UNIVERSIDAD POLITÉCNICA SALESIANA ECUADOR

POSGRADOS

MAESTRÍA EN
ELECTRÓNICA Y AUTOMATIZACIÓN

RPC-SO-30-NO.507 -2019

OPCIÓN DE TITULACIÓN:

ARTÍCULO PROFESIONAL DE ALTO NIVEL

TEMA:

RECONOCIMIENTO DE OBJETOS MEDIANTE TÉCNICAS DE VISIÓN ARTIFICIAL A TRAVÉS DE CÁMARA Y LIDAR ENFOCADO HACIA LOS VEHÍCULOS AUTÓNOMOS TERRESTRES

AUTORES:

BYRON ISRAEL CHUCARALAO MINCHALA
JONNATHAN BERNARDO JARRO PATIÑO

DIRECTOR:
JUAN DIEGO VALLADOLID QUITOISACA

CUENCA – ECUADOR
2025

Autores:

Byron Israel Chucaralao Minchala.
Ingeniero en Mecatrónica.
Candidato a Magíster en Electrónica y Automatización por la Universidad Politécnica Salesiana - Sede Cuenca.
bchucaralao@est.ups.edu.ec

Jonnathan Bernardo Jarro Patiño.
Ingeniero en Mecatrónica.
Candidato a Magíster en Electrónica y Automatización por la Universidad Politécnica Salesiana - Sede Cuenca.
jjarrop@est.ups.edu.ec

Dirigido por:

Juan Diego Valladolid Quitoisaca
Ingeniero Electrónico.
Magister en Control y Automatización Industriales.
jvalladolid@ups.edu.ec

Todos los derechos reservados.

Queda prohibida, salvo excepción prevista en la Ley, cualquier forma de reproducción, distribución, comunicación pública y transformación de esta obra para fines comerciales, sin contar con autorización de los titulares de propiedad intelectual. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual. Se permite la libre difusión de este texto con fines académicos investigativos por cualquier medio, con la debida notificación a los autores.

DERECHOS RESERVADOS
2025 © Universidad Politécnica Salesiana.
CUENCA – ECUADOR – SUDAMÉRICA
BYRON ISRAEL CHUCARALAO MINCHALA
JONNATHAN BERNARDO JARRO PATIÑO
Reconocimiento de objetos mediante técnicas de visión artificial a través de cámara y lidar enfocado hacia los vehículos autónomos terrestres

Reconocimiento de objetos mediante técnicas de visión artificial a través de cámara y LiDAR enfocado hacia los vehículos autónomos terrestres.

B. Chucaralao and J. Jarro

Universidad Politécnica Salesiana - Ecuador

Resumen

En la actualidad se presenta innovaciones para la transformación del sector automotriz, en busca de la mejora de la seguridad vial y eficiencia para el manejo vial. En el presente artículo tiene como objetivo determinar un método para el reconocimiento de objetos a través de sensores LiDAR y cámaras con enfoque en los vehículos autónomos terrestres, este principio es un pilar fundamental en la mejora para la percepción y tomas de decisiones en la conducción autónoma para los entornos en los que se utilicen. La fusión de los datos de los sensores y con mecanismos de inteligencia artificial mediante un software permiten obtener resultados que pueden ser comparados con otros estudios realizados, debido a la presente competencia en los desarrollos para la conducción autónoma. Las pruebas se realizaron en un ambiente amplio de la Universidad Politécnica Salesiana donde transcurrían personas y existen objetos presentes. Los resultados muestran el reconocimientos de peatones como de otros objetos y la visualización del entorno para determinación de distancias, evidenciando la sinergia entre el sensor con la cámara.

Abstract

Nowadays, innovations are presented for the transformation of the automotive sector, in search of the improvement of road safety and efficiency for road management. This article aims to determine a method for object recognition through LiDAR sensors and cameras with a focus on autonomous land vehicles, this principle is a fundamental pillar in the improvement of perception and decision making in autonomous driving for the environments in which they are used. The fusion of data from sensors and artificial intelligence mechanisms through software allows obtaining results that can be compared with other studies conducted, due to the present competition in the developments for autonomous driving. The tests were conducted in a large environment at the Salesian Polytechnic University where people were passing by and objects were present. The results show the recognition of pedestrians and other objects and the visualization of the environment to determine distances, evidencing the synergy between the sensor and the camera.

Palabras clave— Vehículo autónomo; Yolo; Autoware; Red Neuronal .

1 Introducción

En el ámbito de la movilidad autónoma, la identificación de objetos es una de tareas iterativas que involucran parámetros y variables susceptibles de medición y control en tiempo real, constituye un desafío intrigante y de gran relevancia. El avance constante de la automatización en diversos campos de la ingeniería facilita la optimización de recursos, tanto de hardware como de software, en un sistema específico, permitiendo realizar tareas de manera más eficiente y segura.

En el contexto de la ingeniería automotriz, se ha observado un notable desarrollo en los sistemas de automatización de la conducción para vehículos terrestres. La Sociedad de Ingenieros Automotrices (SAE) ha establecido una clasificación de seis niveles para estos sistemas, los cuales abarcan la realización parcial o total de la tarea de conducción dinámica (DDT) de forma sostenida (1).

Para el desarrollo de la conducción autónoma es necesario considerar un sistema de percepción, sistema de planificación, y sistema de control. Todos estos interactúan en conjunto con una serie de sensores, como LiDAR (Light Detection and Ranging), cámara, radar, GNSS/IMU entre otros; estos sensores reciben toda la información del medio externo para enviarla a un procesador CPU/GPU donde, toma toda la información, la clasifica, la procesa y la filtra en datos y señales convenientes para el sistema de control en cuestión (2).

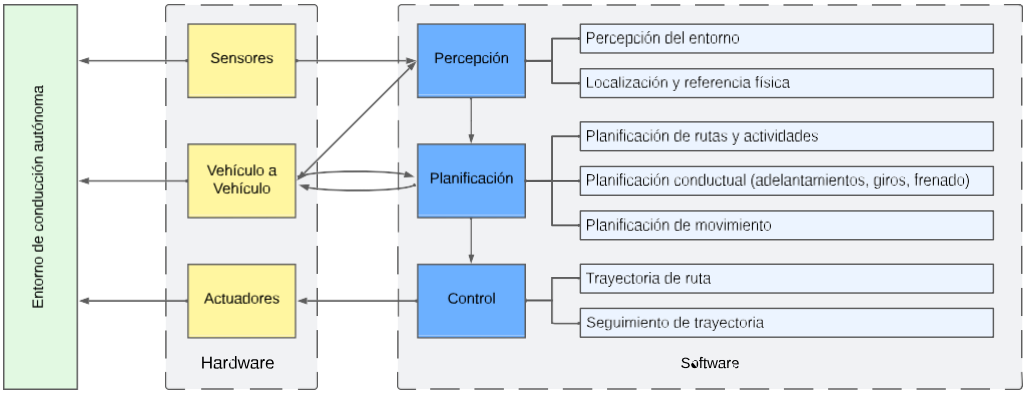


Figura 1: Configuración típica de un sistema de vehículo autónomo.

El sistema de gestión y toma de decisiones para la conducción autónoma se basa en diversos modelos de planificación y percepción, los cuales están fundamentados en la inteligencia artificial. Este enfoque se inclina particularmente hacia el uso de aprendizaje automático (ML) supervisado y no supervisado. En este contexto, el aprendizaje automático desempeña un papel esencial al permitir que el sistema se adapte y mejore continuamente a través de la experiencia, utilizando datos para ajustar dinámicamente sus capacidades de planificación y percepción en tiempo real (3).

Teniendo en cuenta las características mencionadas de la conducción autónoma, es crucial y necesario destacar la importancia en en el avance de la investigación y la implementación de vehículos autónomos.

El estudio realizado por (4), aborda la necesidad de mejorar la seguridad vial mediante la transferencia de tareas críticas de seguridad de los humanos a las máquinas. En particular, se centra en la detección de gestos manuales de los oficiales de tráfico, un área poco explorada hasta ahora en la investigación de vehículos autónomos. La implementación de técnicas de aprendizaje profundo y modelos LSTM demuestra ser efectiva, alcanzando una precisión del 80 % en la predicción de los gestos, lo que resalta su potencial para aplicaciones prácticas. Además, este estudio subraya la importancia de desarrollar un conjunto de datos específico para Ecuador, lo que mejoraría la exactitud y confiabilidad de los vehículos autónomos en condiciones locales de tráfico. La investigación destaca la necesidad de colaboraciones para crear conjuntos de datos más amplios y detallados, cruciales para el desarrollo de modelos más robustos y precisos. Este artículo contribuye significativamente al avance de los sistemas inteligentes de gestión de tráfico y la seguridad de los vehículos autónomos, abriendo nuevas posibilidades para su implementación en entornos reales.

1.1 Revisión Bibliográfica

El reconocimiento de objetos que forma parte del sistema de detección de un vehículo autónomo (AV), es sumamente importante debido a que representa la base fundamental para la toma de decisiones del resto de sistemas.

El sistema de detección de un AV juega un papel crucial al momento de reconocer objetos, ya que constituye el pilar esencial sobre el cual se fundamentan las decisiones de los demás sistemas. La capacidad para reconocer objetos en el entorno del vehículo no solo proporciona información sobre la presencia y ubicación de elementos, sino que también establece las bases necesarias para la toma de decisiones autónomas eficientes (3). En este sentido, el reconocimiento debe obtener una precisión y exactitud bastante alta de la realidad puesto que se establece como un componente vital para garantizar el rendimiento seguro y efectivo de todas las funcionalidades del AV (5).
El desarrollo exponencial de la inteligencia artificial, a permitido el uso del aprendizaje automático para generar modelos y algoritmos que atiendan las necesidades de los sistemas autónomos de conducción, en este sentido cada uno de los sistemas que componen un AV, usan métodos de aprendizaje enfocados a la percepción, planificación y control (6). Sin embargo, abordar cada uno de los sistemas a detalle es bastante extenso, de la misma forma es importante exponer las partes fundamentales que componen y que están inmersas en un AV.

La visión artificial es un campo ampliamente explorado, por lo que se a convertido en una tarea habitual usarla en muchos avances tecnológicos, no obstante el trabajo desarrollado en (7), analiza la eficiencia de modelos basados en redes neuronales de aprendizaje profundo, comparando métricas tales como *precision* y *recall* enfocado hacia el reconocimiento de objetos en distintos escenarios. De igual forma la evaluación realizada en (8) se enfoca en resaltar el rendimiento de redes neuronales de una y dos etapas midiendo el porcentaje de *accuracy* de distintas bases de datos usados en AV. A continuación se puede observar las tablas realizadas por los autores y modificadas con los datos mas relevantes y pertinentes con respecto al rendimiento de los algoritmos para reconocer objetos.

Tabla 1: Rendimiento de modelos usados en distintos escenarios (7).

		Faster R-CNN		Mask R-CNN		YOLOV5	
		Precision	Recall	Precision	Recall	Precision	Recall
Inte riores	Persona	0.72	0.83	0.95	0.90	0.97	0.93
	Silla	1.0	0.11	0.57	0.82	1.0	0.11
Exteriores	Persona	0.91	0.50	0.95	0.46	0.87	0.85
	Silla	0.94	0.68	0.71	0.62	0.91	0.82
	Moto	0.35	1	0.64	1	0.14	1.0

Tabla 2: Rendimiento de modelos entrenados con distintos dataset(8).

	Red	Dataset	Accuracy	Time (ms)
Red Neuronal de una etapa	MSI-CNEM	PASCAL VOC	85.35 %	15.63
	BS3D	KITTI	84.80 %	21.88
	Dense-ACSSD	BSD 100K	84.02 %	28.57
	YOLO	PASCAL VOC	90.90 %	66.67
Red Neuronal de dos etapas	Mobile-Net	Del autor	91.59 %	66
	FMLA-CNN	KITTI	88.83 %	170
	Faster-RCNN	KITTI	91.20 %	200
	RoarNet	KITTI	84.25 %	65

En las Tablas 1 y 2 presentadas se comparan distintos modelos de detección de objetos basados en redes neuronales convolucionales R-CNN (Region-based Convolutional Neural Network) con sus distintas mejoras Faster & Mask; de igual manera se usa YOLO (You Only Look Once) en su

versión 5.
Como se puede evidenciar en los resultados presentados en las tablas anteriores, se concuerda con (7; 8; 9) que existe una eficiencia bastante alta con el algoritmo YOLO, en cuanto a las métricas antes mencionadas, sin embargo se puede destacar que Mobile-Net y Faster-RCNN es otro modelo que destaca por sus buenos niveles de precision, recall y accuracy, sin embargo al ser estos redes de dos etapas, tienen un tiempo de detección mucho mayor, es por eso que YOLO tiene un equilibrio entre métricas y tiempo de detección, y es uno de los mas usados en vehículos autónomos comerciales y de desarrollo.

Ahora por otra parte es imperativo destacar la importancia vital de los sensores, ya que desempeñan una función fundamental al proveer información en tiempo real acerca del entorno circundante. Estos dispositivos de captura de datos son esenciales para el procesamiento de señales que se requiere en cada sistema del AV, facilitando la toma de decisiones apropiadas y la ejecución de maniobras seguras y eficientes. En el artículo (2) se menciona, relata y compara los distintos sensores que existen para la detección de objetos 2D y 3D; para que un algoritmo de detección de objetos a usar sea lo mas eficiente posible se requiere de una fusión de sensor pasivo y activo, tal como menciona (10), la fusión Cámara-LiDAR es de las más utilizada para crear mapas de entornos desconocidos y de esta forma poder determinar la ubicación del AV en dicho entorno mediante la técnica SLAM (Simultaneous Localization and Mapping) para sistemas robóticos y AV´s, por un lado la cámara nos proporciona características 2D del objeto en cuestión, con la desventaja de que es sensible a variables meteorológicas, y esto se compensa con el uso del LiDAR, ya que obtenemos información 3D sobre el entorno que rodea al vehículo autónomo, obteniendo la posición y forma geométrica del objeto (11).

En el trabajo realizado por (11) se evidencia que al integrar datos provenientes de diversos sensores, es imprescindible llevar a cabo la calibración tanto espacial como temporal de los mismos. La calibración espacial se torna esencial para la conversión de los datos tridimensionales captados por el sensor LiDAR al espacio bidimensional de la cámara. La realización de la calibración implica establecer los parámetros extrínsecos LiDAR-cámara, los cuales están expresados mediante las matrices de rotación (R) y transformación (T). En este contexto, la matriz R representa la rotación, mientras que la matriz T describe la transformación que especifica las relaciones entre los sistemas de coordenadas del LiDAR y la cámara.

1.2 Motivación

Con referencia a los estudios anteriores mencionados cabe destacar que en la actualidad la ingeniería se enfrentan al desafío de integrar diversas disciplinas, desde la mecánica hasta la electrónica y la inteligencia artificial, para crear sistemas mecatrónicos altamente complejos (12). Esta convergencia tecnológica no solo amplía el conjunto de habilidades necesarias, sino que también requiere una comprensión profunda de cómo estas disciplinas interactúan para lograr un rendimiento óptimo en los vehículos modernos. Como arquitectos de sistemas complejos, los ingenieros deben dominar no solo la mecánica y la electrónica, sino también la dinámica de vehículos y el control para impulsar la innovación en la creación de sistemas mecatrónicos avanzados y complejos.

Este cambio en la naturaleza de la ingeniería automotriz plantea la necesidad crítica de que las instituciones educativas adapten sus planes de estudio para proporcionar una formación integral adaptativa con los avances tecnológicos. Además el uso y la adquisición de hardware y software para hacer frente a estos retos es muy necesario. Como menciona en su trabajo (13), existen kit´s de desarrollo enfocadas en la conducción autónoma, uno de estos es **PixKit** distribuido por la empresa **"PIX Moving"** y utilizado en el proyecto de investigación "Desarrollo de estrategias de movilidad inteligente, sostenible y aceptación social de vehículos autónomos en la ciudad de Cuenca, empleando técnicas de inteligencia artificial y realidad virtual en plataformas de software y hardware especializados "del Grupo de Investigación en Ingeniería del transporte GIIT de la Universidad Politécnica Salesiana. Conforme a esta idea la motivación de esta redacción técnica principalmente es emplear la fusión de sensores Cámara-LiDAR para el reconocimiento de objetos mediante el uso de inteligencia artificial. Además se propone generar un algoritmo para el reconocimiento de objetos por medio de sensores de percepción horizontal LiDAR y cámaras de visión enfocados hacia los vehículos autónomos terrestres.

1.3 Aportes del estudio

Esta investigación abordará el reto antes descrito, por lo cual se plantea el uso del kit de desarrollo enfocado a la conducción autónoma, mismo que dispone de una combinación hardware y software bastante integro; con sensores, actuadores, chasis, interfaz Humano-Maquina, software libre.

- Desarrollo una metodología adecuada para la adquisición y procesamiento de señales y datos que proveen los sensores disponibles en el kit de desarrollo, con este enfoque se aborda realizar la correcta configuración e instalación de todos los paquetes de software requerido para poder enlazar el medio físico con el sistema embebido que dispone el automóvil.
- Calibración de sensores LiDAR y cámara mediante archivos precargados de configuración para la correcta fusión de sensores del software embebido Autoware en el vehículo.
- Ajuste de los parámetros de la interfaz de visualización (RViz), a través de guías proporcionadas por el fabricante con el fin de usar los recursos disponibles de hardware & software del vehículo; enfocado hacia los métodos y modelos revisados.

2 Sistema de chasis del vehículo

En esta sección se describe de manera general la totalidad de hardware disponible en el kit de desarrollo PixKit utilizado en el trabajo de investigación. En la Figura 2 se observa la infraestructura del vehículo utilizado. Este chasis es completamente de código abierto en Gitlab, utilizando el protocolo de código abierto General Public License version 3 (GPLv3); una licencia de software libre publicada por la Free Software Foundation (FSF). Además proporciona todo el código de control subyacente y planos de ingeniería 3D/2D relacionados con la estrategia de control del vehículo.

El chasis esta construido en su totalidad con materiales de alta resistencia, específicamente con aleaciones de aluminio, cuenta con sistema de propulsión, sistema de dirección, sistema de frenado, sistema eléctrico, electrónico y seguridad, sistema de iluminación del chasis y sistema de comunicación. Destacando la red CAN del chasis PIX es un enfoque de comunicación importante para todo el vehículo. Actualmente hay un CAN BUS en el chasis, al que están conectados todos los dispositivos que necesitan comunicarse a través de CAN. El bus CAN proporciona múltiples entradas de dispositivos, así como una interfaz de diagnóstico que permite monitorear y reportar el estado de varios subsistema (OBD), en el manual de operación (14) se puede acceder a información mas detallada de cada una de los sistemas y especificaciones del dispositivo.

Figura 2: Infraestructura de vehículo PixKit.

El conjunto de dispositivos electrónicos y sensores con el que cuenta el chasis es compatible

con todos los escenarios y aplicaciones de la conducción autónoma. Es importante destacar que el vehículo admite el despliegue de las funciones de conducción autónoma con el software Autoware.

3 Sistema de Software Integrado

El modelo de vehículo de desarrollo esta cargado con el software Autoware el cual es de código abierto basado en ROS (Robot Operating System), desarrollado por la universidad de Nagoya y destinado a la investigación y el desarrollo de conducción autónoma. En la conferencia desarrollada en (15) se describe y contrasta; la arquitectura, características, enfoques, alcances y aplicaciones del software Apollo y Autoware, evidenciando que este último tiene un enfoque más adaptativo y con mejores prestaciones para el ámbito de desarrollo e investigación. En la Figura 3 se presenta la arquitectura que usa Autoware; destacando la configuración típica de un AV. Sensing; en el cual disponemos de gran robustez y compatibilidad para varios sensores. Actuation; en la cual se ajustan cada uno de los parámetros como son, las coordenadas Yaw, Pitch y Roll para el sistema de referencia. El cual se entrelaza directamente con los sistemas de frenado, motores, servomotores, suspensión, etc. Computing; donde se destaca principalmente el control y la autonomía de todo el sistema en conjunto, dividiéndose en tres grandes subsistemas como son el de percepción, planificación y decisión.

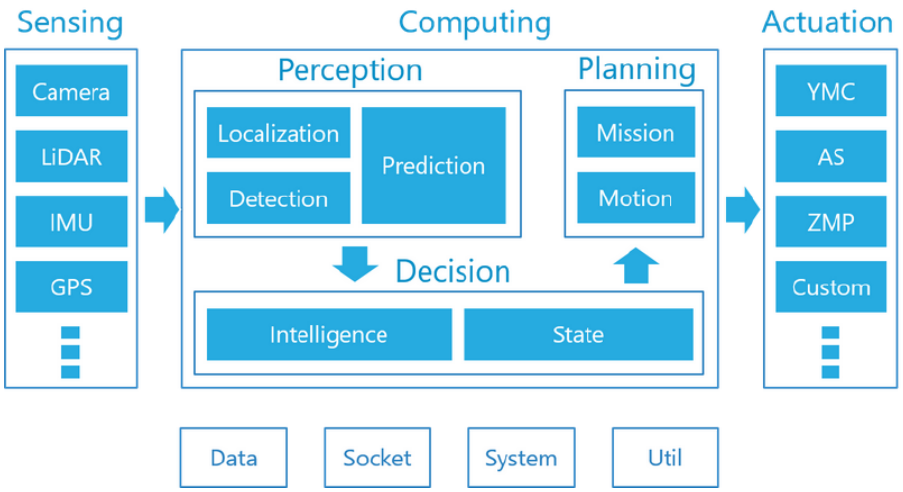


Figura 3: Arquitectura Autoware (14).

ROS utiliza el envió y recepción de mensajes e información a través de temas de forma publish/subscribe para marcos de conexión/cooperación ínter- módulos, de esta forma permitiendo que cada uno de los suscriptores que están enlazados a un tema en especifico puedan compartir información de manera bidireccional. Aquí, el paso de mensajes es un mecanismo de comunicación entre procesos en el que un emisor puede enviar datos a uno o más receptores. Esta característica nos permite diseñar sistemas de distribución. En ROS, se lanzan procesos denominados nodos y cada nodo se ejecuta de forma independiente. En la comunicación entre nodos, siguiendo el esquema publish/subscribe, un nodo escribe mensajes (publish) en un topic y otro nodo lee los mensajes (subscribe) del topic.

En la Figura 5 podemos observar la ventada del Runtime Manager de Autoware, en donde se dispone de un inicio rápido para configurar y ajustar cada uno de los subsistemas del AV, destacando su setup, computing, map, sensing, interface, databes, simulation, status y topic. Las ventanas emergentes muestran los programas ejecutables del LiDAR y del proceso del filtro.

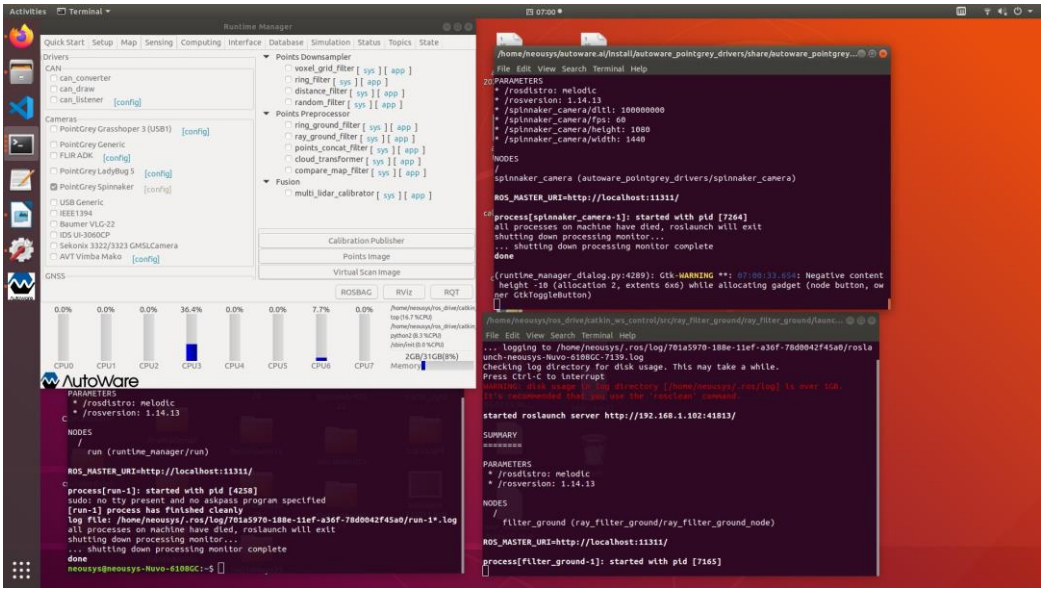


Figura 4: Ejecución del Autoware con sus dependencias.

3.1 Calibración LiDAR

En un sistema de percepción autónomo, la alineación de posición y orientación entre los sensores LiDAR y cámara respecto al sistema de coordenadas del vehículo requiere definir las transformaciones extrínsecas que los relacionan. Este proceso implica calcular y aplicar tanto las traslaciones (x, y & z) como las rotaciones (roll, pitch & yaw) necesarias para transformar datos entre los marcos de referencia individuales de cada sensor y el marco de referencia del vehículo, comúnmente denominado base_link. Estas transformaciones garantizan que la información captada por el LiDAR y la cámara se proyecte correctamente en el espacio del vehículo, permitiendo una representación coherente y precisa de los datos en el entorno del vehículo (16).
Mediante Calibration Tool Kit obtenemos una nube de puntos 3D, la cual es guardada en la dirección points_raw, ademas la herramienta Euclidean Cluster y Voxel Grid Filter se utilizan para optimizar la percepción de datos de LiDAR. Euclidean Cluster agrupa puntos basándose en la distancia euclidiana, lo que permite segmentar objetos individuales en el entorno, como otros vehículos, peatones y obstáculos, mejorando la identificación de cada uno en el mapa de percepción. Por su parte, Voxel Grid Filter reduce la densidad de puntos en la nube, disminuyendo la carga de procesamiento al agrupar puntos en pequeños vóxeles y seleccionando un punto representativo por cada voxel. Esto permite procesar datos LiDAR en tiempo real sin comprometer la estructura del entorno, facilitando el uso eficiente de los recursos en Autoware.

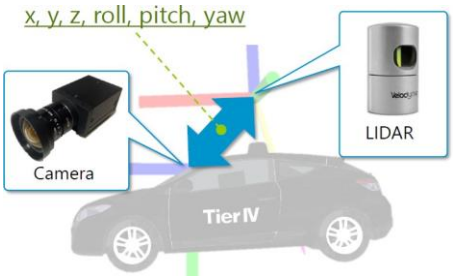


Figura 5: Disposición para la calibración de sensores (14).

3.2 Calibración Cámara

En Autoware, tanto SSD (Single Shot MultiBox Detector) como YOLO (You Only Look Once) son modelos de detección de objetos en tiempo real, cada uno con ventajas específicas. Como menciona (17) SSD ofrece un buen balance entre precisión y velocidad, ya que utiliza múltiples "feature maps"para detectar objetos de diferentes tamaños, lo que resulta útil para captar objetos pequeños en el entorno. YOLO, en cambio, es más rápido en procesamiento debido a su arquitectura de una sola pasada (single pass), dividiendo la imagen en una cuadrícula y realizando detección de objetos en tiempo récord, aunque su precisión con objetos pequeños puede ser inferior. En aplicaciones de Autoware, SSD es ideal cuando se requiere alta precisión en la detección de obstáculos, mientras que YOLO es preferido en situaciones donde la velocidad es prioritaria, como en tareas de detección de objetos en tiempo real.

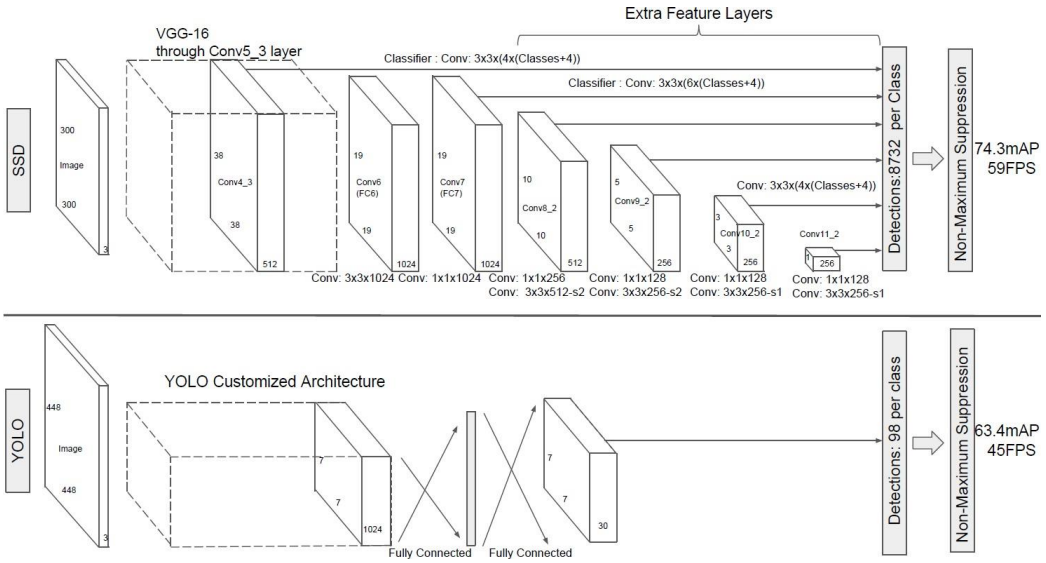


Figura 6: Comparación de modelos de detección presentes en Autoware (17).

Autoware dispone de un sinnúmero de interfaces y herramientas para el correcto manejo de sensores, actuadores y datos de control del AV. Uno de ellos es el Calibration Tool Kit que permite realizar la calibración de parámetros de una cámara, que son necesarios para corregir la distorsión y proyectar imágenes capturadas con precisión. La calibración precisa es esencial para garantizar que los datos recogidos por los sensores se alineen correctamente en un entorno 3D. Para ello es necesario proveer al toolkit varias fotografías a distintas distancias y ángulos, usando un tablero de ajedrez de 7x9 como se presenta en la Figura 6. Estas fotografías ayudan a corrección y reducción de errores en la toma de datos, tener variedad de puntos de vista, calibración de estereocámaras, proveer distintas condiciones de iluminación y generar suficientes datos basados en los puntos detectados de cada imagen.

A continuación el toolkit se encarga de generar un archivo de almacenamiento con información estructurada que es típico para la serialización de datos y usado muy habitualmente para escribir archivos de configuración, este archivo con extensión *.YAML* el cual se carga dentro de Autoware para el funcionamiento y reconocimiento de las correcciones de la cámara a utilizar.

En la Figura 7 disponemos del archivo de calibración de la cámara, en donde; CameraExtrinsicMat: define la matriz extrínseca de la cámara, que describe la transformación desde el sistema de referencia espacial tridimensional. El sistema de coordenadas Universal Tranverse Mercator (UTM) es utilizado por su mayor dependencia en la calibración de sensores por la georreferenciación de datos precisa con respecto a la ubicación tridimensional en un punto en la Tierra.
CameraMat: define la matriz intrínseca de la cámara, que contiene parámetros como la longitud focal y el punto principal de la cámara. DistCoeff: define los coeficientes de distorsión del lente, que describen las distorsiones ópticas presentes en la imagen capturada por la cámara.

ImageSize: tamaño de las imágenes capturadas por la cámara, especificado en píxeles (ancho x alto). En este caso, la resolución es 1440x1080 píxeles.
Reprojection Error: El error de reproyección promedio, que es una medida de la precisión de la calibración. Idealmente, debería ser lo más bajo posible. En este caso, se muestra como 0.0, lo que podría indicar una calibración ideal.
DistModel: Modelo de distorsión utilizado para describir la distorsión de la lente de la cámara. plumb_bob es un modelo común que incluye distorsiones radiales y tangenciales.

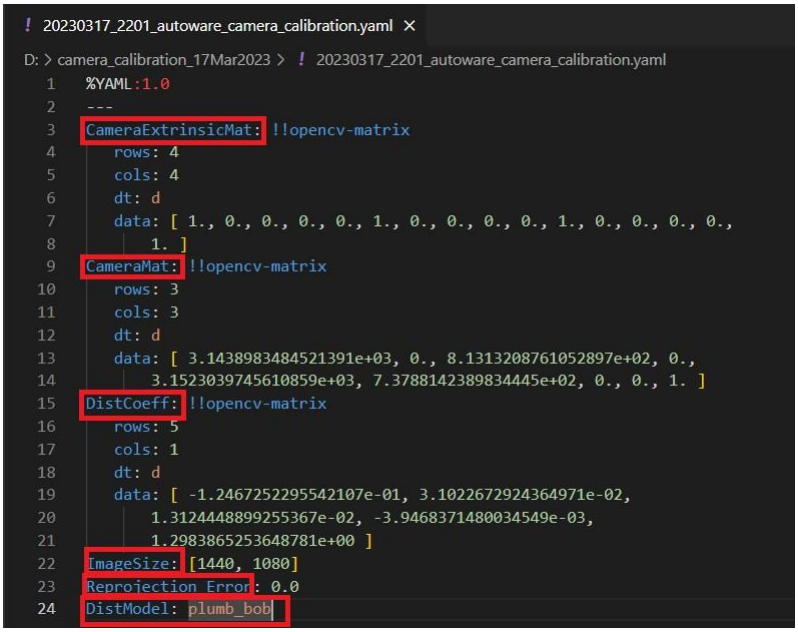


Figura 7: Archivo de calibración de cámara.

3.3 Calibración Cámara-LiDAR

En la Figura 8 se encuentra las lineas de código del archivo YAML que contiene toda la información necesaria para calibrar una cámara con referencia a LiDAR, incluyendo las matrices intrínsecas y extrínsecas, los coeficientes de distorsión, el tamaño de la imagen, y el modelo de distorsión. Estos parámetros son esenciales para corregir y transformar las imágenes capturadas por la cámara para su uso en aplicaciones de visión por computadora, como las utilizadas en vehículos autónomos. De igual forma al usar la función programada Fusion Tool, se necesita calibrar los dos sensores y poder unir la nube de puntos que forma el LiDAR, con los píxeles de la imagen que proyecta la cámara y tenerlos referenciados en un mismo plano y espacio de coordenadas, para posteriormente gracias a estos dos sensores realizar el reconocimiento de los objetos que se encuentran alrededor del vehículo autónomo.

El dataset KITTI es fundamental en la investigación de detección de objetos para sistemas autónomos, ya que ofrece una amplia variedad de escenarios urbanos y rurales con anotaciones precisas para LiDAR y cámaras, lo que permite probar y mejorar algoritmos de percepción y detección. Su riqueza en datos multimodales (LiDAR, cámaras y GPS) lo convierte en una referencia clave para calibraciones precisas y para la validación de métodos de fusión de sensores, esenciales en la percepción de entornos complejos. En Autoware, KITTI se utiliza para evaluar y optimizar los modelos de percepción, apoyando tanto la detección de objetos estáticos como dinámicos en tiempo real, y permite entrenar sistemas de visión por computadora para tareas de reconocimiento y seguimiento de objetos, lo que refuerza la estabilidad y precisión del sistema de percepción del vehículo autónomo (18). Ademas el data set COCO (Common Objects in Context) es usado para entrenar la red neuronal profunda YOLO3, ya que incluye 80 clases de objetos, que abarcan desde vehículos, peatones y bicicletas hasta otros elementos del entorno urbano.

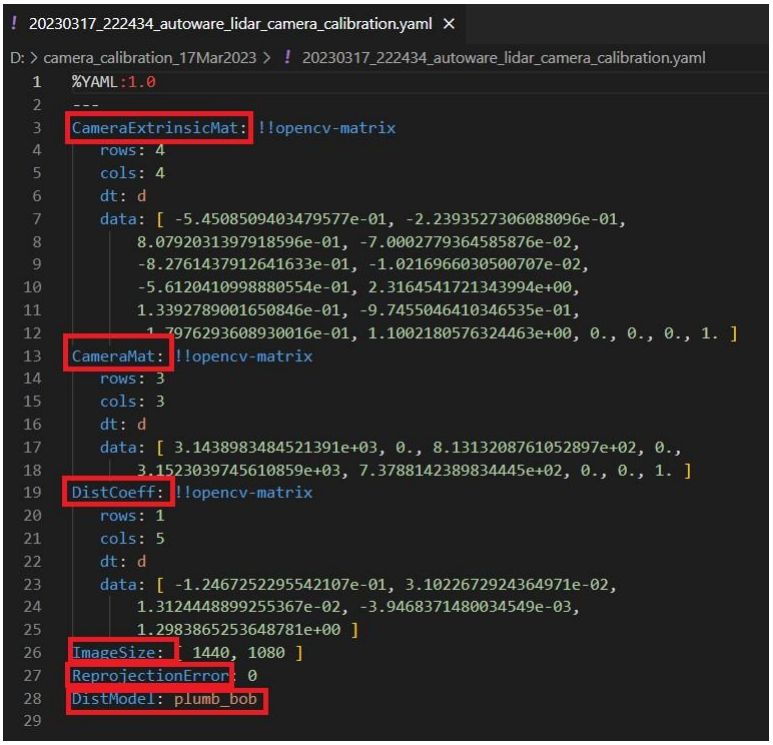


Figura 8: Archivo de calibración de cámara-LiDAR.

4 Resultados y discusión

La detección de objetos mediante cámara y LiDAR en un sistema basado en ROS y Autoware utiliza varios componentes clave para un funcionamiento eficiente y preciso. El spinnaker_camera_driver permite la captura de imágenes de alta calidad desde una cámara Point Grey, mientras que el voxel_grid_filter reduce la densidad de las nubes de puntos del LiDAR para un procesamiento más eficiente.

El nodo vision_darknet_yolo3 utiliza el modelo YOLOv3 para detectar objetos en las imágenes capturadas, proporcionando cuadros delimitadores y clasificaciones de objetos.

Pixel_cloud_fusion y range_vision_fusion se encargan de fusionar los datos de la cámara y el LiDAR, integrando información visual y de distancia para mejorar la precisión de la detección.

Ndt_matching ayuda en la localización precisa del vehículo, alineando las nubes de puntos del LiDAR con un mapa preexistente.

Finalmente, vel_pose_connect sincroniza los datos de velocidad y posición del vehículo, asegurando la coherencia temporal de todos los datos utilizados en el sistema. Juntos, estos componentes permiten una detección robusta y precisa de objetos, esencial para la navegación autónoma segura y eficiente.

Como se puede observar en la Figura 9, se logra detectar las tres personas ubicadas a 15m aproximadamente del punto de referencia, estando aun dentro del rango de visión de LiDAR, y mostrando una velocidad de detección sumamente rápida que es algo primordial para un AV. Es importante destacar que el tiempo de detección y la precisión que presenta YOLO3 en autoware es bastante eficiente, en comparación con el resto de modelos de redes neuronales como se analizó en la tabla 1 y 2

Por otra parte también se realizó otra prueba que se puede apreciar en la Figura 10 con los objetos mas cercanos que interfieran con el rango de visión de la cámara, para poder observar el comportamiento de la detección, y de igual forma resultó ser muy preciso en sus resultados. Además se observa la detección de objetos mediante la fusión de datos entre la cámara y el sensor

LiDAR, obteniendo una percepción amplia y precisa del entorno que se ha realizado las pruebas. Estas características son las principales para una mejor percepción del entorno en una navegación autónoma.

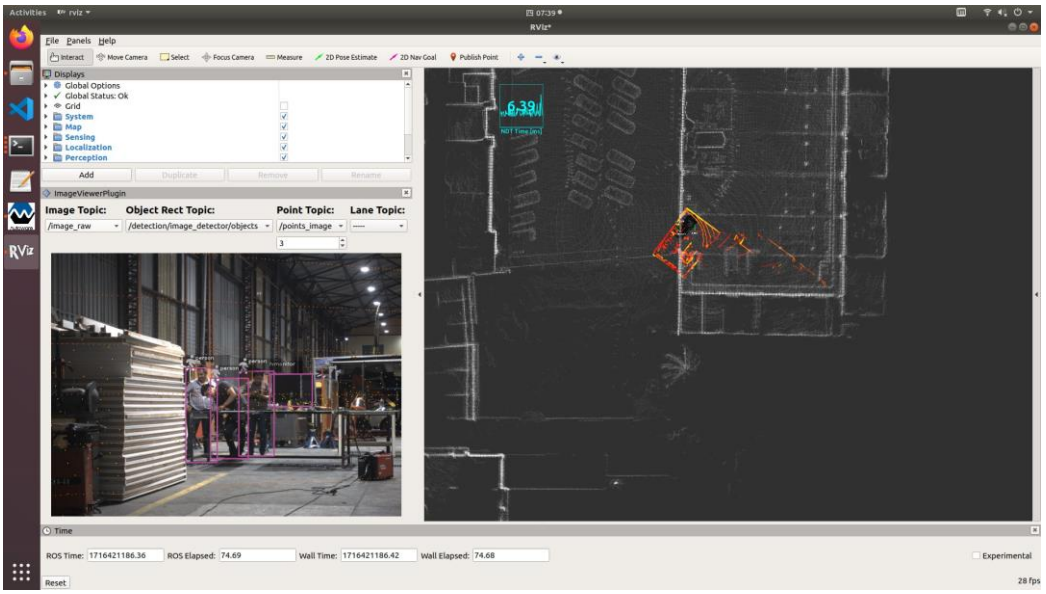


Figura 9: Detección de objetos lejanos con YOLO3.

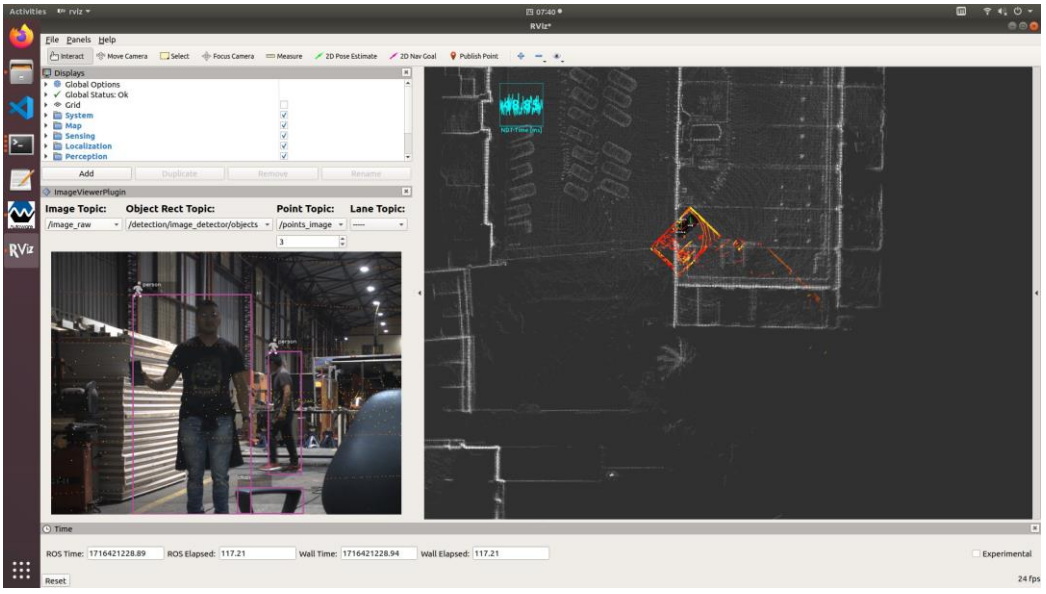


Figura 10: Detección de objetos cercanos con YOLO3.

La interfaz de código abierto que ofrece Autoware para el desarrollo de vehículos autónomos, es una excelente herramienta al momento de integrar y fusionar distintos sensores, como cámara, LiDAR, GNSS, ultrasónicos, etc. De la misma forma tener modelos, métodos y técnicas enfocadas para el mapeo, procesamiento y detección de las señales de cada uno de los sensores que se los agregue. PointGround Spinnaker es una de las técnica de detección y clasificación de objetos justamente diseñada para su uso en sistemas de conducción autónoma, que nos brinda el software Autoware. Esta técnica utiliza el sensor LiDAR para capturar datos del entorno circundante del vehículo. En

el GitHub de Autoware se facilita varios ejemplos, guías y manuales básicos de configuración para el manejo del hardware y software del vehículo autónomo, que son una gran base y ayuda para poder realizar este artículo.

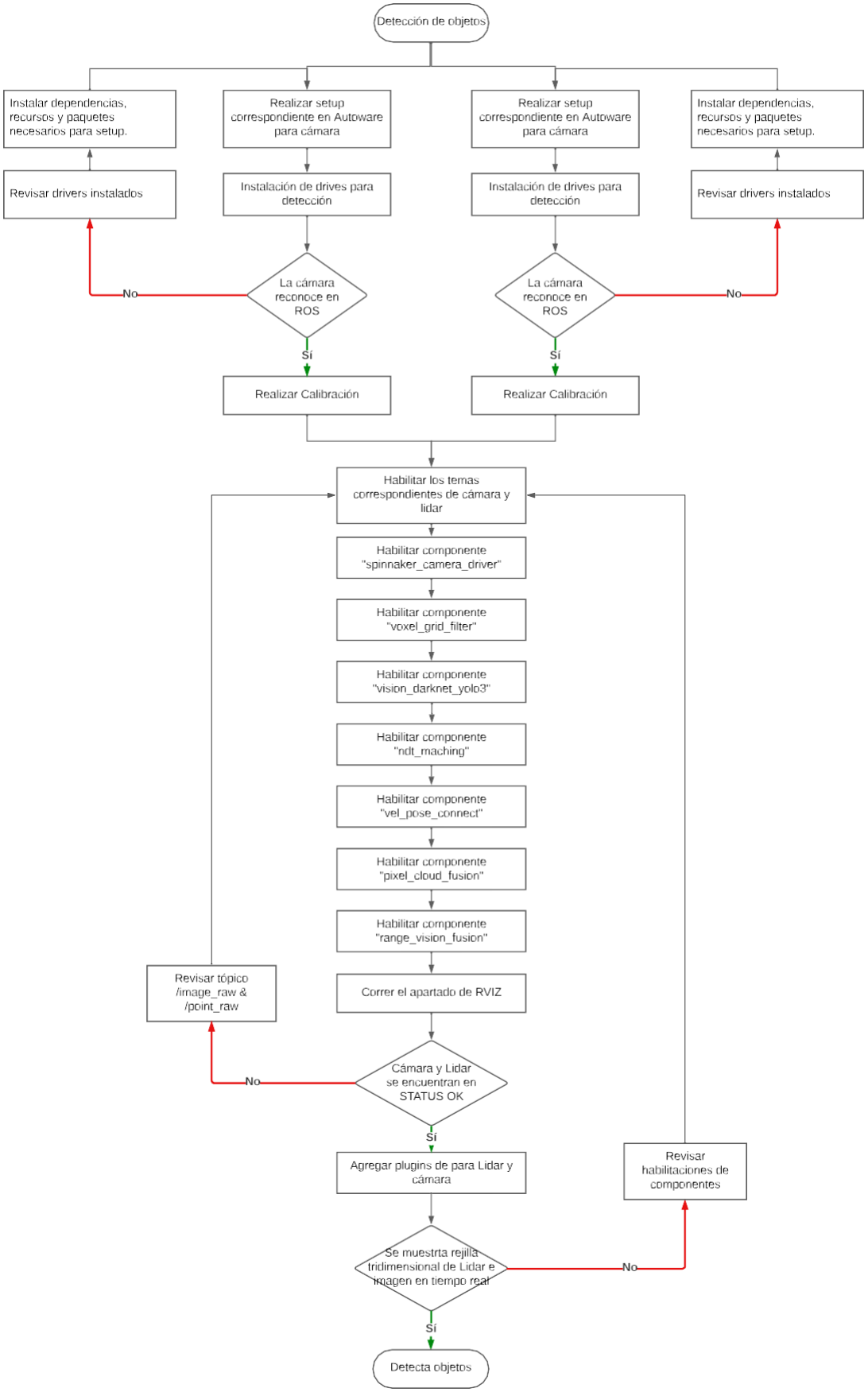


Figura 11: Diagrama de flujo para la detección de objetos mediante LiDAR y cámara.

5 Conclusiones

En este artículo se presenta el reconocimiento de objetos mediante técnicas de visión artificial a través de cámara y LiDAR enfocado hacia los vehículos autónomos terrestres. El hecho de contar con hardware y software especializado en el tópico es un punto de apoyo muy fuerte al momento de realizar la investigación y el desarrollo del presente trabajo.

Autoware un software que utiliza sistemas embebidos para vehículos autónomos, presta herramientas para de desarrollo de proyectos que ayuden en la optimización de los vehículos autónomos. Siendo el caso de este artículo, se logró incorporar los sensores de cámara y LiDAR a una IPC (Industrial Personal Computer), cotando con un hardware (chasis) es realmente beneficioso ya que provee el factor de calibrar y probar los sensores en un entorno y condiciones reales.

Al tener hardware y software orientado en el tópico de este artículo, se efectuó una correcta metodología para realizar una revisión bibliográfica de los equipos e interfaces a usar. Es importante tener en cuenta que todo el software y los archivos de manuales, guías y ejemplos son de código abierto y están presentes en el Github de Autoware que proporciona el mismo fabricante.

La calibración y fusión de los sensores; cámara y LiDAR se enfocaron en la detección de objetos, el uso de YOLOV3 una una red neuronal convolucional (RNN), así como también una técnica Point Ground Spinnaker para la detección de objetos, proporciona la velocidad y precisión esperada en los resultados experimentales, demostrando así las ventajas que ofrece en comparación con otros modelos que son YOLOV2 y SSD que también se pueden usar en el software. Sin embargo queda clara la necesidad y viabilidad de integrar más sensores, ya que la detección de objetos es muy buena, pero el sistema de vehículo autónomo cuenta con puntos ciegos y un rango corto de detección 20m.

La integración de componentes como el spinnaker_camera_driver, voxel_grid_filter, vision_darknet_yolo3, ndt_matching, pixel_cloud_fusion y range_vision_fusion en un sistema basado en ROS y Autoware demuestra la importancia de una arquitectura modular y flexible para la detección de objetos. Esta combinación de herramientas y algoritmos permite una percepción robusta y precisa al fusionar datos de múltiples sensores, como cámaras y LiDAR, facilitando la detección de objetos en entornos complejos y variables.

La calibración adecuada de la cámara y el LiDAR desempeña un papel fundamental en el éxito de la detección de objetos. La calibración intrínseca y extrínseca asegura la precisión en la captura de datos y la alineación espacial entre los diferentes sensores, lo que es crucial para una fusión de datos efectiva y una detección precisa de objetos. Una calibración bien realizada mejora la calidad de los datos de entrada y, por lo tanto, la fiabilidad de los resultados de detección, lo que es esencial para aplicaciones de vehículos autónomos donde la seguridad y la precisión son de suma importancia.

6 Agradecimientos

Los autores agradecen al proyecto «Desarrollo de Estrategias de Movilidad Inteligente, Sostenible y Aceptación Social de Vehículos Autónomos en la Ciudad de Cuenca, Empleando Técnicas de Inteligencia Artificial y Realidad Virtual en Plataformas de Software y Hardware Especializados» del Grupo de Investigación en Ingeniería del Transporte (GITT) de la Universidad Politécnica Salesiana por facilitar los datos utilizados en este documento.

Referencias

[1] S. Int, "Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles j3016$_2$02104," 2021. *F echadeAcceso* : 20*deOctubrede*2023.

[2] A. Ghasemieh and R. Kashef, "3d object detection for autonomous driving: Methods, models, sensors, data, and challenges," *Transportation Engineering*, vol. 8, 6 2022.

[3] P. Antonante, H. Nilsen, and L. Carlone, "Monitoring of perception systems: Deterministic, probabilistic, and learning-based fault detection and identification," *Artificial Intelligence*, p. 103998, 12 2023.

[4] P. Idrovo-Berrezueta, J. P. Ortiz, J. D. Valladolid, and F. E. Espinoza, "Development of an lstm-based model with attention mechanism for detection of traffic's officer hand gestures for autonomous vehicles in ecuador," in *Systems, Smart Technologies and Innovation for Society* (J. P. Salgado-Guerrero, H. R. Vega-Carrillo, G. García-Fernández, and V. Robles-Bykbaev, eds.), (Cham), pp. 106–115, Springer Nature Switzerland, 2024.

[5] M. Carranza-García, P. Lara-Benítez, J. García-Gutiérrez, and J. C. Riquelme, "Enhancing object detection for autonomous driving by optimizing anchor generation and addressing class imbalance," *Neurocomputing*, vol. 449, pp. 229–244, 8 2021.

[6] J. Y. Yoon, J. Jeong, and W. Sung, "Design and implementation of hd mapping, vehicle control, and v2i communication for robo-taxi services," 9 2022.

[7] X. Yu, S. Salimpour, J. P. Queralta, and T. Westerlund, "General-purpose deep learning detection and segmentation models for images from a lidar-based camera sensor," *Sensors*, vol. 23, 3 2023.

[8] Q. Liu, Z. Li, S. Yuan, Y. Zhu, and X. Li, "Review on vehicle detection technology for unmanned ground vehicles," 2 2021.

[9] P. Azevedo and V. Santos, "Comparative analysis of multiple yolo-based target detectors and trackers for adas in edge devices," *Robotics and Autonomous Systems*, vol. 171, 1 2024.

[10] Q. Wang, C. Yan, R. Tan, Y. Feng, Y. Sun, and Y. Liu, "3d-cali: Automatic calibration for camera and lidar using 3d checkerboard," *Measurement: Journal of the International Measurement Confederation*, vol. 203, 11 2022.

[11] N. Sakic, M. Krunic, S. Stevic, and M. Dragojevic, "Camera-lidar object detection and distance estimation with application in collision avoidance system," vol. 2020-November, IEEE Computer Society, 11 2020.

[12] Árpád Fehér, S. Aradi, T. Bécs, and P. Gáspár, "Highly automated electric vehicle platform for control education," vol. 53, pp. 17296–17301, Elsevier B.V., 2020.

[13] J. F. Arango, L. M. Bergasa, P. A. Revenga, R. Barea, E. López-Guillén, C. Gómez-Huélamo, J. Araluce, and R. Gutiérrez, "Drive-by-wire development process based on ros for an autonomous electric vehicle," *Sensors (Switzerland)*, vol. 20, pp. 1–31, 11 2020.

[14] Pixmoving, *PIX Moving PIX LOOP X-by-Wire Chassis User Operating Manual*. Pixmoving, 2023.

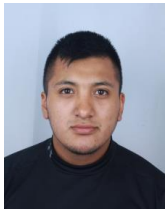
[15] *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*. IEEE.

[16] F. Ren, H. Liu, and H. Wang, "A LiDAR-Camera Joint Calibration Algorithm Based on Deep Learning," *Sensors*, vol. 24, no. 18, 2024.

[17] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *ECCV*, 2016.

[18] P. Rotter, M. Klemiato, and P. Skruch, "Automatic Calibration of a LiDAR—Camera System Based on Instance Segmentation," *Remote Sensing*, vol. 14, no. 11, 2022.

7 Biografía



Byron Chucaralao Byron Chucaralao es estudiante del programa de Maestría en Electrónica y Automatización. El obtuvo su título de Ingeniero en Mecatrónica en el 2022 por la Universidad Politécnica Salesiana. Actualmente se desempeña como Programador-Analista del Departamento de Mantenimiento de la empresa INACI CORP.

Jonnathan Jarro El autor es estudiante del programa de Maestría en Electrónica y Automatización. El obtuvo su título de Ingeniero en Mecatrónica en el 2022 por la Universidad Politécnica Salesiana.

Juan Diego Valladolid Juan Diego Valladolid es actualmente profesor e investigador de la Universidad Politécnica Salesiana. Obtuvo su doctorado en Ingeniería en la Pontificia Universidad Javeriana y su maestría en Automatización y Control Industrial en la Universidad Politécnica Salesiana, en 2010 y 2014 respectivamente. Es profesor del programa de Ingeniería Automotriz y Mecatrónica de la Universidad Politécnica Salesiana. Es miembro del Grupo de Investigación en Ingeniería del Transporte (GIIT). Sus intereses de investigación son la teoría de control no lineal, los vehículos inteligentes, los vehículos eléctricos y la inteligencia artificial.