la figura 3.7 se muestra el diseño, que incluye una ranura para posicionar el cable de la cámara.

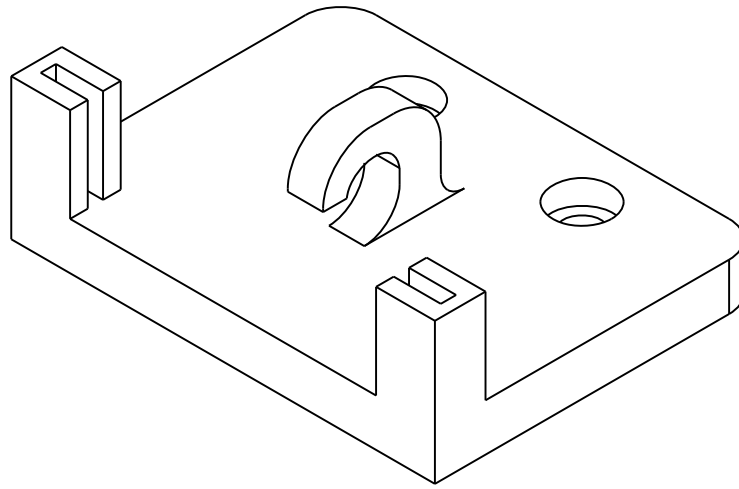


Figura 3.7: Soporte sobre el cual va la cámara digital.

El ensamble completo con todos los componentes diseñados para la construcción del robot se puede ver en la figura 3.8. Todos los acoples mecánicos fueron realizados con tornillos de métrica 3 (M3), cordatos una medida específica para cada unión.

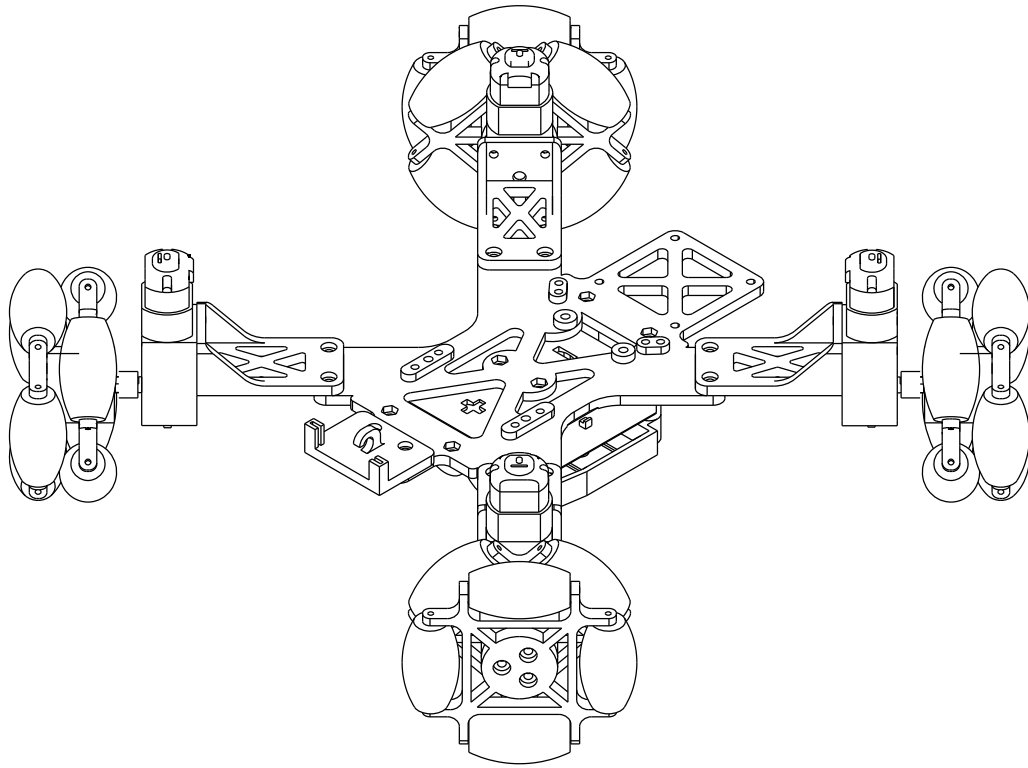


Figura 3.8: Ensamble robot omnidireccional.

Los componentes mencionados y diseñados específicamente para el robot (ruedas omnidireccionales, extensiones para los motores, cuerpo principal, compartimiento para baterías, soporte para la cámara digital y soporte para el sensor óptico) fueron diseñados con la herramienta FreeCAD, un modelador 3D paramétrico. Para el proceso de fabricación se utiliza Sli3r, que convierte los modelos 3D a instrucciones para la impresora 3D. Las piezas son construidas con impresión 3D de filamento fundido (FFF del término en inglés: *Fused Filament Fabrication*) con las siguientes especificaciones:

- Material: PLA 1.75mm
- Diámetro boquilla: 0.4mm
- Altura de capa: 0.2mm
- Temperatura de boquilla: 220°

- Temperatura placa: 60°
- Velocidad impresion: 30mm/s
- Velocidad recorrido: 60mm/s
- Relleno: 15 %
- Patron relleno: rectilíneo

La especificación completa de la configuración de la impresora se puede ver en el anexo A

El desplazamiento del robot es producido por cuatro motores, estos no van en conexión directa con las ruedas, si no que por medio de una caja reductora y esta transmite el movimiento del motor a las ruedas. Los motores son de tamaño estándar 130 alimentados con 12v de corriente continua (DC) con una caja reductora TT a 90°, de relación 120:1, como se muestra en la figura 3.9.

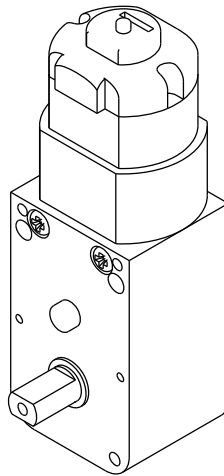


Figura 3.9: Motor del robot omnidireccional.

Los motores son controlados por el *driver* (controlador) l298n, este cuenta con dos puentes H, uno por motor, por lo que el robot utiliza dos controladores. Los controladores son alimentados con 12v y estos son los que suministran la energía a los motores. La fuente de energía del robot es proporcionada por baterías, se utilizan tres baterías de litio 18650 conectadas en serie, cada batería es de 3,7v, por lo que

en total en conjunto se tiene un voltaje nominal de 11.1v, con las baterías al 100 % de su capacidad un voltaje de 12,6v.

Los movimientos del robot son realizados posteriormente al cálculo de las velocidades y dirección de giro de cada rueda, esta tarea es realizada por el microcontrolador, el utilizado es un ATmega328 bajo la infraestructura de Arduino. El microcontrolador toma como entrada la velocidad y dirección a la cual debe moverse y realiza los cálculos para obtener los valores asociados a cada motor. Las salidas de control que van dirigidas a los controladores de los motores y son 3 por cada motor, dos para controlar la dirección del motor y la tercera para la velocidad, así se tiene un total de 12 salidas de control para todos los motores.

4. Pruebas

Las pruebas que se realizan para comprobar si la red neuronal es capaz de evadir obstáculos se dividen en dos etapas: La primera consta en mover el robot sobre la pista de pruebas sin colocar obstáculos, solo con información visual a los costados, la que será captada como flujo óptico, esta permitirá que el robot evite los bordes de la pista y sea capaz de centrarse dentro de esta. En la segunda etapa se mantendrá el esquema de la primera, además de incluir un obstáculo en la pista de pruebas, que el robot tendrá que evadir.

Por cada una de las etapas se comprobará, de dos formas distintas, si la red neuronal está aprendiendo por medio del método de entrenamiento *backpropagation*. La primera es evaluando el aprendizaje con un set de datos distinto al del entrenamiento, aquí se evalúa cuanto difiere la salida esperada con la que entrega la red. Esta diferencia se considera *test error* (error de prueba) y esta será obtenida por la técnica de *Validation Set Approach* (enfoque de conjunto de validación) la que consiste en dejar parte del conjunto de entrenamiento para validar cuan efectivo fue este.

[completar cuantos datos de entrenamiento fueron y cuantos se dejaron para validation set approach] la implementación de la red reordena de forma aleatoria el conjunto de datos y desde este crea un conjunto de datos para las pruebas.

La cuantificación del error se realiza midiendo el porcentaje de efectividad de la red, contando la cantidad de veces que la salida de la red es una respuesta esperada, para saber si la salida satisface esta condición se establece un margen de aceptación de [porcentaje de aceptación], para qué se considere un entrenamiento efectivo la

eficacia de la red debe estar sobre un [porcentaje de efectividad]

La segunda comprobación de entrenamiento se realizará de forma empírica, comparando el comportamiento de la conducción manual del robot, contra la conducción autónoma guiada por la red neuronal, cada vez que el robot se desplace por la pista de pruebas se deja registro de su desplazamiento, por lo que se puede comparar el recorrido realizado de forma manual y el de forma autónoma. Para considerar la prueba exitosa la desviación de las dos trayectorias no debe ser mayor a [medida de desviación].

La obtención del conjunto de datos, para realizar las pruebas, consta de conducir de forma manual el robot repetidas veces por la pista de pruebas, colocando el robot en posiciones definidas al comienzo de la pista y dirigiéndolo hasta el final de esta.

Para la primera etapa, sin importar la posición de inicio, tiene como objetivo mover el robot al centro de la pista direccionándolo solo de forma horizontal [mejorar descripción del movimiento de robot, para que de esta forma la red neuronal pueda aprender a direccionar el robot evadiendo los bordes de la pista].

En la segunda etapa se posicionará un obstáculo dentro de la pista, al igual que en la primera etapa, también se deberá centrar el robot dentro de la pista, pero al enfrentarse a un obstáculo este debe ser evadido.

Para realizar la recopilación de datos, el inicio de la pista será dividido en [n] secciones iguales la que estarán divididas en [cm] una de otra, de donde comenzará el robot, por cada una de estas secciones se registrarán [n] recorridos del robot sobre la pista, obteniendo la información de flujo óptico y el recorrido echo por este. El recorrido registrado será a una velocidad constante [velocidad del robot] y recorrerá una distancia de [distancia a recorrer].

Para la recopilación de datos con obstáculos, se utilizará el mismo esquema anterior, pero se posicionará un obstáculo dentro de la pista en distintas posiciones predefinidas, estas posiciones son representadas dentro de la pista por una matriz, la cual será de [n] filas y [m] columnas, la separación entre las filas y las columnas será igual y de

[cm], y estará posicionada en la pista a [distancia] del inicio de la pista y centrada con respecto a los bordes. Por cada obstáculo se recopilarán los datos [n] veces por cada posición de inicio de la pista.

La construcción de la pista de pruebas necesita de un material que proporcione tracción a las ruedas del robot, por lo que se utiliza una alfombra de cubre pisos de 1 metro de ancho y 4 metros de largo, de color beige para proporcionar mayor contraste frente a los obstáculos, los bordes de la pista tendrán un alto de [cm] con un patrón regular líneas verticales blancas y negras de [cm] de ancho.

5. Resultados

Glosario

Ángulo de ataque:

Caja reductora:

Circuito en serie:

Filtro gaussiano:

Grado de libertad:

Holonómico:

Impresión 3D:

LED:

Modelado Paramétrico:

Motor DC:

Open source:

Overflow:

Paso de hélice:

Puente h:

Robot:

Test Error:

Validation Set Approach:

Voltaje nominal:

Bibliografía

- [1] Opencv: Optical flow, 2017.
- [2] Clay Allen. Quadcopters, 2014.
- [3] Dana Ballard. *Computer vision*. Prentice-Hall, Englewood Cliffs, N.J, 1982.
- [4] Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer vision with the OpenCV library*. .o'Reilly Media, Inc.”, 2008.
- [5] Myriam Cayre, Jordane Malaterre, Sophie Scotto-Lomassese, Colette Strambi, and Alain Strambi. The common properties of neurogenesis in the adult brain: from invertebrates to vertebrates. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 132(1):1–15, 2002.
- [6] Adam Chapweske. Ps/2 mouse interfacing, 2003.
- [7] Adam Chapweske. The ps/2 mouse/keyboard protocol, 2003.
- [8] Sanja Filder. Image pyramid.
- [9] David Forsyth. *Computer vision : a modern approach*. Prentice Hall, Upper Saddle River, N.J. London, 2003.
- [10] Berthold Horn. *Robot vision*. MIT press, 1986.
- [11] Munasinghe Muhandiram Premarathna Jayakody, Munagamage and Rajakaruna. Design and development of an omni-directional wheeled mobile robot. *International Research Symposium on Engineering Advancements 2015 (RSEA 2015)*, 2015.

- [12] Reinhard Klette. *Concise computer vision : an introduction into theory and algorithms*. Springer, London, 2014.
- [13] Tim Morris. *Computer vision and image processing*. Palgrave Macmillan, Basingstoke, 2004.
- [14] Michael Nielsen. Using neural nets to recognize handwritten digits, 2017.
- [15] International Civil Aviation Organization. Unmanned aircraft systems (uas), 2011.
- [16] Linda Shapiro. *Computer vision*. Prentice Hall, Upper Saddle River, NJ, 2001.
- [17] Charles Siskind. *Electrical control systems in industry*. McGraw-Hill, New York, 1963.
- [18] Milan Sonka. *Image processing, analysis, and machine vision*. Thompson Learning, Toronto, 2008.
- [19] C. E. Vandoni. *Proceedings : 1996 CERN School of Computing : Egmond aan Zee, the Netherlands, 8 September-21 September 1996*. CERN, Geneva, 1996.
- [20] Al Williams. *Microcontroller projects using the Basic Stamp*. CMP Books, Lawrence, Kan, 2002.

ANEXOS

A. Configuración impresora 3D

Configuración completa de la impresora 3D al momento de construir las piezas para el robot omnidireccional.

```
# generated by Slic3r 1.2.9 on Tue Jun 20 23:36:24 2017
avoid_crossing_perimeters = 0
bed_shape = 0x0,200x0,200x145,0x145
bed_temperature = 60
before_layer_gcode =
bottom_solid_layers = 3
bridge_acceleration = 0
bridge_fan_speed = 100
bridge_flow_ratio = 1
bridge_speed = 30
brim_width = 0
complete_objects = 0
cooling = 1
default_acceleration = 0
disable_fan_first_layers = 3
dont_support_bridges = 1
duplicate_distance = 6
end_gcode = M104 S0 ; turn off temperature\nG28 X0
; home X axis\nG0 Y149; \nM84 ; disable motors\n
external_fill_pattern = rectilinear
external_perimeter_extrusion_width = 0
```



```
external_perimeter_speed = 80%
external_perimeters_first = 0
extra_perimeters = 1
extruder_clearance_height = 20
extruder_clearance_radius = 20
extruder_offset = 0x0
extrusion_axis = E
extrusion_multiplier = 1
extrusion_width = 0
fan_always_on = 1
fan_below_layer_time = 0
filament_colour = #000000
filament_diameter = 1.75
fill_angle = 45
fill_density = 15%
fill_pattern = rectilinear
first_layer_acceleration = 0
first_layer_bed_temperature = 60
first_layer_extrusion_width = 250%
first_layer_height = 0.4
first_layer_speed = 50%
first_layer_temperature = 215
gap_fill_speed = 15
gcode_arcs = 0
gcode_comments = 0
gcode_flavor = reprap
infill_acceleration = 0
infill_every_layers = 1
infill_extruder = 1
infill_extrusion_width = 0
infill_first = 0
infill_only_where_needed = 0
infill_overlap = 15%
infill_speed = 40
```

```
interface_shells = 0
layer_gcode =
layer_height = 0.2
max_fan_speed = 100
max_print_speed = 80
max_volumetric_speed = 0
min_fan_speed = 70
min_print_speed = 10
min_skirt_length = 0
notes =
nozzle_diameter = 0.4
octoprint_apikey =
octoprint_host =
only_retract_when_crossing_perimeters = 1
ooze_prevention = 0
output_filename_format = [input_filename_base].gcode
overhangs = 1
perimeter_acceleration = 0
perimeter_extruder = 1
perimeter_extrusion_width = 0
perimeter_speed = 30
perimeters = 3
post_process =
pressure_advance = 0
raft_layers = 0
resolution = 0
retract_before_travel = 2
retract_layer_change = 1
retract_length = 5
retract_length_toolchange = 10
retract_lift = 0
retract_restart_extra = 0
retract_restart_extra_toolchange = 0
retract_speed = 40
```

```
seam_position = aligned
skirt_distance = 6
skirt_height = 1
skirts = 2
slowdown_below_layer_time = 20
small_perimeter_speed = 15
solid_infill_below_area = 70
solid_infill_every_layers = 0
solid_infill_extruder = 1
solid_infill_extrusion_width = 0
solid_infill_speed = 80%
spiral_vase = 0
standby_temperature_delta = -5
start_gcode = G28 ; home all axes\n;G1 Z5 F5000 ; lift nozzle\n
support_material = 0
support_material_angle = 0
support_material_contact_distance = 0.2
support_material_enforce_layers = 0
support_material_extruder = 1
support_material_extrusion_width = 0
support_material_interface_extruder = 1
support_material_interface_layers = 3
support_material_interface_spacing = 0
support_material_interface_speed = 100%
support_material_pattern = pillars
support_material_spacing = 2.5
support_material_speed = 40
support_material_threshold = 0
temperature = 210
thin_walls = 1
threads = 2
toolchange_gcode =
top_infill_extrusion_width = 0
top_solid_infill_speed = 50%
```

```
top_solid_layers = 3
travel_speed = 60
use_firmware_retraction = 0
use_relative_e_distances = 0
use_volumetric_e = 0
vibration_limit = 0
wipe = 0
xy_size_compensation = 0
z_offset = 0
```

A.1. La primera sección del primer anexo

Aquí va el texto de la primera sección del primer anexo...

A.2. La segunda sección del primer anexo

Aquí va el texto de la segunda sección del primer anexo...

A.2.1. La primera subsección de la segunda sección del primer anexo

B. El segundo Anexo

Aquí va el texto del segundo anexo...

B.1. La primera sección del segundo anexo

Aquí va el texto de la primera sección del segundo anexo...