

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería



**Implementación de un sistema de visión por computadora
integrable como módulo de ROS a una plataforma robótica
móvil**

Protocolo de trabajo de graduación presentado por María Alejandra
Samayoa Gómez, estudiante de Ingeniería Mecatrónica

Guatemala,

2022

Resumen

Uno de los proyectos que se estará trabajando dentro de la Universidad del Valle de Guatemala en el año 2022 será una plataforma robótica móvil. El siguiente trabajo consiste en la implementación de un sistema de visión por computadora a esta plataforma. Esto se realizará por medio de una cámara Kinect y otro módulo que se defina como el más adecuado entre otras dos opciones. Para esto se realizarán varias pruebas y se le asignará una puntuación a cada módulo para determinar cuál será implementado. Ambos módulos serán adaptados a nodos de ROS. En este medio serán integrados a la plataforma robótica. Se realizarán pruebas básicas con algoritmos de visión por computadora para comprobar su funcionamiento adecuado.

Antecedentes

El siguiente trabajo busca expandir algunos trabajos realizados dentro de la Universidad del Valle de Guatemala e implementar la información de otras investigaciones realizadas en diferentes partes del mundo. En otros trabajos de investigación del exterior se han realizado estudios sobre la implementación de la cámara Kinect para realizar control de visión por computadora. Dentro de la universidad, se han trabajado proyectos de visión por computadora y de la plataforma robótica móvil. A continuación se describen los antecedentes de este proyecto.

Cámara Kinect para visión por computadora

La cámara Microsoft Kinect ha probado ser una herramienta útil y de bajo costo para realizar visión por computadora para la robótica. En el artículo *Enhanced Computer Vision With Microsoft Kinect Sensor: A Review*, los autores realizan una guía con diferentes algoritmos que se pueden utilizar para realizar la visión por computadora a través del módulo de Kinect [1]. Se describen las funciones de reconocimiento de posición, mapeo y de cámara 3D que se pueden lograr con el dispositivo.

En la tesis *Evaluation of Microsoft Kinect 360 and Microsoft Kinect One for robotics and computer vision applications* por Simone Zennaro [2], se logró realizar diferentes pruebas de visión específicamente para aplicarse para el control por visión para una plataforma robótica y comparar los módulos de la cámara Kinect 360 y Kinect One. La intención de este robot fue que lograra seguir a una persona por medio de un sistema de control basado en visión. En el trabajo, desarrolla diferentes algoritmos para poder comparar reconocimiento facial, detección de contornos y percepción de profundidad. Se concluyó que ambas pueden ser implementadas, pero que la cámara de Kinect 360 funciona mejor en interiores y que la cámara Kinect One tiene mejor detalle y funcionamiento en el exterior por tener mejor contraste.

En la Figura 1 se muestra la plataforma robótica en la que se probaron las cámaras.

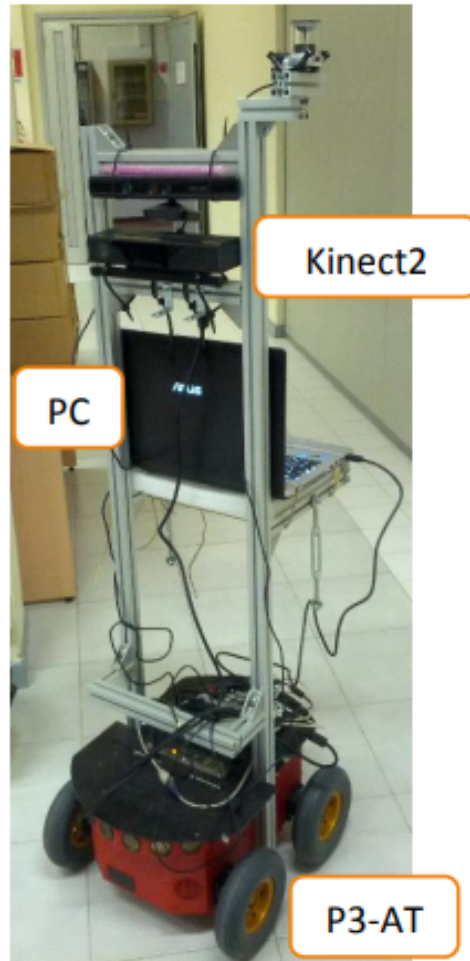


Figura 1: Robot móvil sobre el cuál se le aplicó el sistema de control basado en visión de la cámara Kinect [2].

Plataforma Móvil

Se han realizado diferentes iteraciones de un robot explorador en la Universidad del Valle a través de los años. La versión más reciente del robot se trabajó en el año 2021, donde se buscó trabajar en el desarrollo y mejora de la plataforma que no había tenido seguimiento desde el 2018. A través del trabajo de graduación de Héctor Sagastume [3] y de Javier Archila [4], se buscó obtener una plataforma que funcionaría a control remoto y contara con algunos sensores para su funcionamiento óptimo.

Sagastume cambió el sistema de transmisión de potencia, le agregó sensores infrarojos y de temperatura y realizó el diseño de la carcasa exterior. Por otro lado, Archila se enfocó en realizar el control del robot para que funcionara en modo de auto piloto. Para la lectura de los sensores y el sistema de auto piloto, se utilizó una Raspberry Pi3 como computadora. Se obtuvo el robot que se muestra en la Figura 2 que funcionó a control remoto por medio de conexión al internet.



Figura 2: Robot explorador que servirá como plataforma móvil [3].

Este robot explorador no contó con un sistema de visión avanzado. Se utilizó un módulo de cámara USB que se conectó a la Raspberry Pi3 y el video se mandaba a través de la conexión de bluetooth que se realizó. Sin embargo, el video no formó parte del control que se realizaba para el movimiento del robot.

Visión por computadora

La visión por computadora no es un tema nuevo dentro de la universidad. Ya que es un tema con bastantes aplicaciones, han existido varios trabajos de graduación que la integran a otro proyecto, hacen una comparación entre diferentes módulos disponibles o continúan afinando métodos para realizar la visión por computadora. Entre los trabajos más recientes destacan algunos que serán de importancia para la aplicación que se le quiere dar a la visión por computadora en este proyecto.

En el año 2020, José Guerra, en su trabajo de graduación, utilizó la visión por computación aplicada a la robótica de enjambre [5]. Con esto, se buscó identificar las poses de los miembros del enjambre para tener un control de la posición de estos sobre una mesa de prueba. Además, realizó una comparación breve entre la programación orientada a objetos y el método multihilos para obtener la mejor visión. Esto lo realizó utilizando la plataforma Open CV y luego migrando la programación a Python.

José Ramírez continuó el estudio en el año 2021 [6], utilizando como base el trabajo realizado por Guerra. La aplicación en este caso buscaba detectar los obstáculos dentro del entorno y se buscaba crear una aplicación en la plataforma de Matlab que lograra el mapeo del área sobre la mesa de prueba Robotat. Se logró la migración de los algoritmos a Matlab y el mapeo del área sobre la mesa.

Otro trabajo de visión por computadora que será útil para el proyecto es el estudio realizado por Héctor Klée en el año 2021 [7]. Este consistió en un estudio comparativo entre diferentes módulos de visión para un sistema embebido: JeVois Smart Vision y ESP32-CAM. Se implementó en un pequeño brazo robótico, con la intención de que se realizara un sistema de control basado en visión para que el brazo siguiera a un marcador. Se realizaron

8 diferentes pruebas con ambos módulos y se les otorgó una calificación según su desempeño ante estas pruebas. Al final, se decidió utilizar el sistema JeVois y la plataforma Open CV ya que este tuvo mejor desempeño. Además, se logró realizar el controlador para el brazo, utilizando la visión para reconocer el marcador específico.

Los trabajos de Ramírez y Guerra fueron limitados por la pandemia para realizar la cantidad de pruebas que se querían realizar, y sus aplicaciones consistieron en aplicaciones estacionarias. El trabajo de Klée, aunque se realizó para una aplicación móvil, se enfocó únicamente en la detección de un tipo de marcador y se aplicó la visión por computadora con un solo módulo de cámara.

Justificación

La visión por computadora ha sido un campo creciente en los últimos años, demostrando su versatilidad y utilidad a través de diversas aplicaciones que se le han dado. Uno de los campos en dónde se ha demostrado su utilidad es en el campo de la robótica, dónde se han generado aplicaciones para reconocer obstáculos, mejorar la interacción con los seres humanos y para realizar sistemas de control para el robot.

Este año se creará una plataforma robótica móvil, Rover UVG, que será un compendio de diferentes módulos sobre los cuales se podran realizar diferentes pruebas y se juntan diferentes aplicaciones de la tecnología. Parte de estos módulos será la visión por computadora que se le agregará con este trabajo de graduación. La plataforma que se había trabajado con anterioridad, [4], no contaba con un sistema de visión por computadora avanzado ya que dependía más de sus diferentes sensores que de la cámara que tenía implementada. Se busca mejorar la visión por computadora que tendrá la plataforma. En la universidad anteriormente se han trabajado implementaciones de visión por computadora a aplicaciones estáticas, tal como las mesas de prueba trabajadas en [6]. Además, se han realizado controladores a base de la visión por computadora en aplicaciones como un pequeño brazo robótico al igual que una comparación breve entre algunos módulos de cámaras, como fue el caso en [7].

El trabajo propuesto busca agregar una parte importante a la plataforma que se estará trabajando este año, implementando un sistema de visión por computadora más completo. Se realizarán más pruebas en esta plataforma móvil, algo que se limitó en los trabajos anteriores. Este proyecto consistirá en la aplicación de dos módulos de cámara para la visión por computadora, con la intención de maximizar las características de ambas.

A través de la comparación de los módulos Raspberry Cam y JeVois, se realizarán diferentes pruebas que detallen su funcionamiento. Por otro lado, la comparación que se realizará entre los tipos de algoritmos de implementación determinará las ventajas y desventajas que tienen los dos métodos. El módulo y el algoritmo que se determinen más eficientes serán implementados como módulo de ROS en el proyecto. Además, se le agregará la cámara Kinect como módulo de ROS. Estos se implementarán en el Rover, permitiendo así realizar un sistema de control más detallado y agregándole un nivel de independencia al robot. Como nodos en ROS, el sistema de visión se podría implementar en diferentes proyectos de robótica en el futuro.

Objetivos

Objetivo General

Implementar un sistema de visión por computadora que pueda integrarse como módulo de ROS a una plataforma robótica móvil.

Objetivos Específicos

- Evaluar y comparar diferentes módulos de cámara para definir el más adecuado para la plataforma robótica móvil.
- Adaptar el sistema de visión por computadora con la cámara seleccionada como módulo de ROS, para su integración a la plataforma robótica móvil.
- Adaptar la cámara Kinect como módulo de ROS, para su integración a la plataforma robótica móvil.
- Realizar pruebas básicas de algoritmos de visión por computadora para validar el sistema implementado.

Marco teórico

Visión por computadora

Según [8] la visión por computadora se puede definir como un campo de la inteligencia artificial que es un compendio de algoritmos y técnicas que mezclan un módulo de visión físico con la interpretación de la imagen por medio de programación y diferentes procesamientos que se le realizan a la imagen [8]. Este campo tiene una gran cantidad de aplicaciones en diferentes campos, en especial en el campo de la robótica. Es un campo que ha ido cambiando en grandes cantidades a lo largo de la historia, como se puede observar en la Figura 3.

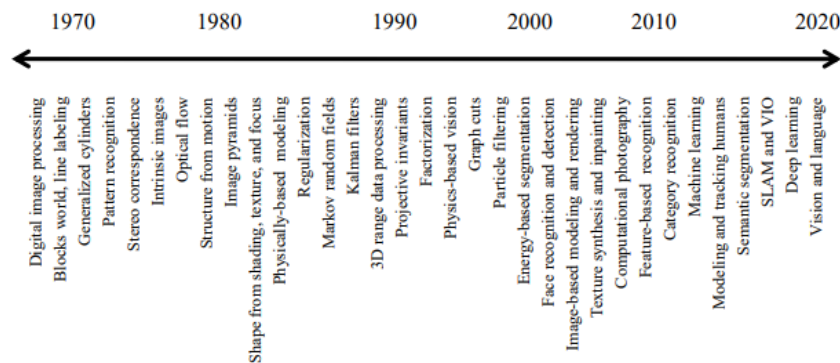


Figura 3: Historia de la visión por computadora, extraído de: [9].

Visión por computadora clásica

En [9], se detalla el proceso de visión por computadora. La mayoría de cámaras realizan un proceso de capturar imágenes en 3 dimensiones y realizar la conversión correspondiente para obtener una imagen en 2. A través de puntos y líneas en dos dimensiones se va formando, pixel por pixel la imagen digital que replica la figura en 3D. Una gran parte de la industria de visión por computadora se basa en mejorar la calidad de la imagen extraída a través de diferentes algoritmos y filtros [9].

Visión por computadora en base a Aprendizaje Automático

La combinación de los campos de visión por computadora y el aprendizaje automático es algo que ha ido evolucionando en los últimos años y sigue en crecimiento. En esencia, los algoritmos de visión por computadora a base de aprendizaje automático tienen la intención de obtener datos y utilizar estos dentro de modelos para realizar predicciones [10]. En la industria se pueden ver en aplicaciones de reconocimiento facial, análisis de texto y clasificación de imágenes, entre otras. Este tipo de visión ha permitido aplicaciones de controladores de visión más complejas, como las que se manejan en los vehículos autónomos que reconocen diferentes señales, obstáculos y comandos y utilizan estas imágenes para reaccionar de la manera correcta.

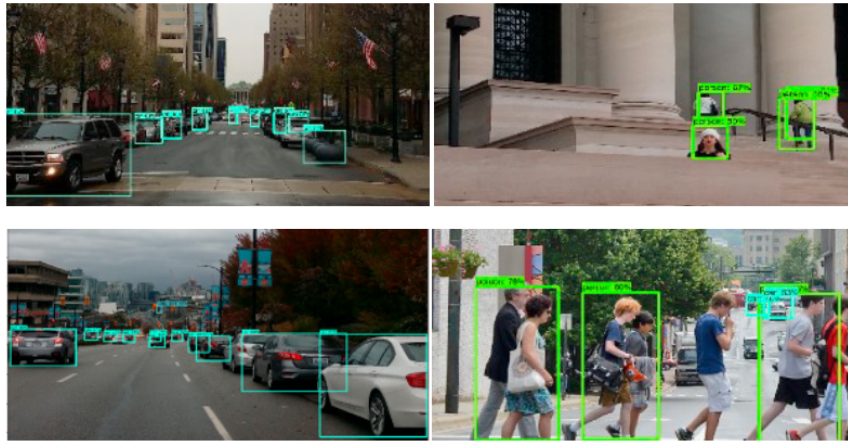


Figura 4: Ejemplo vision por computadora utilizando aprendizaje automatico para reconocimiento [10].

Software utilizado para visión por computadora

El campo de visión por computadora, al ser bastante amplio, cuenta con una gran cantidad de programas sobre los que se pueden trabajar. Entre estos, uno que es bastante común es OpenCV (*Open Source Computer Vision Library*) es una librería que se utiliza para realizar visión por computadora en tiempo real [11]. Estas se pueden manejar dentro de diferentes plataformas como Python y Matlab. Se ha utilizado para realizar la mayoría de funciones de visión por computadora y permite la conexión de varios módulos de visión. Varios módulos de cámara se pueden conectar a través de OpenCV. Programas como Matlab cuentan

también con una gran habilidad de procesamiento de imágenes, ya que tienen funciones especializadas para la visión.

Módulos de Cámaras

Jevois

El módulo JeVois es un módulo de cámara compatible con proyectos de PC, Arduino o Raspberry Pi y es un módulo compatible con diferentes algoritmos de *Open source* que se utilizan para la visión por computadora [12]. Entre sus considerables ventajas están su tamaño compacto, su versatilidad de implementación y las diferentes aplicaciones que se le han dado para realizar visión por computadora en otros proyectos. Cuenta además con un sensor de video, procesador de cuatro núcleos, puerto USB y puertos seriales [12].

Raspberry Pi Cámaras

La empresa de Raspberry Pi tiene en el mercado un módulo de cámara conocido como el Raspicam [13]. Este es un módulo compatible con diferentes modelos de las Raspberry Pi y con extensas librerías para su funcionamiento. Cuenta con la habilidad de tomar fotos y videos en definición alta. Es un módulo pequeño, pero cuenta con un sensor de visión, un cable FPC y diferentes montaduras [13].

Kinect

El módulo de visión Kinect es un módulo de detección de movimiento de la compañía Microsoft que se utilizaba originalmente para la consola Xbox. Sin embargo, en los últimos años se ha convertido en una opción de bajo costo para realizar visión por computadora gracias a su procesamiento de profundidad y su sensor de visión RGB [1].

En los últimos años, se han realizado estudios donde se utiliza el Kinect para aplicaciones de seguimiento de objetos, reconocimiento de posición y gestos, mapeo de áreas, entre otros. La estructura del módulo se muestra a continuación, en la Figura 5.

ROS

ROS, como se conoce por sus siglas en inglés (*Robot Operating System*) es un software de *open source* que cuenta con una gran cantidad de librerías y herramientas que buscan integrar diferentes componentes para el funcionamiento de un robot [14]. El programa funciona al conectar diferentes nodos que representan diferentes actuadores o sensores del robot. ROS permite juntar estos nodos y realizar controladores para el robot de manera más sencilla.

El programa permite igual realizar un gemelo digital lo que permite simular el robot para realizar pruebas al igual que obtener la información en tiempo real de los nodos de los sensores. Se puede conectar a través del programa todo tipo de sensores que se manejen con



Figura 5: Configuración de módulos dentro del Kinect [1].

software. Existen librerías que integran módulos de ROS con software para realizar visión por computadoras, tal como OpenCV.

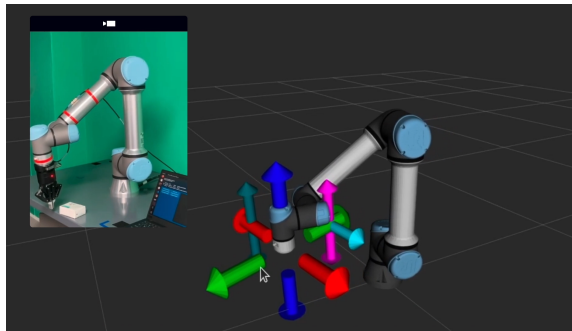


Figura 6: Ejemplo de gemelo digital realizado en ROS [14].

Metodología

Investigación preliminar

Lo primero que se realizará para el proyecto es una investigación a fondo sobre los módulos de cámara, módulos de visión por computadora en ROS y los algoritmos a utilizar para las pruebas y validaciones que se realizarán. Esta parte consiste también en verificar con qué versiones de los módulos se estará trabajando y qué se necesita para la unión de todos los nodos a la plataforma robótica.

Comparación entre módulos

La primera parte del trabajo consiste en la elección del módulo de cámara que se implementará con el Kinect. Entre esto, se estarán comparando los módulos de cámara JeVois y la Raspicam. Una vez se logró la función básica de ambos módulos, empezarán unas series de pruebas para tipos de visión por computadora que sean útiles para la plataforma. A las

cámaras se les estará otorgando una calificación para que la que tenga mejor desempeño sea la que se vaya a utilizar. Para esto será necesario tomar en cuenta el tipo de visión por computadora que realizará este módulo en específico dentro de la plataforma y cómo se estará implementando dentro del robot.

Pruebas con módulo Kinect

Con el módulo Kinect será necesario primero lograr la función básica del módulo. De igual manera se necesitará definir la función específica que cumplirá dentro del robot. Parte de esta etapa con el módulo será también observar bien su posicionamiento en la plataforma robótica y diseñar algún soporte físico si se considera necesario para su posicionamiento adecuado.

Implementación como módulo de ROS

Esta parte del trabajo depende de los requisitos que se definan para la implementación de los diferentes nodos en la plataforma robótica móvil. Sin embargo, se estarán realizando pruebas del funcionamiento de los algoritmos que se definieron en las etapas anteriores del proyecto. Una vez se hayan determinado los requisitos para la unión, se estará verificando que el nodo cumpla con esto y con sus funciones originales.

Pruebas con algoritmos

El siguiente paso consistiría en la investigación a fondo de algunos algoritmos que se pondrán a prueba para comprobar el funcionamiento del nodo en ROS de las cámaras. Según las funciones específicas que se decidan para cada uno de los módulos, se investigarán sobre los algoritmos que cumplan con estas funciones. Con esas pruebas se realizará la evaluación necesaria y se determinarán los algoritmos finales para el nodo.

Cronograma de actividades

		Mes											
		Julio			Agosto				Septiembre				
Fecha		4/07-10/07	11/07-17/07	18/07-24/07	1/08-7/08	8/08-14/08	15/08-21/08	22/08-28/08	29/08-4/09	5/09-11/09	12/09-18/09	19/09-25/09	26/09-30/09
No. Semana		1	2	3	4	5	6	7	8	9	10	11	12
No. Tarea	Actividad	Tiempo de ejecución											
1	Investigación de pruebas a realizar para la comparación de módulos												
2	Investigación de funcionamiento de los módulos de cámaras												
3	Investigación sobre nodos de visión por computadora en ROS												
4	Conexiones iniciales de los módulos de cámara												
5	Pruebas para la comparación entre módulos de cámara												
6	Selección de módulo de cámara a utilizar												
7	Conexiones iniciales del módulo Kinect												
8	Selección de función específica dentro de la plataforma robótica												
9	Verificación de posicionamiento de módulos de cámara en la plataforma y diseño de ajustes si es necesario												
10	Investigación sobre algoritmos de ambos métodos que cumplan con la función específica definida para cada módulo												
11	Pruebas del funcionamiento de nodo de ROS												
12	Selección de algoritmos finales												
13	Integración de algoritmos escogidos como módulo de ROS												
14	Pruebas de algoritmos como módulos de ROS												
15	Verificación de requisitos del nodo de ROS para su integración a la plataforma												
16	Elaboración de trabajo escrito												

Figura 7: Cronograma de actividades.

Índice preliminar

1. Resumen
2. Abstract
3. Introducción
4. Antecedentes
 - 4.1 Cámara Kinect para visión por computadora
 - 4.2 Plataforma móvil
 - 4.3 Visión por computadora
5. Justificación
6. Objetivos
 - 6.1 Objetivo General
 - 6.2 Objetivo Específico
7. Alcance
8. Marco teórico
 - 8.1 Módulos de Cámara
 - 8.2 Visión por Computadora
 - 8.3 ROS
9. Comparación entre módulos
 - 9.1 Pruebas a realizar
 - 9.2 Resultados JeVois
 - 9.3 Resultados RaspiCam
 - 9.4 Selección de módulo
10. Adaptación a ROS
 - 10.1 Algoritmos que se implementarán
 - 10.2 Pruebas del módulo como nodo de ROS
 - 10.3 Resultados de algoritmos adaptado
11. Visión por computadora en la plataforma
 - 11.1 Función específica de ambos módulos de cámara
 - 11.2 Posición mecánica de los módulos en la plataforma
 - 11.3 Resultados de verificación de requisitos para la integración a la plataforma como módulo de ROS
12. Conclusiones

13. Recomendaciones
14. Bibliografía
15. Anexos

Referencias

- [1] J. Han, L. Shao, D. Xu y J. Shotton, “Enhanced Computer Vision With Microsoft Kinect Sensor: A Review,” *IEEE Transactions on Cybernetics*, vol. 43, n.º 5, págs. 1318-1334, 2013. DOI: [10.1109/TCYB.2013.2265378](https://doi.org/10.1109/TCYB.2013.2265378).
- [2] S. Zennaro, “Evaluation of Microsoft Kinect 360 and Microsoft Kinect One for robotics and computer vision applications,” 2014, Tesis de licenciatura.
- [3] H. J. Sagastume, “Diseño Mecánico, Selección de Motores e Implementación de Sensores para un Robot Explorador Modular,” Tesis de licenciatura, Universidad del Valle de Guatemala, 2021.
- [4] J. E. Archila, “Diseño e implementación de capacidades automáticas de navegación para un Robot Explorador Modular,” Tesis de licenciatura, Universidad del Valle de Guatemala, 2021.
- [5] J. Guerra, “Algoritmos de Visión por Computadora para el Reconocimiento de la Pose de Agentes Empleando Programación Orientada a Objetos y Multihilos,” Tesis de licenciatura, Universidad del Valle de Guatemala, 2021.
- [6] J. I. Ramírez, “Herramienta de Software de Visión por Computadora para Aplicaciones de Robótica de Enjambre en una Mesa de Prueba - Fase III,” Tesis de licenciatura, Universidad del Valle de Guatemala, 2021.
- [7] H. A. Klée, “Desarrollo e implementación de algoritmos de visión por computadora clásicos empleando OpenCV en sistemas embebidos,” Tesis de licenciatura, Universidad del Valle de Guatemala, 2021.
- [8] D. Forsyth y J. Ponce, *Computer Vision: A Modern Approach. (Second edition)*. Prentice Hall, nov. de 2011, pág. 792. dirección: <https://hal.inria.fr/hal-01063327>.
- [9] R. Szeliski, *Computer vision: Algorithms and applications*, 2.^a ed. Springer Nature, 2022.
- [10] A. I. Khan y S. Al-Habsi, “Machine Learning in Computer Vision,” *Procedia Computer Science*, vol. 167, págs. 1444-1451, 2020, International Conference on Computational Intelligence and Data Science, ISSN: 1877-0509. DOI: <https://doi.org/10.1016/j.procs.2020.03.355>. dirección: <https://www.sciencedirect.com/science/article/pii/S1877050920308218>.
- [11] OpenCV, *OpenCV modules*, 2021. dirección: <https://docs.opencv.org/4.5.5/>.
- [12] JeVois, *About JeVois*, 2019. dirección: <https://www.jevoisinc.com/pages/what-is-jevois>.
- [13] R. Pi, *Raspberry pi documentation*, 2019. dirección: <https://www.raspberrypi.com/documentation/accessories/camera.html>.
- [14] O. Robotics, *Why Ros?* 2021. dirección: <https://www.ros.org/blog/why-ros/>.