



UNIVERSIDAD AUTÓNOMA CHAPINGO

**POSGRADO EN INGENIERÍA AGRÍCOLA Y USO
INTEGRAL DEL AGUA**

**SISTEMA AUTOMÁTICO CON VISIÓN ARTIFICIAL PARA
DETECCIÓN DE MALEZAS EN MAÍZ**

**TESIS QUE COMO REQUISITO PARCIAL PARA OBTENER EL GRADO:
DOCTOR EN INGENIERÍA AGRÍCOLA Y USO INTEGRAL DEL AGUA**



APROBADA



PRESENTA:

DAVID IVÁN SÁNCHEZ CHÁVEZ

BAJO LA SUPERVISIÓN DE:

DR. AGUSTÍN RUIZ GARCÍA

Chapingo, Estado de México, noviembre de 2023.



SISTEMA AUTOMÁTICO CON VISIÓN ARTIFICIAL PARA DETECCIÓN DE MALEZAS EN MAÍZ

Tesis realizada por David Iván Sánchez Chávez bajo la dirección del comité asesor indicado, aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

DOCTOR EN INGENIERÍA AGRÍCOLA Y USO INTEGRAL DEL AGUA

Director:



Dr. Agustín Ruíz García

Asesor:



Dr. Noé Velázquez López

Asesor:



Dra. Raquel Salazar Moreno

Lector externo:



Dra. Karina Ruby Pérez Daniel

ÍNDICE DE CONTENIDO

1. INTRODUCCIÓN GENERAL.....	1
1.1 Justificación.....	6
1.2 Objetivos	6
1.2.1 Objetivo General.....	6
1.2.2 Objetivos Específicos.....	7
2. REVISIÓN DE LITERATURA.....	8
2.1 Identificación automática de malezas.....	8
2.2 Identificación de malezas con métodos de aprendizaje de máquina tradicional.....	9
2.3 Identificación de malezas con métodos de Aprendizaje profundo	16
2.4 Navegación de vehículos en campos agrícolas	30
2.4.1 Sensores utilizados para navegación en vehículos para uso agrícola	31
2.5 Literatura citada.....	36
3. UN ENFOQUE DE APRENDIZAJE PROFUNDO PARA LA DETECCIÓN DE MAÍZ Y MALEZAS	45
A Deep Learning Approach for Corn and Weed Detection.....	46
3.1 Introducción	48
3.2 Materiales y métodos	52
3.2.1 Métodos.....	52
3.2.2. Materiales y adquisición de datos.....	54

3.2.3 Métricas para evaluar modelos de Detección de Objetos y Segmentación de Instancias.....	58
3.3 Entrenamiento del modelo.....	59
3.3.1. Adquisición de datos.....	59
3.3.2 Proceso de etiquetado de la base de datos.....	60
3.3.3. Modelo de estimación.....	66
3.4 Resultados experimentales y discusión.....	70
3.4.1 RetinaNet.....	70
3.4.2 Red U-Net.....	71
3.4.3 YOLO.....	73
3.5 Conclusiones.....	75
3.6 Referencias	76
4. ROBOT PARA NAVEGACIÓN EN CAMPOS DE MAÍZ PARA EL FIELD ROBOT EVENT	79
Robot for navigation in maize crops for the Field Robot Event 2023	80
4.1 Introducción.....	81
4.2 Materiales y Métodos	83
4.2.1 Sistema de navegación.....	83
4.2.2 Sistema de Riego por Aspersión.....	95
4.2.3 Sistema de Detección de Imágenes	98
4.3. Resultados	100
4.3.1 Tarea 1: navegación	100
4.3.2 Tarea 2: tratamiento (aspersión) de las plantas.....	101

4.3.3 Tarea 3: detección y reconocimiento de posibles obstáculos	101
4.3.4 Tarea 4: obstáculos estáticos y dinámicos.	101
4.3.5 Pruebas	102
4.4 Conclusión.....	108
4.5 Referencias	108
5. APENDICES	112
5.1 Descripción de las tareas del Field Robot Event 2023	112
5.1.1 Tarea 1: Navegación.....	112
5.1.2 Tarea 2: Tratamiento (aspersión) de las plantas.....	114
5.1.3 Tarea 3: detección y reconocimiento de posibles obstáculos	116
5.1.4 Tarea 4: Obstáculos estáticos y dinámicos.....	117
5.2 Imágenes de las tareas del FRE 2023.	119

ÍNDICE DE CUADROS

Cuadro 1. Trabajos que utilizaron aprendizaje de máquina clásico.	15
Cuadro 2. Trabajos que utilizaron aprendizaje profundo.	26
Cuadro 3. Descripción de diferentes conjuntos de datos para la detección de malezas.....	60
Cuadro 4. Configuración del conjunto de datos para entrenar diferentes modelos de estimación.....	67
Cuadro 5. Diagrama de conexión Encoder-Arduino.....	88
Cuadro 6. Conexiones del lidar TFmini.....	95
Cuadro 7. Conexión del controlador Arduino maestro.	96
Cuadro 8. Ángulos medidos para giros a la izquierda del robot.....	102
Cuadro 9. Ángulos medidos para giros a la derecha del robot.	103
Cuadro 10. Desplazamientos del robot medidos en movimiento.	105
Cuadro 11. Desplazamientos del robot medidos en movimiento.	105
Cuadro 12. Desplazamientos del robot medidos en movimiento.	106
Cuadro 13. Promedio de ddd1 y ddd2.	106
Cuadro 14. Resultados de la detección con láseres.	107
Cuadro 15. Métricas del modelo YOLOv5 entrenado.	107

ÍNDICE DE FIGURAS

Figura 1. Ejemplos de tareas de reconocimiento visual.....	52
Figura 2. Campo experimental: Universidad Autónoma Chapingo.	55
Figura 3. Diseño experimental: Cultivo y sus 32 puntos de muestreo.	56
Figura 4. Ubicación de la cámara para la adquisición de imágenes durante los experimentos.	56
Figura 5. La ubicación de la cámara para la adquisición del conjunto de datos con diferentes condiciones de luz y etapas de crecimiento del maíz.....	57
Figura 6. Visión general del proyecto para la detección de maíz y malezas.....	59
Figura 7. Resumen del proceso propuesto para el etiquetado de máscaras. ...	63
Figura 8. Transformación de algunas muestras de maíz durante el proceso propuesto de etiquetado automático de máscaras.	65
Figura 9. Transformación de algunas muestras de malezas durante el proceso propuesto de etiquetado automático de máscaras.	66
Figura 10. Resumen del proceso de etiquetado de segmentación de instancias semiautomático propuesto.	67
Figura 11. Arquitectura propuesta: Fase de entrenamiento.	69
Figura 12. Arquitectura propuesta: Fase de inferencia.	69
Figura 13. Detección de maíz y malezas utilizando el modelo entrenado con la arquitectura RetinaNet. La primera fila muestra algunos ejemplos de predicciones para malezas, y la segunda fila presenta algunos resultados para predicciones de maíz.	71
Figura 14. Máscaras predichas utilizando el modelo entrenado con la arquitectura U-Net. La primera columna representa la imagen RGB original; la segunda columna es la etiqueta generada por el generador automático de máscaras propuesto. La tercera columna representa la máscara predicha usando el modelo de U-Net.....	72

Figura 15. Resultados de detección de maíz y malezas utilizando el modelo basado en YOLOv5.	74
Figura 16. Detección de maíz y malezas utilizando el modelo de la arquitectura propuesta, es decir, la arquitectura de cuasi segmentación de instancias.	75
Figura 17. Pasos de la navegación.....	85
Figura 18. Diagramas de carrocería y chasis.....	91
Figura 19. Componentes mecánicos.	92
Figura 20. Neumático agrícola 3.50-4.....	92
Figura 21. Sistema de suspensión patente 4369.....	93
Figura 22. Nodos usados para la navegación de Voltan.....	94
Figura 23. Diagrama de flujo del sistema de aspersión con detección de la presencia de plantas de maíz.	98
Figura 24. Detección de una cabra utilizando el código desarrollado.	100
Figura 25. Movimientos del robot para alcanzar la siguiente fila.....	104
Figura 26. Voltan con el aspersor para la competencia.	107
Figura 27. Concepto de estructura del campo para la tarea de navegación ...	113
Figura 28. Imágenes de ejemplo para la tarea 3.....	117
Figura 29. Participantes del FRE 2023.	119
Figura 30. Terreno de pruebas de la competencia.	119
Figura 31. Zona de pruebas de navegación.....	120
Figura 32. Tarea 2 aspersión.....	120
Figura 33. Revisión de los jueces tarea 2.	121
Figura 34. Tarea 3 reconocimiento de imágenes.....	121
Figura 35. Ejemplo de imágenes usadas en la tarea 3.	122
Figura 36. Tarea 4 navegación con obstáculos.	122
Figura 37. Campo del concurso durante tarea 4.....	123
Figura 38. Certificado de participación FRE 2023.....	124

DEDICATORIA

A mis padres José Luis Sánchez Enríquez y Angelica Chávez Santibáñez que me han guiado y apoyado siempre, una inspiración de superación y valores.

A mis hermanos Elsa, Luis, Ismael y Diego compañeros valiosos en esta vida, sigan adelante.

Para mi abuelita Clara Enríquez que siempre cuidó de mí y de mi familia. Para mi tío Genaro Sánchez Enríquez que siempre me motivó a seguir aprendiendo. Los echo de menos pero espero que hayan encontrado paz.

AGRADECIMIENTOS

Al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) por el apoyo económico para realizar este trabajo de investigación y mis estudios de doctorado.

A la Universidad Autónoma Chapingo por su apoyo para este trabajo académico.

A los miembros de mi comité asesor por su compromiso y muy valiosa guía con el desarrollo de este trabajo de investigación: Dr. Noé Velázquez López, Dr. Agustín Ruíz García y Dra. Raquel Salazar Moreno. Grandes académicos pero también grandes seres humanos.

A la Dra. Karina Ruby Pérez Daniel por sus enseñanzas y valioso aporte a este trabajo.

A todos los compañeros y personas que contribuyeron en este proyecto de alguna u otra manera, gracias por el trabajo en equipo, especialmente a Alan Hernández Mercado, Guillermo García Sánchez y Mónica Elizabeth Berrocal Aguilar por su valioso apoyo.

A Carlos Alberto López Figueroa y Griselda Trinidad Miguel por su contribución en la captura de imágenes y fase de entrenamiento.

DATOS BIOGRÁFICOS



David Iván Sánchez Chávez es originario del municipio de Alpatláhuac en el estado de Veracruz, México. Curso su educación básica en las escuelas de este municipio: la primaria “Emiliano Zapata” y la telesecundaria “Independencia”. En el año 2006 fue aceptado en la preparatoria Agrícola de la UACH, donde culminó su preparación media superior en el año 2009. Se incorporó a los estudios de licenciatura en el departamento de Ingeniería Mecánica Agrícola de la Universidad Autónoma Chapingo, los cuales concluyó en el año 2013. Obtuvo su título profesional en el año 2015, en el cual ingresó al programa de Maestría en Ingeniería y Uso integral del Agua en la línea de investigación de mecanización agrícola, durante este periodo trabajó en el desarrollo de un vehículo terrestre no tripulado para labores agrícolas, con este proyecto formó parte del equipo que participó por primera vez en el Field Robot Event, competencia internacional de robots para la agricultura, celebrado en Inglaterra en el 2017 fungiendo como capitán. Durante el doctorado ha colaborado en el desarrollo de prototipos como la sembradora electrónica programable para el mismo robot, ha trabajado con sensores para invernaderos, redes neuronales y desarrollo en ROS entre otros temas.

RESUMEN GENERAL

SISTEMA AUTOMÁTICO CON VISIÓN ARTIFICIAL PARA DETECCIÓN DE MALEZAS EN MAÍZ

Entre los desafíos clave en el cultivo de maíz está la interferencia de malezas que compiten por luz, agua, nutrientes y espacio, lo que conduce a pérdidas en la producción. En este trabajo se propuso desarrollar un sistema de detección de malezas en campos de maíz utilizando visión artificial, mediante modelos de aprendizaje profundo y navegación autónoma de un robot para tareas de monitoreo. Para lograrlo se utilizó un conjunto variado de imágenes de maíz y malezas en diversas etapas de crecimiento, ángulos de visión, condiciones climáticas y de iluminación. Posteriormente se propuso una metodología que combina detección de objetos y segmentación semántica procesando imágenes en U-Net y YOLOv5. Además, se probaron métodos de deep learning para compararlos (RetinaNet, U-Net, YOLOv5). Se desarrolló un algoritmo de navegación que integra sensores, como encoder, IMU, RPLIDAR y cámaras en Ubuntu 18.04 con ROS Melodic, para un vehículo terrestre no tripulado (VTNT) creado en la Universidad Autónoma Chapingo, lo que permitió un movimiento autónomo entre hileras de maíz. La metodología propuesta dio los mejores resultados con los siguientes valores de precisión media (mAP): mAP = 0.936 para maíz, mAP = 0.909 para maleza y mAP = 0.9225 para todas las clases, lo que muestra una alta precisión de detección. Además se lograron eficiencias en tiempo de inferencia y tamaño del modelo con 0.0196 segundos y 94.9 MB respectivamente. El Sistema integrado en ROS representó un avance importante en el movimiento autónomo del robot por los campos de maíz. La implementación de este sistema tiene potencial como alternativa para combatir los arvenses cercanos a las plantas cultivadas en la producción de maíz. No obstante, es necesario perfeccionar la adaptabilidad del sistema a las variaciones del entorno mediante la integración de más sensores y la actualización de las librerías y entorno de desarrollo usados.

Palabras clave: Redes neuronales convolucionales, Navegación autónoma, detección de maleza, detección de maíz, visión artificial

Tesis de Doctorado en Ingeniería Agrícola y Uso Integral del Agua, Universidad Autónoma Chapingo

Autor: David Iván Sánchez Chávez

Director de Tesis: Agustín Ruiz García

ABSTRACT

AUTOMATIC SYSTEM WITH MACHINE VISION FOR WEED DETECTION IN CORN

One of the key challenges in corn cultivation is the interference from weeds that compete for light, water, nutrients, and space, leading to losses in production. This study proposed the development of a weed detection system in cornfields using computer vision, through deep learning models and autonomous navigation of a robot for monitoring tasks. A diverse set of images of corn and weeds at various growth stages, viewing angles, weather, and lighting conditions was utilized. Subsequently, a methodology that combines object detection and semantic segmentation using U-Net and YOLOv5 models was proposed. In addition, deep learning methods such as RetinaNet, U-Net, and YOLOv5 were tested for comparison. A navigation algorithm that integrates sensors, such as an encoder, Inertial Measurement Unit (IMU), RPLIDAR, and cameras on Ubuntu 18.04 with ROS Melodic, for an unmanned ground vehicle (UGV) was developed at Chapingo Autonomous University, enabling independent movement between corn rows. The proposed methodology yielded the best results with the following mean Average Precision (mAP) values: mAP = 0.936 for corn, mAP = 0.909 for weeds, and mAP = 0.9225 for all classes, showing high detection accuracy. Efficiencies in inference time and model size were also achieved with 0.0196 seconds and 94.9 MB, respectively. Integrating the system into ROS marked a significant advance in the robot's autonomous movement through the cornfields. Implementing this system has potential as an alternative to combat weeds close to cultivated plants in corn production. Nonetheless, it is necessary to further refine the system's adaptability to environmental variations by integrating more sensors and updating the utilized libraries and development environment.

Keywords: Convolutional neural networks, autonomous navigation, weeds detection, corn detection, machine vision

PhD. Thesis, Ingeniería Agrícola y Uso Integral del Agua, Universidad Autónoma Chapingo

Author: David Iván Sánchez Chávez

Thesis advisor: Agustín Ruiz García

1. INTRODUCCIÓN GENERAL.

Debido al incremento de la población mundial se requiere de nuevas tecnologías y automatización de procesos para hacer frente al aumento en la demanda de alimentos. Esto significa que los campos tienen que aumentar su producción, sobre todo de granos que son la base de la alimentación de muchas sociedades a nivel mundial. Según las Naciones Unidas (ONU,2023) la población mundial alcanzó los 8000 millones de personas a mediados de noviembre de 2022 y se estima que aumentará casi 2000 millones en los próximos 30 años, pasando de los 8000 millones actuales a los 9700 millones en 2050, pudiendo llegar a un pico de cerca de 10.400 millones para mediados de 2080.

La migración de las personas a ciudades u otros países buscando mejores oportunidades de trabajo, deja las regiones productoras con menos fuerza de trabajo. Otro fenómeno es el envejecimiento de las personas que se quedan a trabajar en el campo, quienes no pueden realizar las actividades requeridas en la agricultura. Según el último censo agropecuario del Instituto Nacional de Estadística y Geografía (INEGI, 2023) en México la gente dedicada al sector agropecuario que tiene 65 años o más es el 29.3%, la gente que tiene entre 45 y 64 años es el 43.5%. También tenemos una disminución de las tierras disponibles para cultivar alimentos. Esto debido al cambio de uso del suelo por fenómenos como la urbanización, aunque también existe a veces un mal manejo de los recursos naturales y un mal manejo de cultivo, lo que nos deja con tierras menos fértiles.

Hay labores agrícolas que necesitan la participación de muchas personas o de personal capacitado como son los operadores de maquinaria, ejemplo de ellas son el monitoreo de los cultivos y la cosecha de algunos productos agrícolas, la escasez de personal afecta más en estos casos.

Muchas de las labores que se realizan son pesadas e incluso peligrosas para la integridad de las personas, un ejemplo es la aplicación de agroquímicos, la

exposición larga a estos productos puede causar un daño severo en la salud de las personas.

Por otro lado incluir el uso de sensores nos permite tener disponible gran cantidad de información como el bienestar de las plantas, sus necesidades nutrimentales, de agua o de condiciones del ambiente en las que se encuentran. Todo esto permite tener un mayor control en los sistemas de producción y tomar mejores decisiones en el manejo del cultivo. Si se saben las necesidades del cultivo en tiempo real, el estado del suelo, la presencia o falta de plantas, de plagas o malezas se pueden utilizar mejor los recursos naturales e insumos como el riego, los fertilizantes, los pesticidas y herbicidas.

Las tecnologías actuales nos permiten realizar operaciones agrícolas con mucha precisión y con una buena calidad de trabajo. Además de que si tenemos suficiente información el sistema puede tomar las decisiones que consideremos más adecuadas de manera rápida. Los sistemas inteligentes actuales pueden detectar anomalías en el campo en donde a una persona le llevaría mucho tiempo, especialmente los robots agrícolas, una de las tecnologías en adopción, pueden trabajar jornadas más largas, lo que libera a las personas de las labores más pesadas y peligrosas, dejándolas con el análisis de información y la programación de la toma de decisiones.

A nivel mundial, existe un gran interés por el desarrollo de nuevas tecnologías que puedan ser aplicadas al campo. El avance tecnológico ha permitido que los costos de los equipos sean más accesibles, además de que la apertura a la información a nivel global aumenta la adopción de diferentes métodos y diferentes técnicas que se aplican en los cultivos.

Una de las labores más importantes en la agricultura es la eliminación de malezas que evita la competencia por agua, nutrientes y luz entre el cultivo y otras plantas silvestres que afecta el desarrollo de las plantas cultivadas y por lo tanto el rendimiento. Además de lo anterior, las malezas sirven de hospedero para varias

plagas que a la larga dañan el cultivo. La labor de deshierbe se puede hacer de manera manual o mecanizada, de manera mecánica o usando agroquímicos.

Los herbicidas son la primer herramienta usada para combatir las malezas, pero estos presentan riesgos secundarios para el hombre y la naturaleza. Se necesita entonces pensar en maneras de reducir sus efectos (Jabir et al., 2021). Una opción pueden ser los sistemas inteligentes que detecten malezas automáticamente aplicando los productos solo en áreas infectadas.

En la actualidad se desarrollan sistemas automáticos de detección de malezas que permiten mejorar las labores, aplicando menores cantidades de herbicida o eliminando de manera más precisa los arvenses, si las malezas son identificadas en una etapa temprana esto lleva a reducir los costos de herbicidas (Espejo-García et al., 2020). Estos sistemas emplean diferentes sensores como pueden ser cámaras que tomen imágenes en tiempo real del cultivo, los sensores usados en drones son categorizados principalmente en tres clases dependiendo de la amplitud y número espectral que pueden grabar: sensores RGB (red, green, blue) o VIS (visible), sensores multiespectrales, sensores hiperespectrales (Esposito et al., 2021).

RGB o VIS. Son las cámaras más comunes y ampliamente disponibles de manera comercial. Sus posibles aplicaciones han sido el foco de la mayoría de las investigaciones por años debido a su potencial y bajo costo (Torres-Sánchez et al., 2014).

Sensores multiespectrales. Son usados para un rango más amplio de cálculos de índices de vegetación ya que pueden depender de un número más alto de bandas radiométricas.

Sensores hiperespectrales. Pueden grabar cientos de miles de estrechas bandas radiométricas, usualmente en rangos visibles e infrarrojos. Estas cámaras pueden estar montadas en vehículos no tripulados ya sea terrestres (UGV) o aéreos (UAV).

Uso de UAVs. En el procesamiento de imágenes vehículos aéreos no tripulados (UAVs) son usados para monitorear cultivos y capturar las posibles malezas en los campos. Los UAVs tienen uso en la agricultura debido a su capacidad de cubrir grandes áreas en una cantidad de tiempo pequeña, y al mismo tiempo no causan compactación del suelo o daño en los campos (Herrmann et al., 2020). Aun así la interpretación de los datos recolectados por UAVs en información de importancia es aún una labor tediosa. Esto porque la recolección y clasificación de datos de manera estándar requiere importantes cantidades de esfuerzo manual.

Uso de ugvs. Robots o vehículos terrestres no tripulados (UGV por sus siglas en inglés) han sido desarrollados para desempeñar varias labores agrícolas tales como cosecha, identificación de enfermedades, detección de malezas, poda, aplicación de pesticidas y también siembra de semillas. Los vehículos no tripulados han sido implicados también para la recopilación de datos y el transporte de carga desde la granja agrícola (Ma y Qi, 2017).

Al establecer los sensores y la plataforma adecuada para la toma de datos, lo siguiente es conseguir diferenciar entre las plantas cultivadas y la maleza, para ello se han empleado diferentes técnicas de visión por computadora.

Visión por computadora abreviado como CV, por sus siglas en inglés, es definido como el proceso de analizar imágenes y videos automáticamente para obtener inferencias de importancia o medidas sin intervención humana. Es un campo multidisciplinario que involucra inteligencia artificial y conceptos de aprendizaje de máquina. El objetivo de la visión por computadora es entender el contenido de imágenes y videos (Patricio y Rieder, 2018).

El aprendizaje de máquina o ML por sus siglas en inglés, es una subrama de la inteligencia artificial, usa un enfoque auto aprendido para dar significado a datos presentados. En vez de crear reglas manualmente a través de analizar grandes cantidades de datos, el aprendizaje de máquina gradualmente mejora su

desempeño de predicción a través de capturar información en los datos (Ünal, 2020).

El aprendizaje profundo o DL por sus siglas en inglés, es un tipo de aprendizaje de máquina que utiliza principios de redes neuronales artificiales. La principal característica por la cual las redes de aprendizaje profundo se distinguen de las redes neuronales artificiales es su profundidad o el número de capas ocultas y esa característica las hace capaces de descubrir estructuras latentes dentro de datos sin estructurar y sin etiquetar. Las redes de aprendizaje profundo, las cuales no necesitan intervención humana mientras realizan extracción automática de características, tienen una ventaja significativa sobre previos algoritmos (Ünal, 2020).

Las técnicas de aprendizaje profundo aplicadas a fotografías digitales pueden ayudar a distinguir entre cultivo y maleza más allá de las limitaciones del procesamiento de imágenes convencional. Una red neuronal convolucional profunda (DCNN) es un tipo de inteligencia artificial que es extensivamente utilizada en años recientes. En sus inicios una red neuronal convolucional (CNN) lo máximo que podía hacer era reconocer dígitos escritos a mano. En el presente, los modelos de DCNN son el aparato más excéntrico en visión por computadora capaces de analizar enormes y complejos bancos de datos, teniendo límites computacionales elevados (Subeesh et al., 2022).

Para poder hacer la detección de manera automática se necesita también una plataforma móvil que pueda hacer recorridos sola de manera periódica, para esto se debe trabajar en la navegación de vehículos autónomos. La navegación de estos vehículos es posible debido a la inclusión de múltiples sensores que recopilan información del campo de maíz, y le permiten moverse en los espacios adecuados.

En este trabajo se desarrolló un sistema de visión artificial para la detección de malezas en el maíz, el cual se probó con los métodos de redes neuronales convolucionales: RetinaNet, U-Net y YOLO. Al final se terminó integrando

YOLOv5 y combinando los enfoques de detección de objetos y segmentación automática de U-Net. La otra etapa del sistema fue la integración de los sensores: IMU, encoder, RPLidar 2d, y cámaras para la navegación de un vehículo terrestre no tripulado desarrollado en la UACH, esto para que el vehículo pueda realizar los trabajos de monitoreo y en una etapa posterior la eliminación automática de las malezas.

1.1 Justificación

El aumento poblacional y la migración generan una sociedad en donde habrá más habitantes en el medio urbano que el rural. Esto por supuesto tendrá profundas consecuencias en la agricultura como la escasez de mano de obra y aumento en la demanda de alimentos. Una de las labores agrícolas más importantes es la eliminación de malezas que se puede hacer de manera mecánica, química o manual pero hay que considerar que los métodos manuales de deshierbe o aplicación de herbicidas ponen a los trabajadores en situaciones de riesgo debido a lo pesado de las labores o a la naturaleza de los productos que se utilizan. Además, en la aplicación de productos químicos no siempre se emplean las cantidades adecuadas, su uso desmedido repercute en las condiciones ambientales que rodean nuestro cultivo y trae consecuencias al suelo. Actualmente se está tratando de utilizar cada vez menos agroquímicos y se están prohibiendo algunos productos considerados altamente nocivos. Una alternativa para combatir estos problemas es el desarrollo de máquinas autónomas que sean capaces de realizar las labores agrícolas y mejorar la productividad, que incluyan sistemas inteligentes que detecten malezas automáticamente aplicando los productos solo en áreas infestadas o eliminándolos con otros métodos como los mecánicos.

1.2 Objetivos

1.2.1 Objetivo General

Diseñar y programar un sistema de detección de malezas en maíz para un vehículo autónomo.

1.2.2 Objetivos Específicos

- Crear un banco de imágenes de maíz y malezas.
- Procesar el banco de imágenes con segmentación y etiquetado.
- Implementar arquitecturas de redes neuronales convolucionales usando el banco de imágenes.
- Diseñar y programar un sistema de navegación autónoma en la siembra de maíz para que el vehículo pueda realizar el monitoreo.
- Diseñar y programar un sistema de visión artificial que sea capaz de diferenciar a las malezas del maíz.

2. REVISIÓN DE LITERATURA

2.1 Identificación automática de malezas

En los últimos años la agricultura de precisión avanzó debido a las innovaciones tecnológicas en las áreas de sensores, hardware de cómputo, sistemas de vehículos no tripulados y robots que permitan la identificación específica de malezas presentes en el campo. Para el reconocimiento de imágenes con sistemas inteligentes se ha recurrido al uso de diferentes aplicaciones de aprendizaje de máquina, campo que ha tenido importantes avances en los años recientes. Estos avances sumados al acceso más amplio a equipos computacionales más potentes y a menor precio han contribuido a la exploración de estas técnicas. Los métodos empleados son muy amplios, algunos trabajos han utilizado procesamiento de imágenes a un nivel menos profundo recurriendo a la visión por computadora "clásica", mientras que otros han entrado de lleno a la combinación con los hallazgos más recientes de inteligencia artificial. En algunos trabajos revisados se observan también combinaciones entre los métodos de aprendizaje de máquina más clásicos y las técnicas de aprendizaje profundo.

Los métodos de detección de malezas en campo usando visión por computadora incluyen principalmente procesamiento de imágenes tradicional y aprendizaje profundo. Cuando se realiza la detección de malezas con tecnología de procesamiento de imágenes tradicional, se extraen características tales como color, textura y forma de la imagen y se combinan con métodos de aprendizaje de máquina tradicional tales como bosque aleatorio o algoritmo de support vector machine (SVM) (Sabzi et al., 2020). Estos métodos necesitan diseñar las características manualmente y tienen alta dependencia en métodos de adquisición de imágenes, métodos de preprocesamiento y la calidad de la extracción de características. Con el avance en el poder de cómputo y el incremento en el volumen de datos, los algoritmos de Aprendizaje profundo pueden extraer información de características semánticas multiescala y multidimensional espacial de las malezas a través de redes neuronales

convolucionales (CNNs) debido a sus capacidades mejoradas de expresión de datos para imágenes, evitando las desventajas de los métodos tradicionales de extracción. Por lo tanto, han llamado cada vez más la atención de los investigadores (Wu et al., 2021).

El aprendizaje de máquina es una colección de algoritmos y herramientas por las cuales máquinas pueden entender patrones dentro de los datos y desempeñar razonamiento acerca de una tarea específica. Tiene la capacidad de extraer información importante de los datos usando varios algoritmos. El aprendizaje profundo puede ser considerado como una siguiente frontera del aprendizaje de máquina; es un subconjunto del aprendizaje de máquina que hace amplio uso de redes neuronales artificiales. El advenimiento del aprendizaje profundo ha traído una revolución en el área de análisis de imágenes y visión por computadora (Chandel et al., 2021; Hemanth y Estrela, 2017).

A continuación se agrupan en dos corrientes los trabajos de aprendizaje de máquina y aprendizaje profundo dependiendo de los métodos empleados para la identificación de las malezas, los trabajos que emplean una combinación de métodos se han considerado como de aprendizaje profundo si incluyen el uso de algún clasificador de este tipo, mientras que si no lo han usado se toma como aprendizaje de máquina tradicional aunque combinen varias técnicas.

2.2 Identificación de malezas con métodos de aprendizaje de máquina tradicional

Los métodos de detección de malezas basados en aprendizaje de máquina tradicional y Aprendizaje profundo tienen distintas ventajas, aun cuando hayan surgido en diferentes periodos de tiempo. Los métodos tradicionales de detección de malezas requieren tamaños de muestra pequeños, tienen bajos requerimientos en unidades de procesamiento de gráficos y pueden ser usado en maquinaria y equipo agrícola a un bajo costo (Wu et al., 2021).

Entre los métodos de clasificación empleados con más frecuencia para identificación de malezas se encuentran bosque aleatorio (RF), support vector machine (SVM) y vecinos más cercanos k (KNN).

Clasificador de bosque aleatorio (RF). Breiman (2001) definió el RF como un ensamble de árboles de clasificación, donde cada árbol de decisión emplea un subconjunto de muestras de entrenamiento y variables seleccionadas por un enfoque de embolsado, mientras las muestras restantes son usadas para validación cruzada interna de le desempeño del RF. El clasificador elige las clases de membresía teniendo el mayor número de votos para un caso dado.

Support vector machine (SVM) clasifica puntos de datos basado en hiperplanos los cuales de manera óptima diferencian las clases basado en los datos de entrenamiento (Bakhshipour y Jafari, 2018). Estos hiperplanos son las superficies definidas por combinaciones de características de entrada.

K vecinos más cercanos (KNN) es un algoritmo no paramétrico popularmente usado para regresión y clasificación. Las entradas consisten en ejemplos de entrenamiento de la k más cercana en el espacio de características (Feyaerts y Van Gool, 2001).

Entre los trabajos que han utilizado estos métodos se encuentra el de Ahmad et al. (2018) que emplearon un clasificador basado en adaBoost con Naive Bayes y fueron capaces de discriminar dos especies de malezas obteniendo una precisión de 0.984.

Le et al. (2020) combinaron características de patrones binarios locales basados en máscaras con características filtradas y un coeficiente k, logrando la clasificación de cultivo y malezas usando support vector machine, el método alcanzó una precisión de 0.9863.

Para diferenciar las regiones de zanahoria y maleza en un campo de cultivo Saha et al. (2017) usaron tres etapas: segmentación de imágenes, extracción de características y clasificación. Para la primera etapa usaron k-means clustering,

en la segunda etapa usaron para extraer la información de la región de interés histograma de gradiente orientado (HoG) y en la etapa de clasificación se empleó support vector machine. El método probó ser efectivo clasificando las regiones con un 0.92 de éxito.

Lin et al. (2017) diferenciaron malezas de plantas de maíz usando imágenes hiperespectrales, emplearon para extraer las características más importantes un sistema de clasificación orientado a objetos con segmentación y árbol de decisiones, extrajeron características de forma y textura y después de tipo espectral. Después con estas características construyeron un modelo de clasificación en plantas de maíz y maleza, obteniendo una precisión global arriba de 0.95 esto basado en el algoritmo C 5.0 el cual construye arboles de clasificación.

Kounalakis et al. (2018) a través de un proceso exhaustivo de optimización, consiguieron crear un sistema de reconocimiento de maleza que usa una entrada de imágenes de resolución 200x150, usa SURF para extracción de características un modelo basado en codebook combinado con codificación Fisher, usa un clasificador lineal para los vectores resultantes. Este sistema experimental consiguió una precisión de reconocimiento del 0.8909 en el dataset examinado.

Rozenberg et al. (2021) emplearon un UAV para obtención de imágenes de 11 campos comerciales de cebolla, las cuales se usaron para mapear la localización de malezas en los campos, los autores emplearon dos algoritmos de clasificación, máxima verosimilitud (ML) y máquinas de vectores de soporte (SVM), todos los porcentajes de precisión (OA) fueron arriba del 0.85, el algoritmo ML dio mejores puntuaciones que el de SVM.

Una combinación de árbol de decisiones y técnicas de lógica difusa se usó para desarrollar un modelo difuso que diferencie plantas de cacahuete de malezas, esto por Bakhshipour et al. (2020). Características de color y de textura fueron extraídas de imágenes de cacahuete y tres de sus malezas comunes, se

aplicaron dos técnicas de selección de características, análisis de componentes principales y selección de características basada en correlación. Las cuatro categorías de plantas fueron cacahuete, Hoja de terciopelo (*Abutilon theophrasti*), Margarita falsa (*Eclipta alba*) y Nicandra (*Nicandra physalodes*). Se obtuvieron de las combinaciones de estas técnicas precisiones de 0.8083, 0.80 y 0.7917.

Se realizó un trabajo de identificación de plantas de soja y malezas asociadas a su cultivo por Tang et al. (2018), esto utilizando k-medias después del preprocesamiento de los datos para construir un diccionario de características que después fue utilizado para extraer características de datos etiquetados y entrenar modelos de clasificación para identificación automática de malezas. Se obtuvieron precisiones de identificación para cinco malezas y la soja, usando clasificadores basados en SVM y regresión soft-max con una y dos capas. Las precisiones totales para la clasificación fueron para SVM con una capa de 0.8824, para SVM de dos capas de 0.9026, para clasificador de regresión soft-max de una capa 0.8922, para regresión soft-max de dos capas de 0.9154.

En España se realizó un trabajo de mapeo de malezas en girasoles por Pérez-Ortiz et al. (2015) utilizando información obtenida con un vehículo aéreo no tripulado a través de dos cámaras, una cámara multiespectral de seis bandas, y una cámara de apunta y dispara. Las imágenes se tomaron a tres distintas alturas, 30, 60 y 100 metros. Las imágenes primero se unieron en un solo mosaico para después dividirlo en imágenes de aproximadamente 1000 píxeles por 1000 píxeles. Los píxeles se categorizaron en suelo, cultivo y maleza. Se emplearon y se compararon métodos sin supervisión, semisupervisados y supervisados. Para la identificación de las malezas emplearon seis métodos: k-medias, repetead k-means, SVM semisupervisado, knn, linear SVM y kernel SVM. Los resultados reportan el mean average error (MAE) el cual tiene valores de 0.1962 para k-means, 0.1912 para rk-means, 0.1854 para SVM, 0.1697 para KNN, 0.1333 para linSVM y 0.1268 para ss-SVM.

Se probaron cuatro algoritmos independientes de clasificación para la creación de mapas de maleza en campos de avena, combinando métodos manuales y

automáticos, esto por parte de Gašparović et al. (2020). los datos utilizados provinieron de una cámara rgb montada en un UAV, el terreno se dividió en dos grupos. Los algoritmos de clasificación fueron basados en random forest para extracción de maleza y suelo, se hicieron automático basado en objetos, automático basado en píxeles, manual basado en objetos y manual basado en píxeles. Siguiendo una clasificación no supervisada con k-means para estimar la presencia de suelo o maleza. De los cuatro algoritmos de clasificación empleado el método con mejor precisión fue el método de clasificación automático basado en objetos con una precisión de 0.89 para el grupo A y 0.871 para el grupo B.

Pérez-Ortiz et al. (2016) utilizaron imágenes RGB de un UAV tomadas en campos de maíz y girasol a una altitud de 30 m para mapear las malezas presentes, esto en España. Primero se hizo la segmentación de las imágenes y se etiquetaron los objetos de interés bajo la supervisión de un experto, después se extrajeron diferentes características que serían usadas para la clasificación con métodos de aprendizaje de máquina. Para el análisis de las imágenes y el proceso de segmentación se utilizó el conocido método de k-medias y un paso para refinar objetos (para evitar objetos pequeños. Para la clasificación se usó máquinas de vectores de soporte (SVM). Las precisiones obtenidas en sus experimentos van desde 0.6167 a 0.955.

Alexandridis et al. (2017) estudiaron la detección y el mapeo de la maleza *Silybum marianum* (L.) Gaertn usando clasificadores de aprendizaje de máquina. Para la obtención de las imágenes se empleó una cámara multiespectral (verde-rojo-NIR) a bordo de un UAV, esto en Thessaloniki, Grecia. Los cuatro clasificadores empleados en el trabajo fueron One Class Support Vector Machine (OC-SVM), One Class Self-Organizing Maps (OC-SOM), Autoencoders y One Class Principal Component Analysis (OC-PCA). Como características de entrada se usaron las tres bandas espectrales y la textura. La identificación de la planta de interés alcanzó una precisión general de 0.96 usando OC-SVM, 0.9465 usando OC-SOM, 0.943 para autoencoder y 0.90 para OC-PCA.

En el estudio en Irán en cultivos de cacahuate de Bakhshipour (2021) se emplearon dos métodos de impulso para mejorar el desempeño de cuatro clasificadores, utilizando imágenes obtenidas con un smartphone a 40 cm del suelo con una resolución de 1344 x 2240 píxeles. Los dos métodos son Adaboost.M1 y LogitBoost, mientras que los cuatro clasificadores son Perceptron multicapa (MLP), k vecinos más cercanos (kNN), bosque aleatorio (RF) y máquinas de vectores de soporte (SVM). Se emplearon también cuatro técnicas de filtrado de características Correlation-based Feature Selection (CFS), Information Gain (IG), Gain Ratio (GR), y OneR, con ellas se seleccionaron 10 características más significativas para alimentar los clasificadores impulsados. La combinación de IG-RF fue el clasificador más apropiado entre los modelos evaluados, también se comprobó que las técnicas de impulso tienen un efecto positivo en el desempeño de los clasificadores. La combinación de técnicas más apropiada para la discriminación de plantas en campos de cacahuate fue la de LogiBoost-IG-RF, la cual obtuvo una precisión de 0.9958, mientras que en el dataset de prueba obtuvo 0.95.

De castro et al. (2018) establecieron un experimento en España en tres campos de girasol y dos campos de algodón de los cuales obtuvieron imágenes con un UAV a 30 y 60 m de altura con una cámara que capturó tanto luz visible como NIR. Estas imágenes fueron utilizadas para entrenar algoritmos de clasificación y usarlos en identificación de malezas en los campos. Combinaron modelos de superficie digital (DSMs), ortomosaicos, y la técnica de aprendizaje de máquina bosque aleatorio (RF). Se usó bosque aleatorio como clasificador que identifica las malezas y como etapa final se crearon mapas con la ubicación de las malezas en los campos. La precisión de identificación de maleza para los cinco campos tomados en cuenta fue de 0.84, 0.63, 0.591, 0.879 y 0.811.

En su trabajo De Rainville et al. (2014) en Quebec Canadá se enfocaron en la distinción de maleza y cultivo en campos de soya y maíz usando visión por computadora y análisis morfológico. La novedad en las técnicas usadas está en el proceso de extracción de características que está basado en localización

espacial de vegetación en campos. Son extraídas características del área de distribución de las hojas de maleza, después características del cultivo son inferidas por el modelo mixto. Esas características son después pasadas al clasificador naive bayesiano y al algoritmo gaussiano para diferenciar maleza de planta de cultivo. La técnica que emplearon clasifica un promedio de 0.94 de plantas de cultivo y 0.85 de maleza.

Cuadro 1. Trabajos que utilizaron aprendizaje de máquina clásico.

Autores	Cultivo y malezas	Obtención de imágenes	Métodos de clasificación	Resultado
Ahmad et al. (2018)	Ne	Rgb	adaBoost, Naive Bayes	Precisión de 0.984
Le et al. (2020)	Ne	Rgb	SVM	Precisión de 0.9863
Saha et al. (2017)	zanahoria	Rgb	SVM	Precisión de 0.92
Lin et al. (2017)	maíz	Imágenes hiperespectrales	Arboles de clasificación, 5.0	Precisión global 0.95
Kounalakis et al. (2018)	Ne	Rgb	Clasificador lineal	Precisión de 0.8909
Rozenberg et al. (2021)	cebolla	UAV rgb	ML, SVM	Precisión OA de 0.85
Bakhshipour et al. (2020)	cacahuete, abutilon theophrusti, Eclipta alba y Nicandra physalodes	Rgb	Árbol de decisiones, lógica difusa	Precisión de 0.8083
Tang et al. (2018)	soya	Rgb	svm y regresión soft-max	Precisión de 0.9154 para soft-max de dos capas mean average error (MAE) de 0.1268 para ss-SVM.
Pérez-Ortiz et al. (2015)	girasol	UAV con una cámara multiespectral de seis bandas, y una cámara de apunta y dispara	k-medias, repetead k-means, SVM semisupervisado, knn, linear SVM y kernel SVM	Precisión de 0.9154 para soft-max de dos capas mean average error (MAE) de 0.1268 para ss-SVM.

Gašparović et al. (2020)	avena	UAV rgb	Random forest, k-means	Precisión de 0.89
Pérez-Ortiz et al. (2019)	maíz, girasol	UAV rgb	SVM	Precisión de 0.955
Alexandridis et al. (2017)	Silybum marianum (L.) Gaertn	cámara multispectral (verde-rojo-NIR) a bordo de un UAV	One Class Support Vector Machine (OC-SVM), One Class Self-Organizing Maps (OC-SOM), Autoencoders y One Class Principal Component Analysis (OC-PCA)	Precisión global de 0.96 usando OC-SVM
Bakhshipour (2021)	cacahuete	Smartphone rgb	MLP, kNN, RF, SVM	Precisión de 0.95 para LogiBoost-IG-RF
De castro et al. (2018)	girasol, algodón	UAV con cámara que capturó tanto luz visible como NIR	RF	Precisión de 0.879
De Rainville et al. (2014)	Soya, maíz	Rgb	naive bayesiano y algoritmo gaussiano	Precisión de 0.94 de plantas de cultivo y 0.85 de maleza.

Ne – no especificado

2.3 Identificación de malezas con métodos de Aprendizaje profundo

Aprendizaje profundo se refiere a métodos involucrados con las arquitecturas de las redes neuronales artificiales. Una red neuronal artificial es un sistema diseñado para asemejarse a la organización neuronal del cerebro humano. Una definición más compleja sería que una red neuronal es un modelo computacional hecho de neuronas artificiales conectadas unas con otras resultando en una arquitectura de red. Esta arquitectura tiene parámetros específicos llamados pesos, los cuales se modifican a través del algoritmo de retropropagación de errores hasta que el valor predicho por la red se acerque al valor real. Ajustándolos podemos mejorar la precisión de nuestro modelo. Este tipo de redes

contienen muchas capas, cada una con una misión específica. Su número determina la complejidad de la red (Jabir et al., 2021).

Uno de los tipos de redes neuronales empleados por los autores revisados en el trabajo de identificación de malezas son las redes neuronales recurrentes, estas trabajan bajo el principio de guardar la salida de una capa y alimentar esta a las entradas para ayudar a predecir la salida de la capa, el trabajo de Jabir et al. (2021) es un ejemplo donde se usó este tipo de red.

Otro de los tipos de red son las redes neuronales convolucionales (RNC o CNN), las cuales son redes neuronales artificiales usadas para reconocimiento y procesamiento de imágenes (Duda et al., 2020). Una CNN es una poderosa inteligencia artificial (AI) especialmente pensada para manejar información de píxeles, un sistema de procesamiento de imágenes que emplea aprendizaje profundo para desempeñar labores generativas y descriptivas. Comúnmente usa visión por computadora, lo cual incluye reconocimiento de imagen y video, así como sistemas de recomendación y procesamiento de lenguaje natural (Zhang et al., 2019).

Las redes neuronales convolucionales (CNN) son una clase de redes neuronales profundas de feedforward que han sido efectivamente aplicadas a visión por computadora. En contraste con una red neuronal artificial, la cual puede ser inviable en algunos problemas de gran alcance, una CNN puede aprender problemas complejos rápidamente debido a pesos compartidos y las capas más complejas utilizadas (Jabir et al., 2021).

La arquitectura de una CNN incluye capa de convolución, capa pooling, capa dropout, y una capa completamente conectada. Una CNN de arquitectura profunda puede ser construida combinando varias capas convolucionales, capas pooling, capas dropout y capa completamente conectada (Liang et al., 2019).

La mayoría de las investigaciones en la detección automática de malezas han usado arquitecturas populares de CNN como AlexNet, VGG, Inception-ResNet,

o combinaciones de CNNs con otros procedimientos. Los estudios que tienen pequeños bancos de datos para entrenar sus modelos de redes han usado el poder de la aumentación de datos para incrementar artificialmente el número de imágenes de entrenamiento y de esta manera mejorar la precisión. Se usa traslaciones, transposiciones y reflexiones así como modificar las intensidades de los canales rgb, la mayoría de los trabajos incluyen pasos con preprocesamiento de imágenes (Jabir et al., 2021).

Jabir et al. (2021) compararon redes neuronales convolucionales (CNN) y redes neuronales recurrentes (RNN) para detección de maleza y concluyeron que las CNN ofrecen mejores resultados en esta labor agrícola, lo cual se ve reflejado en la popularidad con la que cuentan estas en los trabajos revisados.

Ahmad et al. (2021) emplearon modelos preentrenados para clasificar imágenes de malezas encontradas en campos de maíz y soya, usaron VGG16, ResNet50 e inceptionV3. Después usaron un modelo de detección de objetos basado en la librería de YOLOv3, y lo entrenaron para identificar diferentes especies de malezas en una imagen. El mejor modelo de clasificación de imágenes fue VGG16 con una precisión de 0.989.

Otros que usaron en sus experimentos de identificación de malezas CNNs fueron Andrews et al. (2019). Ellos tomaron imágenes con una cámara RGB de 5 MP e identificaron tres plantas diferentes cuyos nombres en inglés son los siguientes: ladies finger, brinjal y Ragi. Obtuvieron precisiones muy buenas que son para ladies finger 1, para brinjal 0.875 y para Ragi 1.

Badhan et al. (2021) usaron dos conjuntos de datos uno de datos uno de un cultivo de calabaza y otro de cebolla, usaron una red neuronal convolucional propia y también el modelo de ResNet-50 el cual ofreció mayor precisión en la identificación de malezas, 0.846 para los datos de calabaza y 0.90 para los datos de la cebolla. Para crear los datos para el entrenamiento extrajeron cuadros de video con openCV y etiquetaron de manera manual los datos. Esta labor es

necesaria para que en la imagen se tenga la ubicación de las plantas ya sea maleza o cultivo.

Algunos trabajos se han realizado con componentes sencillos y económicos como es el caso del uso de una red de arquitectura VGG16 por Jogi et al. (2020) usando una cámara web y una raspberry pi para obtener imágenes en un cultivo de raji.

Khan et al. (2021) usaron imágenes de UAV de alta resolución (la cámara de un equipo DJI Spark) sobre dos campos uno de chícharo y uno de fresa ubicados en Pakistan, consiguieron detectar maleza con una precisión de 0.953, mientras que la precisión global (maleza y cultivo) fue de 0.9473. Emplearon Residual Network 101 (ResNet-101) que tiene 101 capas.

Farooq et al. (2019a) utilizaron imágenes hiperespectrales (un sensor de 61 bandas) de cuatro malezas nombradas: Hyme, Alli, Azol, y Hyac, esto para probar una CNN y compararla con el desempeño de histogramas de gradientes orientados (Hog) en Matlab obteniendo mejores resultados con la CNN, esto con valores de 0.8631, 0.9061 y 0.9472 en las pruebas de CNN para diferentes tamaños de batch, por otro lado Hog dio precisiones de 0.6819, 0.7178 y 0.7434.

Liang et al. (2019) en Taiwán propusieron un sistema de identificación de malezas de bajo costo usando imágenes rgb tomadas por drones como datos para entrenamiento y aplicando redes neuronales convolucionales (CNN) para construir el modelo de identificación. Compararon diversos métodos para la clasificación de malezas entre ellos local binary pattern (LBP) y histogram of oriented gradient (HOG). Su sistema identifica malezas con una precisión de hasta 0.988.

Fu et al. (2020) plantearon un algoritmo de reconocimiento de malezas basado en el VGG llamado VGG inception. Redujeron el número de capas convolucionales para reducir los parámetros de la red, agregaron la estructura inception que puede mantener las características principales y tiene mejor

precisión de clasificación, por último usaron métodos de aumento y transferencia para prevenir el problema del overfitting. Usaron el conjunto de datos the kaggle. Obtuvieron precisiones de más de 0.98 en la detección de malezas, en un campo real puede alcanzar 0.80.

Potena et al. (2017) lograron que un UGV equipado con una cámara multispectral hiciera detección automática de malezas en tiempo real, esto empleando imágenes RGB+near infrared (NIR) que funcionan como entrada a dos CNN, la primera realiza segmentación de la imagen para dejar solo los pixeles que pertenecen a vegetación posteriormente la segunda CNN clasifica los pixeles en las categorías de cultivo y maleza. Lograron obtener en la detección de plantas con la primera red 0.968 de precisión, mientras que la detección de maleza combinando las dos redes obtuvo 0.926.

Peteinatos et al. (2020) emplearon imágenes RGB obtenidas de las especies *Zea mays*, *Helianthus annuus*, *Solanum tuberosum*, *Alopecurus myosuroides*, *Amaranthus retroflexus*, *Avena fatua*, *Chenopodium album*, *Lamium purpureum*, *Matricaria chamomila*, *Setaria spp.*, *Solanum nigrum* y *Stellaria media* para entrenar CNNs (VGG16, ResNet-50, and Xception) en un total de 93000 imágenes que contenían una sola especie por imagen. Obtuvieron una precisión de entre 0.77 a 0.98 en la detección y discriminación de las diferentes especies vegetales.

Sabzi et al. (2020) usaron la red neuronal artificial -algoritmo cultural (ANN-CA) la cual es un híbrido. El algoritmo cultural desempeña acciones de optimización similares a los algoritmos genéticos, solo que se toma en cuenta la evolución cultural y el espacio social. El objetivo final del CA es encontrar las elites y su desarrollo para complemento cultural. Para su entrenamiento extrajeron de imágenes de papa y cinco malezas *Malva neglecta*, *Portulaca oleracea*, *Chenopodium album L*, *Secale cereale L* y *Xanthium strumarium* características de color, textura, forma. De estas se seleccionaron seis: desviación estándar de saturación componente (S) en el espacio HSV, la diferencia entre el primer y séptimo momentos invariantes, valor medio del componente (H) en el espacio de

color HSI, relación área a longitud, componente (Cb) en el espacio de color YCbCr y desviación estándar de componente (I) en el espacio de color YIQ. La clasificación dio una precisión de 0.98 identificando la papa de las especies de maleza.

Tamouridou et al. (2017) identificaron maleza de *Silybum marianum* usando la arquitectura de red neuronal perceptrón multicapa con determinación automática (MLP-ARD), alimentando con datos obtenidos de imágenes de una cámara multispectral (verde-rojo-infrarrojo cercano) montada en un UAV, las entradas fueron las tres bandas espectrales y la capa de textura. La red alcanzó una precisión de 0.9954.

Sharpe et al. (2019) usaron tres arquitecturas de red neuronal convolucional (detectnet, vggnet y googlenet) para la identificación de la maleza *Geranium carolinianum* presente en cultivos de fresa en cuatro sitios de Florida, Estados Unidos, usando imágenes recortadas de la maleza para entrenamiento de las redes obtuvieron una detección con Fscore=0.77 para la vggnet, Fscore=0.62 para la GoogLeNet, y un Fscore=0.94 para DetectNet.

Haq (2021) utilizó para identificación de maleza una red neuronal convolucional que incluye el método learning vector quantization como una segunda capa totalmente conectada, el cual llamó modelo CNN LVQ. Trabajó con 4400 imágenes obtenidas con un UAV en las cuales identificó las categorías de suelo, soya, pasto, y latifolia. Obtuvo precisiones de 1 para suelo, 0.9979 para soya, 0.9858 para pasto y 0.9832 para latifolia.

Sharpe et al. (2020) utilizaron una red convolucional, yolov3, para identificar vegetación (una categoría) y entre tres tipos de vegetación (tres categorías) presente en cultivos de tomate y fresa establecidos en Florida, Estados Unidos. La red de una sola categoría presentó un Fscore=0.93 mientras que la red que distinguía entre tres categorías (hojas anchas, juncias y pastos) presentó un Fscore=0.95. Se demostró que el uso de YOLOv3 es una opción para identificación de malezas.

Dasgupta et al. (2020) utilizaron en su proyecto imágenes obtenidas por un UAV y para la detección de malezas una combinación de redes neuronales convolucionales que incluye inception v3, y yolov3, este trabajo se enfocó en la predicción de cultivo logrando una precisión de 0.8929.

Gao et al. (2020) emplearon una red neuronal convolucional (CNN) basada en la arquitectura de tiny YOLOv3 para detección de *C. sepium* y remolacha azucarera. Combinaron imágenes sintéticas con imágenes de campo. El modelo resultante fue probado en 100 imágenes de campo mostrando que la combinación de imágenes sintéticas e imágenes de campo para entrenar el modelo pudieron mejorar la precisión media (mAP) de 0.751 a 0.829 comparado con el uso de solamente imágenes de campo. Se compararon también el desempeño del modelo con los modelos YOLOv3 y tiny YOLO donde el modelo que propusieron obtuvo mejores precisiones de 0.761 para *C. sepium* y 0.897 para la remolacha azucarera.

Se realizó un trabajo en huertos de naranja para identificar malezas, esto por parte de Torres-Sospedra y Nebot (2014). Consistió en dos etapas, en la primera se determinaron características de las imágenes de los huertos (árboles, troncos, suelo y cielo). En la segunda etapa se detectaron las malezas solo en las áreas que fueron identificadas como suelo. Debido a los cambios ambientales durante la identificación de malezas (clima y condiciones de luz), se usó un entrenamiento que incluía patrones de ruido para las redes neuronales. Se utilizaron tres arquitecturas de redes neuronales las cuales obtuvieron desempeños de 0.9405, 0.9501 y 0.9466.

El trabajo realizado por Lottes et al. (2020) estima la localización del tallo de las malezas, lo que permite que un robot agrícola desempeñe tratamiento mecánico preciso, al mismo tiempo provee el área en pixeles cubierta por malezas. El método que utilizaron fue una red convolucional Fully Convolutional Network (FCN) que desempeña tanto la estimación de cobertura de las malezas como la posición del tallo, utilizaron las clases cultivo, maleza dicotiledónea, maleza pasto y suelo. La precisión de la ubicación de tallo varía para diferentes terrenos en los

que se hicieron pruebas llegando valores máximos de 0.949, en el caso de la segmentación del área de malezas alcanzó 0.935. Cabe mencionar que para algunas categorías la precisión alcanzo valores muy bajos de menos de 0.20.

Se realizó un trabajo de identificación de malezas en campos de césped por Yu et al. (2019). En este se emplearon tres modelos de redes neuronales convolucionales profundas, DetectNet, GoogLeNet y VGGNet. Se emplearon para ello imágenes obtenidas con una cámara digital a resolución de 1920 x 1080 pixeles. Se entrenaron las redes de dos maneras, primero para identificación de malezas y césped de manera individual, una especie de maleza y césped fueron las categorías, y la segunda manera fue poniendo como categorías todas las especies de maleza encontradas y el césped. Las precisiones de los clasificadores en la validación para una sola especie fueron las siguientes: GoogLeNet para *Hydrocotyle* spp. 0.5, GoogLeNet *Hedyotis corymbosa* 0.5753, GoogLeNet *Richardia scabra* 0.5, VGGNet *Hydrocotyle* spp. 1, VGGNet *Hedyotis corymbosa* 0.9901, VGGNet *Richardia scabra* 0.9823. Las precisiones para múltiples especies al mismo tiempo en la validación fueron: GoogLeNet *Hydrocotyle* spp. 0.5470, GoogLeNet *Hedyotis corymbosa* 0.6468, GoogLeNet *Richardia scabra* 0.5484, VGGNet *Hydrocotyle* spp. 0.9746, VGGNet *Hedyotis corymbosa* 0.9381, VGGNet *Richardia scabra* 0.9124. Las precisiones para malezas creciendo en césped bermuda latente fueron: DetectNet *Poa annua* 1, GoogLeNet *Poa annua* 0.5, VGGNet *Poa annua* 0.9360. Para *Poa annua* y otras especies de malezas con hoja ancha: DetectNet 1, GoogLeNet 0.5, VGGNet 0.5397. Puede observarse que los mejores resultados en las primeras dos categorías fueron dados por VGGNet, y en la tercera categoría por DetectNet.

Rasti et al. (2019) utilizaron el enfoque de scattering transform para clasificación de imágenes de maleza y cultivo en cultivos de alta densidad (mache). Esta técnica se ha mostrado que supera el desempeño de enfoques mono escala (gray-level co-occurrence matrix, local binary patterns), mostraron un enfoque para el diseño óptimo de scatter transform basado en contraste de energía. Las

imágenes fueron tomadas por un robot agrícola y el clasificador entrenado solo en datos sintéticos muestra una precisión del 0.85 probado en datos reales.

Farooq et al. (2019b) utilizaron un método de extracción de características usando redes neuronales convolucionales (CNN) y patrón binario local basado en super pixel (LBP). Las características espaciales de mediano y alto nivel son aprendidas usando la CNN. Después características de textura local son extraídas con LBP y usadas como entrada para vectores de máquinas de soporte (SVM) que clasifican la maleza. El método combinado utilizado alcanzó mejores valores de precisión que el uso de las técnicas por separado, la precisión general en el primer dataset fue de 0.8970 y en el segundo dataset de 0.9635.

Tannouche et al. (2022) evaluaron la habilidad del aprendizaje profundo y las redes neuronales convolucionales para detectar maleza en varios cultivos usando imágenes obtenidas por una cámara colocada en aspersores jalados por un tractor para poder usar aspersion localizada y ultralocalizada de herbicida en la región de Beni Mellal en Marruecos, etiquetaron las imágenes en dos categorías: imagen con maleza e imagen libre de maleza. Se estudió la detección de malezas a través de seis CNNs recientes conocidas por su velocidad y precisión, estas fueron VGGNet (16 y 19), GoogLeNet (Inception V3 y V4) y MobileNet (V1 y V2). El primer experimento se realizó con las arquitecturas de las CNNs desde cero y el segundo experimento se realizó con versiones preentrenadas. Los resultados mostraron que Inception V4 obtuvo la precisión más alta con 0.9941 para la versión desde cero y 0.9951 para la versión preentrenada, mientras que la más rápida y ligera con 14 MB fue MobileNet V2.

Fawakherji et al. (2021) emplearon métodos para aumentar las imágenes disponibles para entrenar redes de clasificación de malezas y cultivo, esto a través de Generative Adversarial Networks (GANs), donde escenas enteramente sintéticas son agregadas a los datos de entrenamiento. Iniciando de imágenes reales, crearon muestras semiartificiales reemplazando las clases de objetos más relevantes (cultivo, maleza) con su contraparte sintetizada. Además de datos RGB utilizaron información del infrarrojo cercano (NIR), generando imágenes

sintéticas multiespectrales de cuatro canales. Demostraron que su modelo es capaz de generar este tipo de imágenes mencionadas y que el uso de dichas imágenes sintéticas mejora el desempeño de redes convolucionales de segmentación semántica. Las redes tomadas en cuenta fueron Bonnet, Unet-Resnet, U-Net y segNet, los mejores resultados se dieron en los datasets que combinaban imágenes sintéticas con las imágenes originales, siendo la precisión más alta para identificación de cultivo de 0.97 empatando UNET-ResNet, U-Net y SegNet, y la más alta para maleza de 0.72 correspondiente a U-Net.

Dos Santos Ferreira et al. (2017) emplearon redes neuronales convolucionales para detección de maleza en cultivo de soya, clasificando las malezas en pasto y hoja ancha, con el objetivo de emplear el herbicida específico para cada maleza. Para esto se empleó un campo de soya en Brasil en el cual se tomaron imágenes juntando una base de datos de más de quince mil imágenes del suelo, soya, maleza de hoja ancha y de pasto. Como red neuronal convolucional se empleó CaffeNet que es una réplica de la conocida AlexNet. Para comparar resultados se emplearon también SVM, AdaBoost y bosque aleatorio en conjunto con una colección de técnicas de extracción de características de forma, color y textura. Como resultado su trabajo consiguió para las convnets una precisión arriba de 0.98.

Ante el problema de identificación de malezas en el cultivo de trigo Xu et al. (2021) propusieron un framework basado en información multimodal para maleza pasto. Primero se recodificó la imagen de canal sencillo en una nueva imagen de tres canales como la estructura de imagen RGB, la cual es adecuada para extracción de características de redes neuronales convolucionales (CNN). En segundo lugar se realiza la detección de objeto multiescala a través de fusionar los mapas de características por diferentes capas convolucionales. La estructura de red de tres canales está diseñada para tomar en cuenta la independencia de RGB e información profunda, y la información multimodal complementaria. Los resultados experimentales mostraron que comparado con el método de detección de malezas basado en imágenes RGB, la precisión del método propuesto se

mejora. Para malezas pasto se tuvo una precisión media promedio (mAP) de 0.361 y para las malezas de hojas anchas 0.429, la precisión general indicada por IoG es de 0.893.

Wang et al. (2022) trabajaron en China con imágenes rgb y multiespectrales obtenidas por un UAV para proponer una red neuronal basada en optimización biónica para detectar la densidad de la maleza y el crecimiento del cultivo. Esta usa la red AlexNet preentrenada para transferencia de aprendizaje, esto porque la tasa de aprendizaje de la capa agregada es difícil de afinar, el peso y la tasa de aprendizaje de sesgo de la nueva capa completamente conectada es elegido con optimización de enjambre de partículas (PSO) y algoritmo murciélago (BA) para encontrar la combinación óptima en el pequeño dataset. Los datos son transportados a la red neuronal convolucional (CNN) recolectando imágenes RGB y multiespectrales de 5 bandas de 3 tipos de maleza (*Chenopodium album*, *Humulus Scandens* y *Xanthium sibiricum*) y 3 tipos de cultivo (maíz, cacahuate y trigo) como bases de datos. Finalmente, se implementan 6 tipos de clasificaciones. Al mismo tiempo, una CNN basada en model-agnostic meta-learning (MAML) se propone para realizar el aprendizaje de redes neuronales artificiales con pocas muestras y alta eficiencia, y se verifica su precisión en el grupo de prueba. Las redes neuronales optimizadas por dos algoritmos biónicos de optimización son comparadas con la CNN basada en MAML. Los resultados muestran que la combinación de tasa de aprendizaje a través de BA es la mejor, y su precisión puede alcanzar 0.9939 para imágenes RGB, 0.9953 para imágenes multiespectrales y 0.9602 para una pequeña muestra de 6-shot.

Cuadro 2. Trabajos que utilizaron aprendizaje profundo.

Autores	Cultivo malezas	y Obtención de imágenes	Métodos de clasificación	Resultados destacados
Ahmad et al. (2021)	maíz y soya	RGB	VGG16, ResNet50, inceptionV3, YOLOv3	precisión de 0.989 con VGG16

Andrews et al. (2019)	ladies finger, brinjal y Ragi	RGB		CNN propia	Precisión de 1 para ragi
Badhan et al. (2021)	calabaza, cebolla	RGB		CNN propia, ResNet-50	Precisión de 0.90 para cebolla con ResNet-50
Khan et al. (2021)	chícharo, fresa	UAV rgb		ResNet-101	Precisión global de 0.9473
Farooq et al. (2019)	hyme, Alli, Azol, y Hyacs	Imágenes hiperespectrales		CNN propia, Hog	Precisión de 0.9472 para CNN
Liang et al. (2019)	ne	UAV rgb		CNN propia, Hog, LBP	Precisión de 0.988
Fu et al. (2020)	ne	RGB		VGG inception	Precisión de 0.98
Potena et al. (2017)	ne	UGV cámara RGB+NIR	con	CNN propia	Precisión de 0.968
Peteinatos et al. (2020)	Zea mays, Helianthus annuus, Solanum tuberosum, Alopecurus myosuroides, Amaranthus retroflexus, avena fatua, Chenopodium album, Lamium purpureum, Matricaria chamomila, Setaria spp., Solanum nigrum y Stellaria media	RGB		VGG16, ResNet-50, Xception	Precisión de 0.77 a 0.98 para diferentes especies vegetales
Sabzi et al. (2020)	papa, malva neglecta, Portulaca	RGB		ANN-CA	Precisión de 0.98

	oleracea, Chenopodium album L, Secale cereale L y Xanthium strumarium				
Tamourido et al. (2017)	Silybum marianum	UAV con cámara multispectral	MLP-ARD	precisión de 0.9954	
Sharpe et al. (2019)	fresa, Geranium carolinianum	RGB	detectnet, vggnet y googlenet	Fscore=0.94 para DetectNet	
Haq (2021)	soya, pasto, y latifolia	UAV rgb	CNN LVQ	Precisión de 0.9979 para soya	
Sharpe et al. (2020)	tomate, fresa	RGB	YOLOv3	Fscore=0.95	
Dasgupta et al. (2020)	ne	UAV rgb	Combinación de inception v3 y yolov3	precisión de 0.8929	
Gao et al. (2020)	c. sepium, remolacha azucarera	RGB	tiny YOLOv3	mAP de 0.897 para remolacha azucarera	
Torres-Sospedra y Nebot (2014)	naranja	RGB	ANN	Precisión de 0.9501	
Lottes et al. (2020)	ne	RGB	FCN	Precisión de 0.949	
Yu et al. (2019)	césped	RGB	DetectNet, GoogLeNet y VGGNet	Precisión de 0.9901 para VGGNet Hedyotis cormybose	
Rasti et al. (2019)	mache	Robot agrícola RGB	Scattering transform	precisión del 0.85	
Farooq et al. (2019)	ne	RGB	CNN, LBP, SVM	Precisión de 0.9635	

Tannouche et al. (2022)	ne	RGB	VGGNet (16 y 19), GoogLeNet (Inception V3 y V4), MobileNet (V1 y V2)	precisión de 0.9951 para Inception V4
Fawakherji et al. (2021)	ne	RGB, NIR	Bonnet, Unet-Resnet, U-Net y segNet	Precisión de 0.97
Dos Santos Ferreira et al. (2017)	soya	RGB	CaffeNet, SVM, AdaBoost, bosque aleatorio	Precisión de 0.98 para CaffeNet
Xu et al. (2021)	trigo, pasto	RGB	CNN propia	Precisión de 0.893
Wang et al. (2022)	maíz, cacahuete y trigo, Chenopodium album, Humulus Scandens, Xanthium sibiricum	UAV, imágenes RGB y multiespectrales	AlexNet, CNN basada en MAML	Precisión de 0.9953 para AlexNet con imágenes multiespectrales

ne – no especificado

En este campo este trabajo propone la detección de malezas en el maíz utilizando dos categorías: maleza y maíz, para eso se combinó el enfoque de detección de objetos y segmentación semántica usando YOLOv5 y U-net, para el entrenamiento de las redes neuronales convolucionales se utilizó imágenes con diversas etapas de crecimiento de las plantas, distintas perspectivas, distintas condiciones de luz y condiciones de tiempo atmosférico lo que busca que el sistema sea más robusto y se adapte mejor a las condiciones variantes de una siembra de maíz real. Además de esto se buscó que el modelo tuviera un tiempo de inferencia bajo y tamaño compacto para hacerlo más adecuado para una implementación en tiempo real.

2.4 Navegación de vehículos en campos agrícolas

Los vehículos terrestres no tripulados tienen una gran importancia para la agricultura 4.0 y sus alcances, existe una tendencia a investigar el desarrollo de sistemas robóticos para la agricultura. Esto ha permitido que hoy en día tengamos vehículos más autónomos e inteligentes que realizan diversas actividades, algunas muy específicas. Si los robots tienen la capacidad de detectar el estado en el que están las plantas, sus necesidades y las condiciones de su ambiente pueden hacer un mejor uso de todos los recursos que necesitamos en la producción agrícola y esto puede terminar en una mayor eficiencia.

Actualmente tenemos una gran cantidad de robots para labores agrícolas en desarrollo y algunos ya en el mercado que realizan labores tan diversas como trasplantar, podar, desmalezar, cosechar, plantar, rociar agroquímicos o fertilizar. Además la mayoría de estos mecanismos que se utilizan en la agricultura son de tipo eléctrico por lo tanto la adopción de estos robots nos permite reducir las emisiones de CO₂ aportando a sistemas de producción sostenibles, que ayuden a reducir la contaminación y combatir los problemas del cambio climático.

Para la navegación autónoma, podemos denotar dos posicionamientos. El primero es el posicionamiento global del robot, el cual se plantea hacerlo mediante GPS, esto nos permitirá llevar el robot a la parcela y una vez ahí el robot tiene que pasar a combinar esto con su posición local, la cual obtiene utilizando la información de los sensores que tiene a bordo. La combinación de estas dos cosas nos permite realizar labores agrícolas de manera más precisa.

Los dos métodos más empleados para detectar las líneas de un cultivo y poder navegar en medio de él es el uso de los a rplidar principalmente 3d y el uso de la visión artificial. Un lidar de alta calidad tiene ventajas sobre las cámaras debido a las características de los entornos agrícolas, donde suele haber variaciones en las condiciones de iluminación. Si se quiere emplear cámaras, entonces que tienen la ventaja de ser más baratas, es necesario combatir los cambios de tonalidades y de niveles de contraste a los que se va a encontrar el robot a

diferentes horas del día. A continuación veremos algunos sensores utilizados en vehículos agrícolas según el análisis de Fasiolo et al. (2023).

2.4.1 Sensores utilizados para navegación en vehículos para uso agrícola

Los sistemas de navegación de diversos robots agrícolas utilizan principalmente sensores de posicionamiento global, de medición inercial, lasers rotativos RPLiDAR y cámaras de distintos tipos.

Posicionamiento global

De acuerdo a Fasiolo et al. (2023) los sensores de este tipo se fundamentan en las tecnologías GNSS. Actualmente, los sistemas de localización que operan a través de satélites se han consolidado y su costo se ha reducido, permitiendo que los receptores GNSS sean ampliamente utilizados en distintas aplicaciones agrícolas, como en la navegación autónoma de tractores (Yamasaki, Morie, y Noguchi, 2022). Específicamente, el GNSS puede determinar la posición absoluta tridimensional del receptor (por ejemplo, en relación con el sistema geodésico mundial WGS84) al calcular la distancia a no menos de cuatro satélites. El Sistema de Posicionamiento Global (GPS), gestionado por el Departamento de Defensa de los Estados Unidos, es la constelación GNSS más reconocida y utilizada (Jackson, Davis, y Gebre-Egziabher, 2018; Guo et al., 2018). Sin embargo, una limitación notable del GNSS es su menor fiabilidad en zonas con densa vegetación. Esto se debe a la propagación multipath de las señales, que afecta la precisión posicional, especialmente en comparación con áreas de cielo abierto (Kaartinen et al., 2015).

Adicionalmente, al instalar tres antenas en una estructura móvil, se puede deducir la orientación del robot en seis grados de libertad (DOF) con una frecuencia de actualización aproximada de 1 Hz (Hirokawa y Ebinuma, 2009). Según estudios como el de Kragh et al. (2017), es viable estimar la dirección del vehículo usando técnicas que involucran múltiples receptores GNSS integrados.

Medición inercial

En aplicaciones agrícolas, se utiliza con frecuencia un IMU (Unidad de medida inercial) de 9 grados de libertad (DOF) para derivar la "attitude" (roll, pitch, yaw) del robot, según lo discutido por Lan et al. (2019). Este dispositivo combina un acelerómetro triaxial (que mide aceleraciones lineales), giroscopios (que registran la tasa angular) y un magnetómetro (usualmente utilizado como referencia de dirección).

A pesar de que los sensores IMU pueden emitir datos a una velocidad de hasta 1 kHz, en sistemas que se desplazan lentamente se suele optar por frecuencias más bajas con el objetivo de reducir la deriva, como se menciona en Manish et al. (2021) y Lowe et al. (2021). Gracias a su tamaño compacto, bajo peso y eficiente consumo de energía, los IMU son ideales para ser integrados en plataformas robóticas pequeñas.

Sin embargo, hay desafíos a considerar: al instalar un IMU de 9-DOF, es fundamental ser consciente de que campos magnéticos perturbados localmente pueden causar inexactitudes en la orientación que el magnetómetro detecta. La magnitud de esta perturbación depende de factores como la posición relativa y la orientación del equipo de medición con respecto a elementos ferromagnéticos, imanes permanentes y corrientes eléctricas intensas, tal como detallan Vitali et al. (2020). Además, es crucial realizar una calibración extrínseca, que es el proceso mediante el cual se alinean y registran los marcos de referencia. Esta calibración tiene como objetivo determinar la posición del IMU en relación al centro de rotación del robot o a cualquier otro sensor (ejemplo, LiDAR) con el que esté asociado el IMU. Si no se lleva a cabo esta calibración, las mediciones de posición, velocidad y aceleración podrían resultar inconsistentes, lo que las vuelve inadecuadas para el acoplamiento del sensor.

Sensores LiDAR

Los sensores LiDAR (Light Detection and Ranging), se utilizan en vehículos autónomos para la agricultura para obtener información geométrica de la

vegetación, permitiendo la reconstrucción del mapa del entorno circundante en forma de nube de puntos. Los dispositivos LiDAR operan midiendo el tiempo que tarda un pulso láser en ser emitido por el sensor, reflejarse en el entorno y regresar al instrumento, utilizando para ello sistemas microelectromecánicos o matrices de fase óptica (Choudhary et al., 2021; Baek e Im, 2022; Garrido et al., 2015).

En cuanto a su clasificación, los sistemas LiDAR pueden ser 2D o 3D. Los LiDAR 2D cuentan con un único anillo láser que realiza mediciones en un plano, mientras que los LiDAR 3D tienen múltiples anillos. Entre las soluciones comerciales más destacadas, el SICK LMS 111 es el dispositivo 2D más utilizado, y el Velodyne VLP16 es el dispositivo 3D predominante.

Los sensores LiDAR poseen un rango de medición considerable. Por ejemplo, el Velodyne VLP16 puede medir hasta 100 m, lo que es útil tanto en ambientes cerrados, como invernaderos, como en escenarios al aire libre, como viñedos y huertos.

Los sensores LiDAR 2D son valiosos para la localización y evitar colisiones, especialmente en presencia de obstáculos dinámicos. Para crear nubes de puntos 3D con estos sensores, deben montarse de forma que el plano láser sea ortogonal a la trayectoria del robot. Sin embargo, los dispositivos LiDAR 3D son más adecuados para la reconstrucción de mapas, ya que ofrecen una nube de puntos más densa y un campo de visión más amplio.

A pesar de sus ventajas, los sensores LiDAR enfrentan desafíos en exteriores, especialmente bajo la lluvia. Las gotas de lluvia pueden interferir con los rayos láser, resultando en nubes de puntos ruidosas e inexactas (Wichmann et al., 2021).

Radar

El radar (radio detection and ranging) es un sistema que utiliza ondas de radio para identificar la distancia, dirección o velocidad de objetos. Entre las ventajas

del radar como herramienta de percepción en robótica, está su desempeño en condiciones adversas. Mientras que el LiDAR puede verse afectado por factores ambientales como el polvo, la niebla, la lluvia y la nieve, el radar es notablemente menos susceptible a estos elementos (Wichmann et al., 2021).

Rouveure et al. (2018) realizaron pruebas con el radar PELICAN, que proporciona una vista panorámica de 360° del entorno en un rango que va desde los 5 hasta los 100 metros. Sin embargo, es esencial señalar que este radar específico solo puede generar imágenes 2D del entorno.

Otra de las ventajas destacadas de los radares sobre otros sistemas de percepción, como las cámaras y los sensores LiDAR, es su capacidad para penetrar eficazmente en la vegetación y detectar objetivos que podrían estar ocultos (Cheng et al., 2023).

Sensores visuales

Estos dispositivos capturan imágenes RGB y son fundamentales para tareas como la localización, evitar obstáculos y mapeo (Durmuş et al., 2015; Barbosa et al., 2019).

Las cámaras monoculares son las más básicas, pero a pesar de su simplicidad, tienen la capacidad de crear reconstrucciones 3D del entorno utilizando técnicas como estructura desde el movimiento (SfM), según se evidencia en el estudio de Sylvain et al. (2015).

Avanzando en complejidad, encontramos las cámaras estéreo que, con sus dos sensores de imagen, simulan la visión binocular humana para percibir profundidad. En escenarios con baja textura, su precisión puede mejorarse con patrones infrarrojos (IR). El trabajo de Aghi et al. (2020) es un claro ejemplo de cómo se pueden utilizar estas cámaras para la navegación autónoma en viñedos.

Las cámaras de profundidad son dispositivos como el Kinect V2, ofrecen valores de rango utilizando luz estructurada o sensores de tiempo de vuelo (ToF). Las

aplicaciones de estas cámaras se reflejan en trabajos como los de Rosell-Polo et al. (2017) y Matsuzaki et al. (2018), que las emplean para distintas tareas de percepción.

Existen también sistemas de múltiples cámaras. Estos dispositivos, presentes en estudios como (Ball et al., 2016) y (Vulpi et al., 2022), combinan varias cámaras calibradas externamente para ofrecer un campo de visión más amplio. Ejemplos actuales incluyen la Giroptic 360CAM (Kragh et al., 2017) y la Ladybug3 (Underwood et al., 2017), que brindan visiones panorámicas de 360°.

A pesar de sus capacidades, es vital recordar las limitaciones de estos sensores. Principalmente en entornos exteriores, enfrentan desafíos en condiciones adversas o con baja iluminación, como señala Patrício et al. (2018). Además, el rango de medición de las cámaras ToF es más limitado en comparación con tecnologías como el LiDAR.

Sensores espectrales

En algunos trabajos se han utilizado este tipo de sensores para la detección de vegetación y utilizar esta información para la navegación aunque, en general, estos sensores brindan información más detallada del estado del cultivo.

Los sensores espectrales juegan un papel esencial en la agricultura de precisión, permitiendo la monitorización de la salud vegetal al analizar la respuesta espectral de la vegetación a la radiación electromagnética (Bannari et al., 1995). Estos sensores no solo se centran en el espectro visible, sino que también examinan la diferencia en reflectancia entre las bandas rojas y cercanas al infrarrojo (NIR), así como la reflectancia espectral en la zona de transición, también conocida como "red edge" (French et al., 2016). Las cámaras espectrales específicas se diseñan para capturar estas longitudes de onda, siendo útiles para calcular índices de vegetación, como el NDVI (Pettorelli et al., 2005).

En cuanto a la navegación de vehículos agrícolas, los sensores espectrales tienen el potencial de mejorar la detección de obstáculos, ya que las cámaras multiespectrales pueden identificar obstáculos en el campo que no son fácilmente visibles en el espectro visible. Además, pueden guiar a vehículos agrícolas hacia áreas específicas del campo que requieren atención basada en la salud vegetal, optimizando así la aplicación de insumos.

Entre las aplicaciones prácticas, los robots móviles equipados con cámaras multiespectrales se han utilizado en viñedos para generar imágenes NDVI (Clamens et al., 2021) y en campos agrícolas para detectar plantas y evaluar el estado de los cultivos (Fawakherji et al., 2021; Chebrolu et al., 2017). Las cámaras hiperespectrales, con una resolución radiométrica más alta que las multiespectrales, se han empleado en huertos y campos de cultivo (Cubero et al., 2020). Además, se ha demostrado que las cámaras digitales RGB estándar pueden adaptarse para capturar el canal NIR, proporcionando resultados equivalentes a las cámaras multiespectrales en la monitorización de la salud de los cultivos (Cubero et al., 2020).

2.5 Literatura citada

Aghi, D., Mazzia, V., & Chiaberge, M. (2020). Local motion planner for autonomous navigation in vineyards with a RGB-D camera-based algorithm and Deep learning synergy. *Machines*, 8(2), 27.

Ahmad, A., Saraswat, D., Aggarwal, V., Etienne, A., Hancock, B. (2021). Performance of deep learning models for classifying and detecting common weeds in corn and soybean production systems. *Computers and Electronics in Agriculture*. 184. 106081. <https://doi.org/10.1016/j.compag.2021.106081>.

Ahmad, J., Muhammad, K., Ahmad, I., Ahmad, W., Smith, M. L., Smith, L. N., Jain, D. K., Wang, H., Mehmood, I. (2018). Visual features based boosted classification of weeds for real-time selective herbicide sprayer systems. *Computers in Industry*. 98. 23-33. <https://doi.org/10.1016/j.compind.2018.02.005>.

Alexandridis, T. K., Tamouridou, A. A., Pantazi, X. E., Lagopodi, A. L., Kashefi, J., Ovakoglou, G., Polychronos, V., & Moshou, D. (2017). Novelty Detection Classifiers in Weed Mapping: *Silybum marianum* Detection on UAV Multispectral Images. *Sensors*, 17(9). <https://doi.org/10.3390/s17092007>

Andrews, S., Ramesh, K., Nandhakumar, T., Kamalraj, T., & Ragupathi, P. (2019). Plant variety and weed growth identification: Trending towards machine learning.

International Journal of Recent Technology and Engineering, 8(2), 6232-6237.
doi:10.35940/ijrte.B3293.078219

Badhan, S., Desai, K., Dsilva, M., Sonkusare, R., & Weakey, S. (2021). Real-time weed detection using machine learning and stereo-vision. Paper presented at the 2021 6th International Conference for Convergence in Technology, I2CT 2021, doi:10.1109/I2CT51068.2021.9417989

Baek, E.-T., & Im, D.-Y. (2022). ROS-based unmanned mobile robot platform for agriculture. *Appl. Sci.*, 12(9), 4335.

Bakhshipour, A. (2021). Cascading feature filtering and boosting algorithm for plant type classification based on image features. *IEEE Access*, doi:10.1109/ACCESS.2021.3086269

Bakhshipour, A., Zareiforoush, H. (2020). Development of a fuzzy model for differentiating peanut plant from broadleaf weeds using image features. *Plant Methods* 16, 153. <https://doi.org/10.1186/s13007-020-00695-1>

Bakhshipour, A.; Jafari, A. (2018). Evaluation of support vector machine and artificial neural networks in weed detection using shape features. *Comput. Electron. Agric.* 145. 153–160.

Ball, D., Upcroft, B., Wyeth, G., Corke, P., English, A., Ross, P., Patten, T., Fitch, R., Sukkarieh, S., & Bate, A. (2016). Vision-based obstacle detection and navigation for an agricultural robot. *J. Field Robotics*, 33(8), 1107–1130.

Bannari, A., Morin, D., Bonn, F., & Huete, A. (1995). A review of vegetation indices. *Remote Sens. Rev.*, 13(1–2), 95–120.

Barbosa, W.S., Oliveira, A.I., Barbosa, G.B., Leite, A.C., Figueiredo, K.T., Vellasco, M.M., & Caarls, W. (2019). Design and development of an autonomous mobile robot for inspection of soy and cotton crops. In 2019 12th International Conference on Developments in ESystems Engineering (pp. 557–562). IEEE.

Breiman, L. (2001). Random forests. *Mach. Learn.* 45. 5–32.

Chandel, N., Chakraborty, S., Rajwade, Y., Dubey, K., Tiwari, M.K., Jat, D. (2021). Identifying crop water stress using deep learning models. *Neural Comput. Applic.* 33. 5353–5367. <https://doi.org/10.1007/s00521-020-05325-4>.

Chebrolu, N., Lottes, P., Schaefer, A., Winterhalter, W., Burgard, W., & Stachniss, C. (2017). Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields. *Int. J. Robot. Res.*, 36(10), 1045–1052.

Cheng, Y., Changsong, P., Mengxin, J., & Yimin, L. (2023). Relocalization based on millimeter wave radar point cloud for visually degraded environments. *J. Field Robot.* Recuperado de <http://dx.doi.org/10.1002/rob.22162>.

Choudhary, A., Kobayashi, Y., Arjonilla, F.J., Nagasaka, S., & Koike, M. (2021). Evaluation of mapping and path planning for non-holonomic mobile robot navigation in narrow pathway for agricultural application. In 2021 IEEE/SICE International Symposium on System Integration (pp. 17–22).

Clamens, T., Alexakis, G., Duverne, R., Seulin, R., Fauvet, E., & Fofi, D. (2021). Real-time multispectral image processing and registration on 3D point cloud for vineyard analysis. In 16th International Conference on Computer Vision Theory and Applications.

Cubero, S., Marco-Noales, E., Aleixos, N., Barbé, S., & Blasco, J. (2020). Robhortic: A field robot to detect pests and diseases in horticultural crops by proximal sensing. *Agriculture*, 10(7), 276.

Dasgupta, I., Saha, J., Venkatasubbu, P. et al. (2020). AI Crop Predictor and Weed Detector Using Wireless Technologies: A Smart Application for Farmers. *Arab J Sci Eng* 45, 11115–11127. <https://doi.org/10.1007/s13369-020-04928-2>

De Castro, A. I., Torres-Sánchez, J., Peña, J. M., Jiménez-Brenes, F. M., Csillik, O., & López-Granados, F. (2018). An automatic random forest-OBIA algorithm for early weed mapping between and within crop rows using UAV imagery. *Remote Sensing*, 10(2) doi:10.3390/rs10020285

De Rainville, F. -, Durand, A., Fortin, F. -, Tanguy, K., Maldague, X., Panneton, B., & Simard, M. -. (2014). Bayesian classification and unsupervised learning for isolating weeds in row crops. *Pattern Analysis and Applications*, 17(2), 401-414. doi:10.1007/s10044-012-0307-5

Dos Santos Ferreira, A., Matte Freitas, D., Gonçalves da Silva, G., Pistori, H., & Theophilo Folhes, M. (2017). Weed detection in soybean crops using ConvNets. *Computers and Electronics in Agriculture*, 143, 314-324. doi:10.1016/j.compag.2017.10.027

Duda P, Jaworski M, Cader A, Wang L. (2020). On training deep neural networks using a streaming approach. *Journal of Artificial Intelligence and Soft Computing Research*. 10(1):15–26. <https://doi.org/10.2478/jaiscr-2020-0002>.

Durmuş, H., Güneş, E.O., Kırıcı, M., & Üstündağ, B.B. (2015). The design of general purpose autonomous agricultural mobile-robot: "AGROBOT". In 2015 Fourth International Conference on Agro-Geoinformatics (pp. 49–53). IEEE.

Espejo-Garcia, B., Mylonas, N., Athanasakos, L., Fountas, S., (2020). Improving weeds identification with a repository of agricultural pre-trained deep neural networks. *Computers and electronics in Agriculture*. 175, 105593. <https://doi.org/10.1016/j.compag.2020.105593>.

Esposito, M., Crimaldi, M., Cirillo, V., Sarghini, F., Maggio, A. (2021). Drone and sensor technology for sustainable weed management: a review. *Chemical and biological technologies in agriculture*. 8 (1). 8:18.

Farooq, A., Hu, J., & Jia, X. (2019a). Analysis of Spectral Bands and Spatial Resolutions for Weed Classification Via Deep Convolutional Neural Network. *IEEE Geoscience and Remote Sensing Letters*, 16(2), 183–187. <https://doi.org/10.1109/LGRS.2018.2869879>

Farooq, A., Jia, X., Hu, J., & Zhou, J. (2019b). Multi-resolution weed classification via convolutional neural network and superpixel based local binary pattern using remote sensing images. *Remote Sensing*, 11(14) doi:10.3390/rs11141692

- Fasiolo, D. T., Scalera, L., Maset, E., & Gasparetto, A. (2023). Towards autonomous mapping in agriculture: A review of supportive technologies for ground robotics. *Robotics and Autonomous Systems*, 169. <https://doi.org/10.1016/j.robot.2023.104514>
- Fawakherji, M., Potena, C., Pretto, A., Bloisi, D. D., & Nardi, D. (2021). Multi-spectral image synthesis for Crop/Weed segmentation in precision farming. *Robotics and Autonomous Systems*, 146 doi:10.1016/j.robot.2021.103861
- Feyaerts, F.; Van Gool, L. (2001). Multi-spectral vision system for weed detection. *Pattern Recognit. Lett.* 22. 667–674.
- French, A.N., Gore, M.A., & Thompson, A. (2016). Cotton phenotyping with LiDAR from a track-mounted platform. In *Autonomous Air and Ground Sensing Systems* (Vol. 9866). International Society for Optics and Photonics.
- Fu, L., Lv, X., Wu, Q., & Pei, C. (2020). Field Weed Recognition Based on an Improved VGG With Inception Module. *International Journal of Agricultural and Environmental Information Systems (IJAEIS)*, 11(2), 1-13. <http://doi.org/10.4018/IJAEIS.2020040101>
- Gao, J., French, A.P., Pound, M.P. et al. (2020). Deep convolutional neural networks for image-based *Convolvulus sepium* detection in sugar beet fields. *Plant Methods* 16, 29. <https://doi.org/10.1186/s13007-020-00570-z>
- Garrido, M., Paraforos, D.S., Reiser, D., Vázquez Arellano, M., Griepentrog, H.W., & Valero, C. (2015). 3D maize plant reconstruction based on georeferenced overlapping LiDAR point clouds. *Remote Sens.*, 7(12), 17077–17096.
- Gašparović, M., Zrinjski, M., Barković, Đ., & Radočaj, D. (2020). An automatic method for weed mapping in oat fields based on UAV imagery. *Computers and Electronics in Agriculture*, 173 doi:10.1016/j.compag.2020.105385
- Grocholsky, B., Nuske, S., Aasted, M., Achar, S., & Bates, T. (2011). A camera and laser system for automatic vine balance assessment. In *American Society of Agricultural and Biological Engineers* (p. 1). Louisville, Kentucky.
- Guo, J., Li, X., Li, Z., Hu, L., Yang, G., Zhao, C., Fairbairn, D., Watson, D., & Ge, M. (2018). Multi-GNSS precise point positioning for precision agriculture. *Precis. Agric.*, 19(5), 895–911.
- Haq, M. A. (2021). CNN based automated weed detection system using UAV imagery. *Computer Systems Science and Engineering*, 42(2), 837-849. doi:10.32604/csse.2022.023016
- Hemanth, D.J., Estrela, V.V., (2017). *Deep Learning for Image Processing Applications*. IOS Press.
- Herrmann, I.; Bdolach, E.; Montekyo, Y.; Rachmilevitch, S.; Townsend, P.A.; Karnieli, A. (2020). Assessment of maize yield and phenology by drone-mounted superspectral camera. *Precis. Agric.* 21. 51–76.
- Hirokawa, R., & Ebinuma, T. (2009). A low-cost tightly coupled GPS/INS for small UAVs augmented with multiple GPS antennas. *Navigation*, 56(1), 35–44.

INEGI. (2023, septiembre). Resultados del Censo agropecuario México. Recuperado de https://www.inegi.org.mx/contenidos/programas/cagf/2022/doc/CA2022_RONAL.pdf

Jabir, B., Rabhi, L., Falih, N. (2021). RNN- and CNN-based weed detection for crop improvement: An overview. *Foods and raw materials*. vol. 9, no. 2, pp. 387-396. <http://doi.org/10.21603/2308-4057-2021-2-387-396>

Jackson, J., Davis, B., & Gebre-Egziabher, D. (2018). A performance assessment of low-cost RTK GNSS receivers. In *IEEE/ION Position, Location and Navigation Symposium, PLANS* (pp. 642–649).

Jay, S., Rabatel, G., Hadoux, X., Moura, D., & Gorretta, N. (2015). In-field crop row phenotyping from 3D modeling performed using structure from motion. *Comput. Electron. Agric.*, 110, 70–77.

Jogi, Y., Rao, P. N., Raksha, Shetty, S., Shreekari, (2020). CNN based Synchronal recognition of Weeds in Farm Crops. 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA). 1373-1378. doi: 10.1109/ICECA49313.2020.9297449.

Kaartinen, H., Hyypä, J., Vastaranta, M., Kukko, A., Jaakkola, A., Yu, X., Pyörälä, J., Liang, X., Liu, J., Wang, Y., Kaijaluoto, R., Melkas, T., Holopainen, M., & Hyypä, H. (2015). Accuracy of kinematic positioning using global satellite navigation systems under forest canopies. *Forests*, 6(9), 3218–3236.

Khan, S., Tufail, M., Khan, M.T., Khan, Z.A., Anwar, S. (2021). Deep learning-based identification system of weeds and crops in strawberry and pea fields for a precision agriculture sprayer. *Precision Agriculture*. 22 (6). 1711-1727.

Kounalakis, T., Triantafyllidis, G.A. & Nalpantidis, L. (2018). Image-based recognition framework for robotic weed control systems. *Multimed Tools Appl* 77, 9567–9594. <https://doi.org/10.1007/s11042-017-5337-y>

Kragh, M.F., Christiansen, P., Laursen, M.S., Larsen, M., Steen, K.A., Green, O., & Karstoft, H. (2017). Fieldsafe: Dataset for obstacle detection in agriculture. *Sensors*, 17(11), 2579.

Lan, H., Elsheikh, M., Abdelfatah, W., Wahdan, A., & El-Sheimy, N. (2019). Integrated RTK/INS navigation for precision agriculture. In *32nd International Technical Meeting of the Satellite Division of the Institute of Navigation* (pp. 4076–4086).

Le, V. N. T., Ahderom, S., Apopei, B., Alameh, K. (2020). A novel method for detecting morphologically similar crops and weeds based on the combination of contour masks and filtered Local Binary Pattern operators. *GigaScience*. 9(3). g1aa017, <https://doi.org/10.1093/gigascience/g1aa017>

Liang, W. -C., Yang, Y. -J., Chao, C. -M. (2019). Low-Cost Weed Identification System Using Drones. *Seventh International Symposium on Computing and Networking Workshops (CANDARW)*. 260-263. doi: 10.1109/CANDARW.2019.00052.

- Lin, F., Zhang, D., Huang, Y., Wang, X., & Chen, X. (2017). Detection of corn and weed species by the combination of spectral, shape and textural features. *Sustainability (Switzerland)*, 9(8) doi:10.3390/su9081335
- Lottes, P., Behley, J., Chebrolu, N., Milioto, A., & Stachniss, C. (2020). Robust joint stem detection and crop-weed classification using image sequences for plant-specific treatment in precision farming. *Journal of Field Robotics*, 37(1), 20-34. doi:10.1002/rob.21901
- Lowe, T., Moghadam, P., Edwards, E., & Williams, J. (2021). Canopy density estimation in perennial horticulture crops using 3D spinning LiDAR SLAM. *J. Field Robotics*, 38(4), 598–618.
- Ma, K., Qi, Z. (2017). A Human-Centered design of general-purpose unmanned electric vehicle chassis for agriculture task payload. *Journal of Computing and Information Science in Engineering*, JCISE-16-1041, 17(3): 031004.
- Manish, R., Lin, Y.-C., Ravi, R., Hasheminasab, S.M., Zhou, T., & Habib, A. (2021). Development of a miniaturized mobile mapping system for in-row, under-canopy phenotyping. *Remote Sens.*, 13(2), 276.
- Matsuzaki, S., Masuzawa, H., Miura, J., & Oishi, S. (2018). 3D semantic mapping in greenhouses for agricultural mobile robots with robust object recognition using robots' trajectory. In *2018 IEEE International Conference on Systems, Man, and Cybernetics* (pp. 357–362). IEEE.
- ONU. (2023, septiembre). Desafíos globales: población. Recuperado de <https://www.un.org/es/global-issues/population>
- Patricio, D. I., Rieder, R. (2018). Computer vision and artificial intelligence in precision agriculture for grain crops: a systematic review. *Computers and Electronics in Agriculture*. 153. 69–81.
- Pérez-Ortiz, M., Peña, J. M., Gutiérrez, P. A., Torres-Sánchez, J., Hervás-Martínez, C., & López-Granados, F. (2015). A semi-supervised system for weed mapping in sunflower crops using unmanned aerial vehicles and a crop row detection method. *Applied Soft Computing*, 37, 533–544. <https://doi.org/https://doi.org/10.1016/j.asoc.2015.08.027>
- Pérez-Ortiz, M., Peña, J. M., Gutiérrez, P. A., Torres-Sánchez, J., Hervás-Martínez, C., & López-Granados, F. (2016). Selecting patterns and features for between- and within- crop-row weed mapping using UAV-imagery. *Expert Systems with Applications*, 47, 85-94. doi:10.1016/j.eswa.2015.10.043
- Peteinatos, G. G., Reichel, P., Karouta, J., Andújar, D., & Gerhards, R. (2020). Weed identification in maize, sunflower, and potatoes with the aid of convolutional neural networks. *Remote Sensing*, 12(24), 1-22. doi:10.3390/rs12244185
- Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J., & Stenseth, N.C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol. Evol.*, 20(9), 503–510.

Potena, C., Nardi, D., & Pretto, A. (2017). Fast and accurate crop and weed identification with summarized train sets for precision agriculture. *Advances in Intelligent Systems and Computing*, 531, 105-121. doi:10.1007/978-3-319-48036-7_9

Rasti, P., Ahmad, A., Samiei, S., Belin, E., & Rousseau, D. (2019). Supervised image classification by scattering transform with application to weed detection in culture crops of high density. *Remote Sensing*, 11(3) doi:10.3390/rs11030249

Rosell-Polo, J.R., Gregorio, E., Gené, J., Llorens, J., Torrent, X., Arnó, J., & Escola, A. (2017). Kinect V2 sensor-based mobile terrestrial laser scanner for agricultural outdoor applications. *IEEE/ASME Trans. Mechatronics*, 22(6), 2420–2427.

Rouveure, R., Faure, P., & Monod, M. (2018). Description and experimental results of a panoramic k-band radar dedicated to perception in mobile robotics applications. *J. Field Robotics*, 35(5), 678–704.

Rozenberg, G., Kent, R. & Blank, L. (2021). Consumer-grade UAV utilized for detecting and analyzing late-season weed spatial distribution patterns in commercial onion fields. *Precision Agric* 22, 1317–1332. <https://doi.org/10.1007/s11119-021-09786-y>

Sabzi, S., Abbaspour-Gilandeh, Y., & Arribas, J. I. (2020). An automatic visible-range video weed detection, segmentation and classification prototype in potato field. *Heliyon*. 6(5). doi:10.1016/j.heliyon.2020.e03685

Saha, D., Hamer, G., & Lee, J. Y. (2017). Development of inter-leaves weed and plant regions identification algorithm using histogram of oriented gradient and k-means clustering. Paper presented at the Proceedings of the 2017 Research in Adaptive and Convergent Systems, RACS 2017, 2017-January 136-139. doi:10.1145/3129676.3129700

Sharpe, S. M., Schumann, A. W., & Boyd, N. S. (2019). Detection of carolina geranium (*geranium carolinianum*) growing in competition with strawberry using convolutional neural networks. *Weed Science*, 67(2), 239-245. doi:10.1017/wsc.2018.66

Sharpe, S.M., Schumann, A.W., Yu, J. et al. (2020). Vegetation detection and discrimination within vegetable plasticulture row-middles using a convolutional neural network. *Precision Agric* 21, 264–277. <https://doi.org/10.1007/s11119-019-09666-6>

Subeesh, A., Bhole, S., Singh, K., Chandel, N.S., Rajwade, Y.A., Rao, K.V.R., Kumar, S.P., Jat, D. (2022). Deep convolutional neural network models for weed detection in polyhouse grown bell peppers. *Artificial Intelligence in Agriculture*. 6. 47-54. <https://doi.org/10.1016/j.aiia.2022.01.002>.

Tamouridou, A. A., Alexandridis, T. K., Pantazi, X. E., Lagopodi, A. L., Kashefi, J., Kasampalis, D., Kontouris, G., & Moshou, D. (2017). Application of Multilayer Perceptron with Automatic Relevance Determination on Weed Mapping Using UAV Multispectral Imagery. *Sensors*, 17(10). <https://doi.org/10.3390/s17102307>

- Tang, J., Zhang, Z., Wang, D. et al. (2018). Research on weeds identification based on K-means feature learning. *Soft Comput* 22, 7649–7658. <https://doi.org/10.1007/s00500-018-3125-x>
- Tannouche, A., Gaga, A., Boutalline, M., & Belhouideg, S. (2022). Weeds detection efficiency through different convolutional neural networks technology. *International Journal of Electrical and Computer Engineering*, 12(1), 1048-1055. doi:10.11591/ijece.v12i1.pp1048-1055
- Torres-Sánchez J, Peña JM, de Castro AI, López-Granados F. (2014). Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. *Comput Electron Agric.*103:104–13.
- Torres-Sospedra, J., & Nebot, P. (2014). Two-stage procedure based on smoothed ensembles of neural networks applied to weed detection in orange groves. *Biosystems Engineering*, 123, 40-55. doi:10.1016/j.biosystemseng.2014.05.005
- Ünal, Z. (2020). Smart Farming Becomes Even Smarter With Deep Learning—A Bibliographical Analysis. *IEEE Access*. 8. 105587-105609. doi: 10.1109/ACCESS.2020.3000175.
- Underwood, J., Wendel, A., Schofield, B., McMurray, L., Kimber, R. (2017). Efficient in-field plant phenomics for row-crops with an autonomous ground vehicle. *J. Field Robotics*, 34(6), 1061–1083.
- Vitali, R.V., McGinnis, R.S., & Perkins, N.C. (2020). Robust error-state Kalman filter for estimating IMU orientation. *IEEE Sensors J.*, 21(3), 3561–3569.
- Vulpi, F., Marani, R., Petitti, A., Reina, G., & Milella, A. (2022). An RGB-D multi-view perspective for autonomous agricultural robots. *Comput. Electron. Agric.*, 202, 107419.
- Wang, S., Han, Y., Chen, J., He, X., Zhang, Z., Liu, X., & Zhang, K. (2022). Weed density extraction based on few-shot learning through UAV remote sensing RGB and multispectral images in ecological irrigation area. *Frontiers in Plant Science*, 12 doi:10.3389/fpls.2021.735230
- Wichmann, M., Kamil, M., Frederiksen, A., Kotzur, S., & Scherl, M. (2021). Longterm investigations of weather influence on direct time-of-flight LiDAR at 905nm. *IEEE Sens. J.*, 22(3), 2024–2036.
- Wu, Z., Chen, Y., Zhao, B., Kang, X., & Ding, Y. (2021). Review of weed detection methods based on computer vision. *Sensors*, 21(11) doi:10.3390/s21113647
- Xu, K., Zhu, Y., Cao, W., Jiang, X., Jiang, Z., Li, S., & Ni, J. (2021). Multi-modal deep learning for weeds detection in wheat field based on RGB-D images. *Frontiers in Plant Science*, 12 doi:10.3389/fpls.2021.732968
- Yamasaki, Y., Morie, M., & Noguchi, N. (2022). Development of a high-accuracy autonomous sensing system for a field scouting robot. *Comput. Electron. Agric.*, 193, 106630.

Yu, J., Sharpe, S. M., Schumann, A. W., & Boyd, N. S. (2019). Deep learning for image-based weed detection in turfgrass. *European Journal of Agronomy*, 104, 78-84. doi:10.1016/j.eja.2019.01.004

Zhang C, Lin Y, Zhu L, Liu A, Zhang Z, Huang F. (2019). CNN-VWII: An efficient approach for large-scale video retrieval by image queries. *Pattern Recognition Letters*. 123:82–88. <https://doi.org/10.1016/j.patrec.2019.03.015>.

3. UN ENFOQUE DE APRENDIZAJE PROFUNDO PARA LA DETECCIÓN DE MAÍZ Y MALEZAS

D. Ivan Sanchez-Chavez^a, Noe Velazquez-Lopez^{a*}, Karina R. Perez-Daniel^b, C. Alberto Lopez-Figueroa^b, Miguel A. Hernandez-Perez^c, Agustin Ruiz-Garcia^a

^aUniversidad Autónoma Chapingo, Carr. Federal México-Texcoco Km 38.5, Texcoco, C.P. 56230, México. ^bUniversidad Panamericana. Facultad de Ingeniería, Augusto Rodin 498, C.P. 03920, Ciudad de México, México. ^cInstituto de Ingeniería, Universidad Veracruzana, Juan Pablo II s/n, Boca del Río, C.P. 94294, Veracruz, México

RESUMEN

El maíz es uno de los cultivos más valiosos en términos de suministro de alimentos y producción económica a nivel mundial en el sector agrícola. Sin embargo, las malezas representan uno de los mayores problemas asociados con posibles pérdidas financieras y de producción en los rendimientos de los cultivos. Recientemente se han realizado avances significativos en el desarrollo de sistemas de control de malezas basados en inteligencia artificial y métodos de detección. Algunos de estos métodos se basan en modelos de visión por computadora. Pero en la mayoría de los casos, los modelos se entrenan con un conjunto de datos de imágenes en una posición, distancia y etapa de crecimiento fijas de la maleza, lo que dificulta y hace inexacta la detección debido a problemas de generalización. Por lo tanto, en este trabajo proponemos un enfoque de aprendizaje profundo para distinguir con precisión entre el maíz y las malezas a nivel de píxel. Así este documento presenta 3 contribuciones principales. La primera se centra en el conjunto de datos. En este sentido, proporcionamos una base de datos que contiene imágenes de maíz y malezas considerando una variedad de etapas de crecimiento, diferentes ángulos de visión y una variedad de condiciones climáticas y de iluminación. La segunda se da por un proceso de etiquetado de segmentación de instancias semiautomático y la tercera se da por una arquitectura de cuasi segmentación de instancias para detectar malezas y maíz basada en U-Net y YOLOv5. Finalmente, para demostrar la eficiencia de la metodología propuesta, se realiza una comparación con los métodos existentes en la literatura, considerando no solo en términos de mAP, sino también en términos de tiempo de inferencia y tamaño del modelo, ya que estas métricas son importantes para la detección en tiempo real en dispositivos con restricciones de hardware.

Palabras clave: Detección de maíz, Detección de malezas, Cuasi segmentación de instancias, Yolov5, U-Net

A Deep Learning Approach for Corn and Weed Detection

D. Ivan Sanchez-Chavez^a, Noe Velazquez-Lopez^{a*}, Karina R. Perez-Daniel^b, C. Alberto Lopez-Figueroa^b, Miguel A. Hernandez-Perez^c, Agustin Ruiz-Garcia^a

^aUniversidad Autonoma Chapingo, Carr. Federal México-Texcoco Km 38.5, Texcoco, Z.C. 56230, Mexico

^bUniversidad Panamericana. Facultad de Ingeniería, Augusto Rodin 498, Z. C. 03920, Mexico City, Mexico.

^cInstituto de Ingeniería, Universidad Veracruzana, Juan Pablo II s/n, Boca del Río, Z. C. 94294, Veracruz, Mexico

Abstract

Corn is among the most valuable crops in terms of food supply and economic output globally in the agricultural sector. However, weeds represent one of the biggest problems associated with potential financial and production losses in crop yields. Significant advances have recently been made in developing artificial intelligence-based weed control systems and detection methods. Recently, some of these methods are based on computer vision models. However, in most cases, the models are trained with an image dataset at a fixed position, distance, and growth stage of the weed, which makes detection difficult and inaccurate due to generalization issues. Therefore, in this work, we propose a Deep-learning approach to accurately distinguish between corn and weeds at the pixel level. Hence, this paper presents 3 main contributions. The first one is focused on the dataset. In this sense, we provide a database containing images of corn and weeds considering a variety of growth stages, different viewing angles, and a variety of weather and lighting conditions. The second one is given by a semi-automated instance segmentation labeling process and the third one is given by a quasi-instance segmentation architecture for detecting weeds and corn based on U-Net and YOLOv5. Finally, to demonstrate the efficiency of the proposed methodology, a comparison with existing methods in the literature is carried out, considering not only in terms of mAP, but also in terms of inference time and model size since these metrics are important for real-time detection on devices with hardware restrictions.

Keywords: Corn detection, Weed detection, Quasi-instance segmentation, Yolov5, U-Net

Nomenclatura

Símbolos

AP Precisión promedio

FN Clasificaciones falsas negativas

FP Clasificaciones falsas positivas

GT Región real

HSV Espacio de color matiz, saturación y valor

IoU Métrica de intersección sobre unión

M Número total de clases

m Número de clase

N Número total de muestras

n Número de muestra

PR Región predicha

TN Clasificaciones verdaderamente negativas

TP Casos verdaderamente positivos bien clasificados

Abreviaturas

BB Bounding box, Cuadro delimitador

CNN Red neuronal convolucional

DL Aprendizaje profundo

FCN Red completamente convolucional

KNN K vecinos más cercanos

mAP Precisión promedio media

ML Aprendizaje automático

RCNN Red neuronal convolucional basada en regiones

RF Bosque aleatorio

RGB Espacio de color rojo, verde, azul

SVM Máquina de vectores de soporte

UAV Vehículo aéreo no tripulado (dron)

YOLO You only look once (sistema de detección de objetos)

3.1 Introducción

El maíz es un cultivo altamente productivo debido a su amplia gama de usos industriales y agrícolas, como alimentos para animales, biocombustibles y edulcorantes alimentarios. Asimismo, el maíz es uno de los productos primarios más consumidos en el mundo, incluyendo países como Estados Unidos, Colombia, Corea del Sur, Japón y México, entre otros. De hecho, Estados Unidos representó el 26% del consumo mundial de maíz durante 2020-2021 (World of Corn, 2021). En cuanto a México, el cultivo de maíz es uno de los elementos esenciales de la dieta mexicana. El consumo promedio de maíz blanco en México de 2010 a 2020 fue de 22 millones de toneladas, mientras que el consumo de maíz amarillo fue de 14 millones de toneladas (Servicio de Información Agroalimentaria y Pesquera, 2021). Por ejemplo, en el año agrícola 2020 para su siembra, se estimó una producción de maíz grano de aproximadamente 27 millones 707 mil y 775 toneladas (SADER, 2021).

Uno de los principales desafíos para lograr una mayor productividad del maíz es el control, manejo o eliminación de las malezas. Esta acción evita la competencia por agua, nutrientes y luz entre el cultivo y otras plantas silvestres, lo que beneficia el desarrollo y aumenta el rendimiento del cultivo (World of Corn, 2021). Además, las malezas pueden introducir bacterias y virus o albergar diversas plagas que finalmente dañan el cultivo. Por otro lado, las malezas en la producción de maíz resultan en pérdidas económicas sobre el rendimiento anual de los países (Soltani et al., 2016). Por lo tanto, la identificación de malezas desempeña un papel vital en la gestión y control de las malezas. El deshierbe se puede hacer manualmente, mecanizado o mediante agroquímicos.

Los herbicidas son la herramienta principal utilizada para combatir las malezas, pero presentan riesgos secundarios para el hombre y el medio ambiente, como la contaminación del suelo y del agua y la permanencia en los cultivos. Por ejemplo, el glifosato ya ha sido prohibido en México (SEGOB, 2020), Austria, Alemania, Francia, Italia y Luxemburgo, Tailandia, Bermudas y Sri Lanka, entre

otros. Alternativamente, la identificación manual de malezas también es insostenible o impráctica en muchas situaciones debido al costo y los recursos humanos. Por lo tanto, ha sido necesario considerar nuevas formas de reducir las malezas y sus efectos (Jabir et al., 2021). Una opción ha sido sistemas inteligentes que detectan automáticamente las malezas aplicando productos solo a las áreas infectadas o con métodos mecánicos localizados.

Actualmente, se están desarrollando sistemas automáticos de detección de malezas para mejorar el trabajo aplicando menores cantidades de herbicida y eliminando las malezas con mayor precisión. Por lo tanto, se necesitan estrategias de control de malezas más sostenibles, que incluyan la implementación de tecnologías de agricultura de precisión, para que se puedan tomar medidas tempranas y adecuadas (Espejo-Garcia et al., 2020). Además, se han desarrollado sistemas utilizando diferentes sensores, como cámaras, que toman imágenes en tiempo real del cultivo para detectar anomalías. Otros sistemas utilizan sensores y cámaras hiperespectrales integradas en drones que proporcionan conjuntos de datos de imágenes UAV (Vehículos Aéreos No Tripulados).

Se han desarrollado algunas técnicas de visión por computadora y aprendizaje automático (ML) (Subeesh et al., 2022) para eliminar las malezas con precisión. Sin embargo, un proceso automático para la eliminación de malezas requiere, en primer lugar, una forma automatizada de detectarlas. Uno de los enfoques de ML más populares es la Máquina de Vectores de Soporte (SVM), que se ha utilizado ampliamente para la detección de malezas (Wu et al., 2021). En este sentido, (Kamath et al., 2020) investigaron un sistema de clasificación construido usando clasificadores SVM y Random Forest (RF) para clasificar malezas en el cultivo de arroz. Las imágenes fueron tratadas, extrayendo textura, color y forma después de restar el fondo; obtuvieron una precisión del 91.36%. (Islam et al., 2021) utilizaron tres algoritmos de ML: Random Forest, SVM y k-Vecinos más cercanos (KNN) para detectar malezas en imágenes UAV de un campo de chile ubicado en Australia. Obtuvieron precisiones de detección de malezas del 96% usando

Random Forest, 94% con SVM y 63% con KNN. (Aparna & Radha, 2019) propusieron un sistema que ayuda a la detección de malezas con datos de cultivos y malezas. Las plantas se identifican como una región clasificada saliente utilizando la SVM con una precisión del 85.1%. (Shirzadifar et al., 2020) probaron la precisión de seis algoritmos de clasificación, paralelepípedo, distancia de Mahalanobis, máxima verosimilitud, mapeador de ángulo espectral, SVM y árbol de decisiones, para clasificar especies de malezas. Utilizaron una cámara multiespectral montada en un cuadricóptero Phantom3 UAV. El método SVM resultó en la mejor precisión del 98-99% bajo clasificación basada en objetos. (Louargant et al., 2018) utilizaron imágenes tomadas a 3m sobre el suelo con una cámara multiespectral en campos de maíz y remolacha azucarera. Los píxeles fueron clasificados espacialmente primero para especificar que eran malezas o cultivos. Las imágenes fueron reclasificadas en malezas y cultivos utilizando un clasificador SVM según las características espectrales. La precisión promedio de detección de malezas fue del 89% para el método que combinaba aspectos espaciales y espectrales, 79% para el método solo espacial y 75% para el método solo espectral. (Rani et al., 2022) utilizaron algoritmos de regresión logística y SVM para clasificar malezas y cultivos. El método utiliza un histograma de gradientes para extraer características; este modelo se integró en un robot de campo. Con esto, el sistema aplica herbicidas solo donde es necesario, obteniendo una precisión del 83%. (Louargant et al., 2018) utilizaron un UAV para obtener imágenes de 11 campos comerciales de cebolla, que se utilizaron para mapear la ubicación de las malezas en los campos; utilizaron dos algoritmos de clasificación, máxima verosimilitud y SVM, y todos los porcentajes de precisión estuvieron por encima del 85%. Sin embargo, todos estos algoritmos están diseñados para clasificar imágenes; asumen que hay una sola instancia por imagen y no pueden identificar la ubicación de un objeto en particular.

Por otro lado, existen otros enfoques para procesar y clasificar datos de imágenes basados en técnicas de ML. Uno de los enfoques más precisos para procesar imágenes utilizando técnicas de aprendizaje automático son los enfoques basados en redes de aprendizaje profundo (DL), específicamente, aquellos que

se basan en las Redes Neuronales Convolucionales (CNN) para la extracción de características, como el algoritmo You Only Look Once (YOLO), que es un algoritmo de detector de objetos ampliamente utilizado. En 2021, YOLOv3 se utilizó para entrenar y probar un modelo para la identificación de malezas (Etienne et al., 2021). Utilizaron 374 imágenes RGB con un sistema aéreo de campos de maíz y soja a 30 m y 10 m sobre el nivel del suelo. Los mejores resultados se observaron con imágenes obtenidas a 10 m, donde hubo un 91.48% de precisión en la detección de monocotiledóneas y un 86.13% para malezas del tipo dicotiledóneas. (Xie et al., 2019) utilizaron un YOLOv3 para identificar malezas en un implemento móvil diseñado para eliminarlas; incluye una unidad de procesamiento gráfico. El clasificador se entrenó durante 20,000 iteraciones. Durante el proceso, la pérdida promedio (avg loss) alcanza 0.135, se detiene el entrenamiento y el mAP (precisión promedio media) es del 75.4%. (Partel et al., 2019) crearon un aplicador de herbicidas inteligente para distinguir entre malezas y cultivos. Utilizaron las redes YOLOv3 y YOLOv4 para detectar malezas; las imágenes se recopilaban a 10, 20 y 30 m sobre el nivel del suelo. La red YOLOv3 alcanzó una precisión del 87.2% y YOLOv4 del 90.1%. (Huang et al., 2021) utilizaron una red YOLOv4 entrenada para detectar malezas y aplicar herbicidas de manera selectiva; esto se logró utilizando cámaras RGB en una UAV y una unidad de procesamiento de GPU en tiempo real. La precisión del sistema fue del 96%.

En este trabajo, proponemos un enfoque de aprendizaje profundo para distinguir con precisión entre el maíz y las malezas, ya que este es uno de los cultivos más importantes. Esto se lleva a cabo en diferentes etapas de crecimiento de los cultivos, por lo tanto, presentamos una metodología para un proceso de etiquetado de segmentación de instancias semiautomático y una arquitectura de segmentación cuasi-instantánea. El documento está organizado como sigue. La Sección 2 describe el proceso utilizado para generar el conjunto de datos. La Sección 3 presenta la arquitectura propuesta para la segmentación cuasi-instantánea. En la Sección 4, se muestran los resultados obtenidos con la

arquitectura propuesta y se comparan con otros métodos. Finalmente, en la Sección 5, se presentan las conclusiones y se sugieren trabajos futuros.

3.2 Materiales y métodos

3.2.1 Métodos

Los paradigmas de las tareas de reconocimiento visual

Los algoritmos de detección de objetos y segmentación semántica son uno de los enfoques principales en las tareas de reconocimiento. Los algoritmos de detección de objetos tienen como objetivo predecir la clase y la ubicación dentro de una imagen de una clase de objeto conocida utilizando un cuadro delimitador (BB). Por el contrario, los algoritmos de segmentación semántica determinan la categoría específica de cada píxel dentro de una imagen. Sin embargo, no pueden distinguir los límites de diferentes instancias de la misma clase dentro de una imagen. Para resolver este problema, se introdujeron los algoritmos de segmentación de instancias. Los algoritmos de segmentación de instancias pueden distinguir entre diferentes instancias de una clase (Xiongwei et al., 2020). En general, cada tarea de reconocimiento visual está relacionada con un tipo particular de formato de etiquetado. La Fig. 1 muestra un ejemplo de las tareas de reconocimiento visual mencionadas anteriormente.

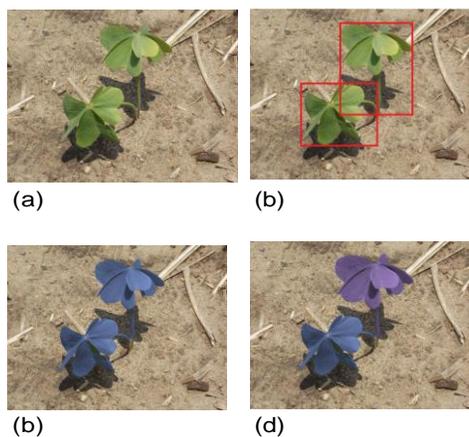


Figura 1. Ejemplos de tareas de reconocimiento visual. (a) Imagen RGB original. (b) Detección de objetos. (c) Segmentación semántica. (d) Segmentación de instancias.

Por lo general, los algoritmos de segmentación de instancias superan la precisión, exactitud y las métricas de intersección sobre unión de otras tareas de reconocimiento visual, ya que resuelve los problemas de detección de objetos y segmentación semántica. Sin embargo, la segmentación de instancias es costosa en términos de tiempo de procesamiento, inferencia y recursos computacionales (Huang et al., 2020), especialmente para aplicaciones integradas en dispositivos con restricciones de hardware. Actualmente, se han propuesto algunos enfoques de segmentación de instancias para la detección de malezas (Champ et al., 2022; Quan et al., 2021). Estos enfoques se basan en las arquitecturas Mask R-CNN y BlendMask, respectivamente, que son métodos de segmentación de instancias muy populares. Sin embargo, los métodos de segmentación de instancias puras requieren más complejidad computacional, afectando la velocidad de detección del modelo, especialmente durante la inferencia en dispositivos con restricciones de hardware. Por esta razón, este trabajo propone un método de segmentación cuasi-instantánea basado en las arquitecturas U-Net y Yolov5, donde la primera red se utiliza para la segmentación semántica y la segunda red se utiliza para la detección de objetos, las dos tareas de un algoritmo de segmentación de instancias.

YOLO

El algoritmo YOLO (You Only Look Once) fue introducido en (Redmon et al., 2016). YOLO fue considerado el primer detector de objetos en tiempo real. YOLO divide la imagen en celdas no superpuestas de 7×7 píxeles. El algoritmo tiene como objetivo detectar el centro de hasta dos objetos. YOLO predice la presencia de un objeto en una celda y las coordenadas del cuadro delimitador del objeto de una clase particular, en caso de que esa celda tenga un objeto. Sin embargo, YOLO evolucionó, y su versión YOLOv5 (DoCong, 2021) fue presentada en 2020. A diferencia de su predecesor, YOLOv5 propone utilizar una columna vertebral, un cuello y una cabeza en su arquitectura. La columna vertebral se utiliza como extractor de características en diferentes granularidades, el cuello se utiliza para la agregación de características, y la cabeza está a cargo de los pasos de predicción y regresión en base a las características del cuello (DoCong, 2021).

RetinaNet

RetinaNet (Lin et al., 2020) también es un detector de objetos. Retinanet comprende 3 etapas principales, la columna vertebral, la subred de clasificación y la subred de regresión. La columna vertebral se utiliza para la extracción de características utilizando una Red Convencional Totalmente (FCN), que genera mapas de características a diferentes escalas de la imagen para realzar las características más relevantes en cada escala. La extracción de características a diferentes escalas permite la detección de objetos de diferentes tamaños. Estas características pasan a la subred de clasificación, que determina la existencia de un objeto y la etiqueta de clase en un mapa de características dado, si lo hay. La subred de clasificación se procesa al mismo tiempo que la subred de regresión. La subred de regresión se encarga de definir el solapamiento relativo entre el cuadro predicho y su cuadro de verdad de base (ground-truth) (Lin et al., 2020).

Red U-Net

U-Net (Ronneberger et al., 2015) es un algoritmo de segmentación semántica. U-Net requiere los datos de imagen originales y sus etiquetas representadas por una máscara binaria por imagen. La arquitectura de U-Net se compone de dos etapas principales. La primera etapa consiste en una parte del codificador, que se utiliza para la extracción de características basada en bloques consecutivos de Red Neuronal Convolutiva (CNN). La parte del codificador genera un gran número de mapas de características con baja dimensionalidad. Luego, estos mapas de características se decodifican mediante operaciones convolucionales transpuestas, mientras que los mapas de características en cada nivel se enriquecen con las capas de codificación en el mismo nivel. La etiqueta predicha final (máscara binaria) se obtiene aplicando la función de activación sigmoide a los mapas de características decodificados finales (Ronneberger et al., 2015).

3.2.2. Materiales y adquisición de datos.

La recolección de datos se llevó a cabo en un campo experimental de la Universidad Autónoma Chapingo en México, en el área ubicada en las coordenadas 19° 29'38.4"N y 98° 52' 45.2"O. El campo tenía maíz híbrido

plantado con fecha de siembra el 4 de julio de 2019. Las fotografías se tomaron durante los primeros 40 días del ciclo de vida del cultivo. La cámara utilizada para esta experimentación es una cámara Nikon COOLPIX L320 de 16.1 megapíxeles con lente de vidrio NIKKOR ED con zoom súper telefoto 26x. La Figura 2 ilustra el campo experimental utilizado para este trabajo.

La zona del cultivo tiene una disposición rectangular de 32 puntos de muestreo para cubrir todo el campo, como se muestra en la Figura 3. El experimento consiste en tomar imágenes de maíz y malezas en un lapso de tiempo de dos horas de 11:00 a 13:00 dos veces a la semana (martes y viernes) en cada uno de los 32 puntos.

El experimento considera imágenes tomadas en dos ubicaciones y distancias diferentes. Por lo tanto, la distancia entre la cámara y el maíz o la maleza se ilustra en la Fig. 4 y se describe de la siguiente manera:

- a) Horizontalmente a 1 m de altura.
- b) Verticalmente a nivel del suelo y a 0.3 m de distancia.



Figura 2. Campo experimental: Universidad Autónoma Chapingo.

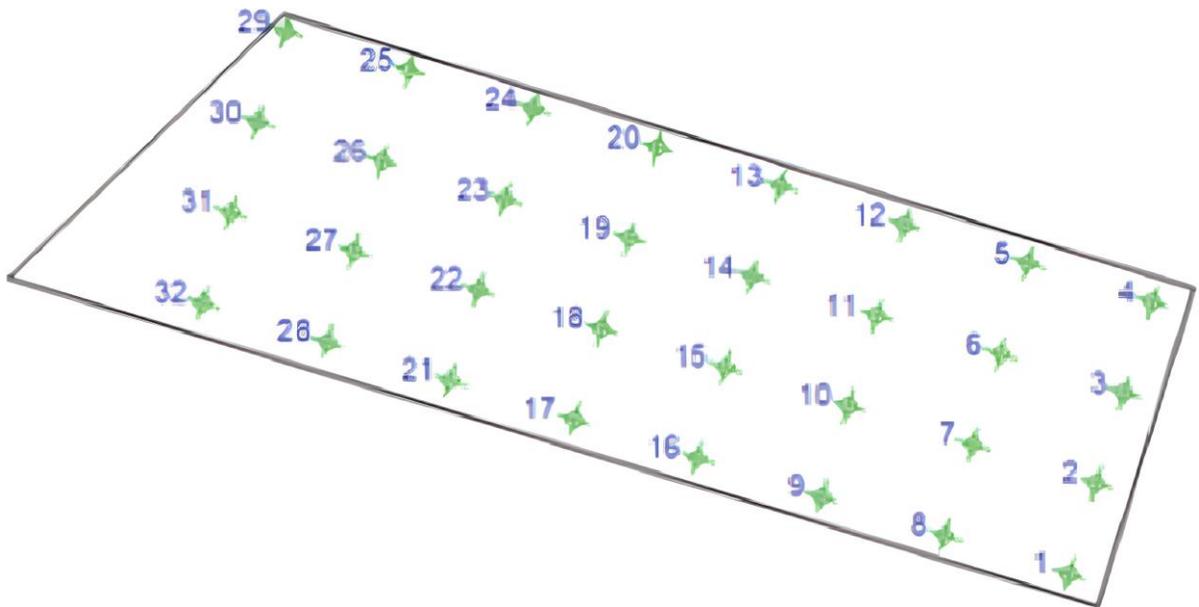


Figura 3. Diseño experimental: Cultivo y sus 32 puntos de muestreo.

Según las ubicaciones y distancias mostradas en la Fig. 4, se tomaron imágenes de maíz y malezas en 5 puntos de vista diferentes en la dirección del surco, como se ilustra en la Fig. 5 y se describe a continuación:

a) Horizontalmente (lente paralela al suelo) a 1 m de altura.

- 45° a 1 m de altura y a 0,3 m de distancia.

- 45° a 0,5 m de altura y a 0,3 m de distancia.

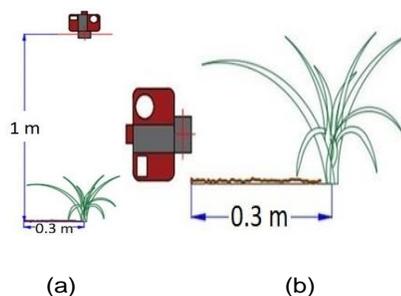


Figura 4. Ubicación de la cámara para la adquisición de imágenes durante los experimentos.

- 45° a 0,3 m de altura y a 0,3 m de distancia.

b) Verticalmente (lente perpendicular al suelo) a nivel del suelo y a 0,3 m de distancia.

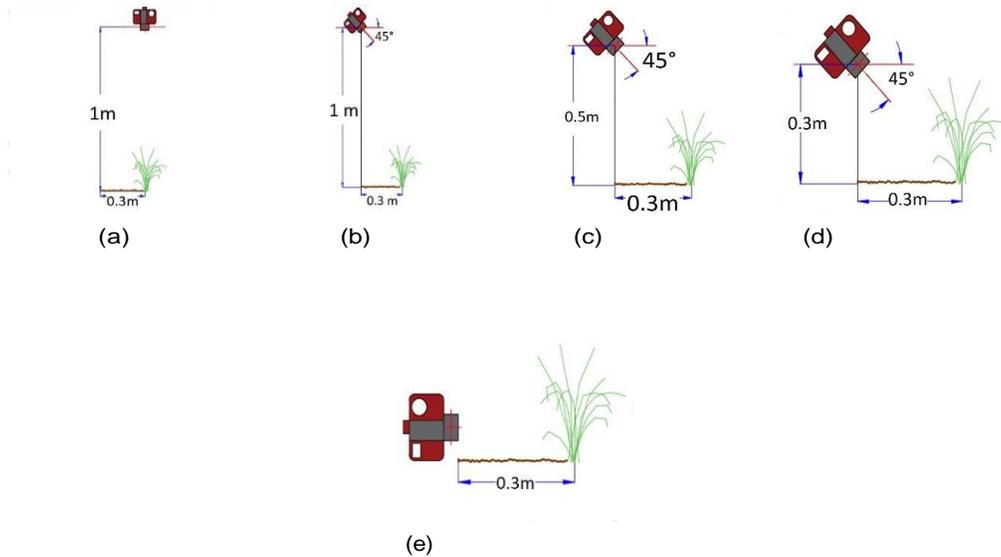


Figura 5. La ubicación de la cámara para la adquisición del conjunto de datos con diferentes condiciones de luz y etapas de crecimiento del maíz.

Se recopiló un banco de imágenes RGB de 1.028 imágenes de maíz y 1.103 imágenes de malezas de varias especies, para un total de 2.131 imágenes, con una resolución de 1920 × 1080 píxeles, y se redimensionaron a una resolución fija de 640 × 480 píxeles.

Todas las imágenes fueron etiquetadas utilizando el software labellmg en el formato YOLO (You Only Look Once). Algunas de las imágenes contienen una sola instancia de maleza o maíz. Sin embargo, otras tienen varias instancias. Por lo tanto, recopilamos un total de 3821 instancias de maíz y malezas. Esto es 2000 instancias de malezas y 1821 instancias de maíz. El conjunto de datos propuesto consta de diversas imágenes de maíz y malezas con diferentes iluminaciones, clima, distancia y condiciones de punto de vista de los datos, considerando diferentes etapas de crecimiento según los primeros 40 días del ciclo de vida del cultivo. Cabe señalar que esta diversidad en los datos es una característica deseable que contribuye a la generalización del aprendizaje y evita el sobreajuste en el modelo.

En general, la detección de malezas y maíz también puede verse como un problema de detección de objetos o un problema de segmentación de instancias. Por lo tanto, este trabajo propone un método de ensamblaje para la detección de objetos y la segmentación semántica para lograr un algoritmo de detección de maíz y malezas preciso.

3.2.3 Métricas para evaluar modelos de Detección de Objetos y Segmentación de Instancias

El rendimiento de la arquitectura propuesta para la detección de maíz y malezas se evalúa mediante las siguientes métricas:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$mAP = \frac{\sum_{n=1}^N AP(n)}{N} \quad (4)$$

$$AP = \sum_{m=1}^M Precision(m)\Delta Recall(m) \quad (5)$$

$$IoU = \frac{|GT \cap PR|}{|GT \cup PR|} \quad (6)$$

Donde TP representa las muestras bien clasificadas para los casos positivos verdaderos y TN para los negativos verdaderos. Por otro lado, las muestras clasificadas incorrectamente están dadas por FP en el caso de clasificaciones positivas falsas y FN en casos negativos falsos. En un escenario de detección de objetos, la precisión se refiere a la probabilidad de que los cuadros delimitadores predichos correspondan a los cuadros delimitadores de entrada. La recuperación da la probabilidad de que los cuadros delimitadores de entrada sean detectados correctamente. El mAP representa la media de la curva de Precisión y Recuperación para todas las clases detectadas. Por lo tanto, IoU es la métrica de Intersección sobre Unión, cuantificando la región de superposición entre una

región predicha (PR) y una región real (GT), considerando el numerador como el área de superposición y el denominador como el área de unión.

3.3 Entrenamiento del modelo

La estrategia propuesta para la detección de maíz y malezas se muestra en la Fig. 6. Este proceso comprende varias etapas descritas a continuación.

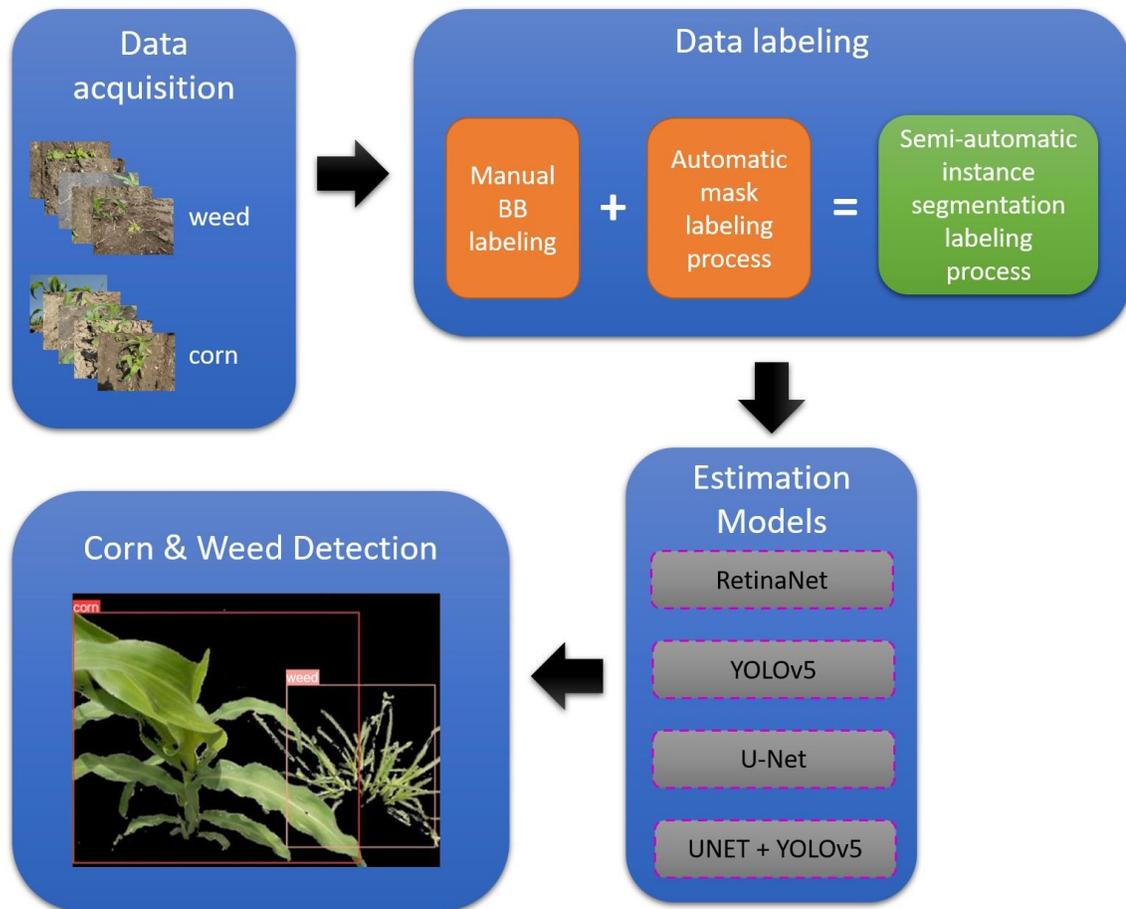


Figura 6. Visión general del proyecto para la detección de maíz y malezas.

3.3.1. Adquisición de datos

La configuración de adquisición del conjunto de datos se describe detalladamente en la Subsección 2.2. El conjunto de datos considera instancias bajo diferentes condiciones climáticas y etapas de crecimiento durante los primeros 40 días del ciclo de vida del cultivo. Se proporciona una descripción detallada del conjunto de datos en el cuadro 3. Además, el cuadro 3 resume las características de algunos conjuntos de datos recientes utilizados para la detección de malezas. De

este cuadro, podemos notar que la mayoría considera condiciones de iluminación y clima constantes, distancia de adquisición limitada y solo una etapa de crecimiento de malezas. Estas características pueden limitar la generalización del proceso de aprendizaje de detección de malezas. (Etienne et al., 2021) proporciona un conjunto de datos muy rico; sin embargo, su conjunto de datos se centra en detectar monocotiledóneas y dicotiledóneas a partir de imágenes aéreas, lo que puede llevar a imprecisiones en detecciones que implican capturas de imágenes terrestres.

3.3.2 Proceso de etiquetado de la base de datos

Actualmente, existen dos paradigmas principales para el etiquetado de datos de imágenes: el proceso de etiquetado de cuadros delimitadores (BB), comúnmente utilizado para algoritmos de detección de objetos, y el proceso de etiquetado de máscaras, utilizado con frecuencia para tareas de segmentación semántica.

Cuadro 3. Descripción de diferentes conjuntos de datos para la detección de malezas.

Fuente	Alcance de detección	# de imágenes	# de instancias	de Tipo datos	de Resolución	Ángulo de adquisición	Consideraciones adicionales de adquisición
(Etienne et al., 2021)	Maleza (monocotiledóneas y dicotiledóneas)	374	25,560	RGB UAV	416 416	xAéreo	Variaciones en luz, clima y etapa de crecimiento
(Xie et al., 2019)	Maleza	559	No mencionado	RGB	416 416	xSuperior	Luz, clima, distancia y etapa de crecimiento constantes
(Partel et al., 2019)	Malezas, girasoles, plantas de pimienta y	1,000	No mencionado	RGB considera instancias	1024 480	xSuperior	Luz, clima, distancia y etapa de

	junco (o cyperus) invasor			reales y sintéticas			crecimiento constantes
(Urmashv et al., 2021)	Malezas (ambrosia, amaranto, quinoa, bromo y enredadera)	5,00	No menciónado	RGB	3024 x4032	Superior	Variaciones en luz
(Chang et al., 2021)	Malezas	200	No menciónado	RGB	416 x416	Superior	Variaciones en luz y clima
(Quan et al., 2021)	Malezas (Solanum nigrum, pasto de cebada o Echinochloa crus-galli, Abutilon theophrasti Medicus)	5,85	No menciónado	RGB	1024 x1024	Superior	Variaciones en luz, y frontal clima y punto de vista
Propuesto	Maíz y malezas	2,1	3,821	RGB	640 x480	Horizontal-	Variaciones en luz, clima, etapas de a 1 m de crecimiento, diferentes puntos de vista y verticalmente a una distancia de 30 cm

Para el alcance del enfoque propuesto, el proceso de etiquetado manual se realizó utilizando el software labellmg, donde un experto en el campo de la agricultura de precisión etiquetó todas las muestras. Los BB recopilados se almacenaron utilizando el formato de anotación YOLO. Por otro lado, el proceso de etiquetado de máscaras se realizó utilizando un proceso automático basado

en la segmentación de color en el espacio de color HSV. Este proceso se describe a continuación.

Procesamiento automático de etiquetado de máscaras

Usualmente, el proceso de etiquetado de enmascaramiento implica un proceso de etiquetado manual, así como las técnicas para obtener cuadros delimitadores. Sin embargo, obtener una máscara de una imagen es necesario para dibujar un polígono que se ajuste a la forma del objeto de etiquetado o dibujar manualmente el contorno del objeto. Este proceso es mucho más lento que la tarea de etiquetado basada en BB. Por esta razón, este trabajo propone utilizar un proceso automático de etiquetado de máscaras basado en características de color.

Las hojas de malezas y maíz comparten similitudes en color. Por lo tanto, proponemos utilizar esta información para crear un filtro de color para eliminar el fondo de las imágenes y generar las máscaras correspondientes para el maíz y las malezas en lugar de crear manualmente la etiqueta basada en máscaras para cada instancia. Por lo tanto, el proceso automático de etiquetado de máscaras propuesto sigue el diagrama de flujo presentado en la Fig. 7.

Cada imagen se convierte del espacio de color RGB al espacio de color HSV para generar las máscaras correspondientes. Se utiliza el espacio de color HSV ya que tiende a ser más adecuado para la segmentación semántica en imágenes con fondos uniformes (Khamar et al., 2015), como en el caso de la segmentación de maíz y malezas. Luego, para identificar los píxeles de color verde, determinamos los rangos de color para cada canal HSV en base a heurísticas. El rango de valores de píxeles para el filtro de color en el espacio de color HSV se estableció de la siguiente manera: $25 \leq H \leq 102$ para el canal H, $52 \leq S \leq 255$ para el canal S y $72 \leq V \leq 255$ para el canal V.

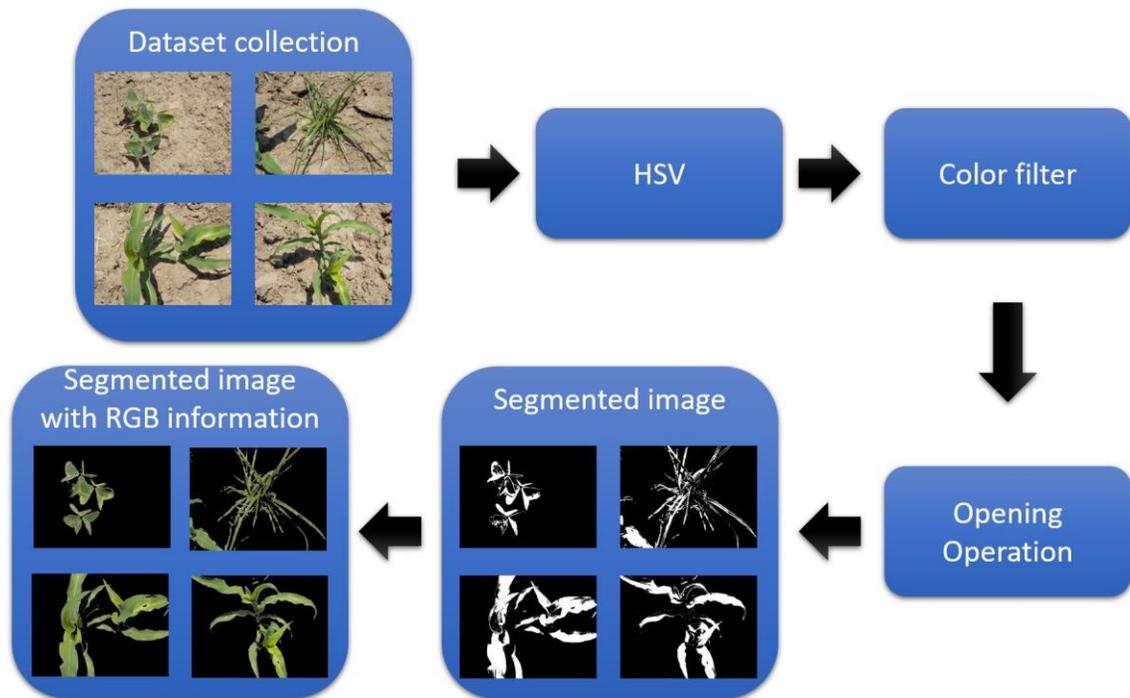


Figura 7. Resumen del proceso propuesto para el etiquetado de máscaras.

Estos umbrales se utilizan para generar las máscaras correspondientes para malezas y maíz, donde los píxeles en este rango se establecen en 255 y los píxeles fuera del rango se establecen en 0. Luego, obtenemos una máscara binaria que sufre la operación morfológica de apertura de este proceso para refinar la máscara segmentada y eliminar las líneas pequeñas. La Fig. 8(a) y la Fig. 9(a) muestran algunas muestras del conjunto de datos de maíz y malezas, respectivamente. El resultado del proceso automático de etiquetado de máscaras se muestra en la Fig. 8(b) para el caso del conjunto de datos de maíz y en la Fig. 9(b) para el caso del conjunto de datos de malezas.

Por otro lado, la Fig. 8(c) y la Fig. 9(c) muestran los resultados de aplicar la operación AND a nivel de bits entre las imágenes RGB originales y sus máscaras correspondientes. El proceso propuesto de etiquetado automático de máscaras puede verse como un algoritmo para la segmentación semántica. Sin embargo, ningún algoritmo de segmentación semántica puede distinguir entre múltiples objetos de la misma categoría. Por lo tanto, no puede determinar los límites de un objeto y otro. Por esta razón, en este trabajo, se propone un proceso de

etiquetado de segmentación de instancias semiautomático para superar esta limitación.

Proceso de segmentación de instancias semiautomático

El proceso de segmentación de instancias semiautomático propuesto considera tanto el conjunto de imágenes segmentadas con información RGB como los cuadros delimitadores etiquetados manualmente. Este proceso se ejemplifica en la Fig. 10.

Las Fig. 8(d) y 9(d) muestran ejemplos de algunas imágenes de maíz y malezas segmentadas utilizando el proceso de segmentación de instancias semiautomático propuesto, respectivamente.

El proceso de segmentación de instancias semiautomático propuesto permite una obtención más rápida de etiquetas que el proceso de etiquetado manual tradicional. Además, mantiene la información de color y textura del objeto de interés, así como las etiquetas BB, mientras permite el aprendizaje a partir del objeto específico de interés, como las etiquetas basadas en la segmentación.



(a)



(b)



(c)



(d)

Figura 8. Transformación de algunas muestras de maíz durante el proceso propuesto de etiquetado automático de máscaras. (a) Imágenes RGB originales. (b) Máscaras de la estrategia propuesta. (c) Resultado de la operación AND a nivel de bits entre (a) y (b). (d) Muestra algunas muestras de las etiquetas de segmentación de instancias semiautomáticas propuestas.

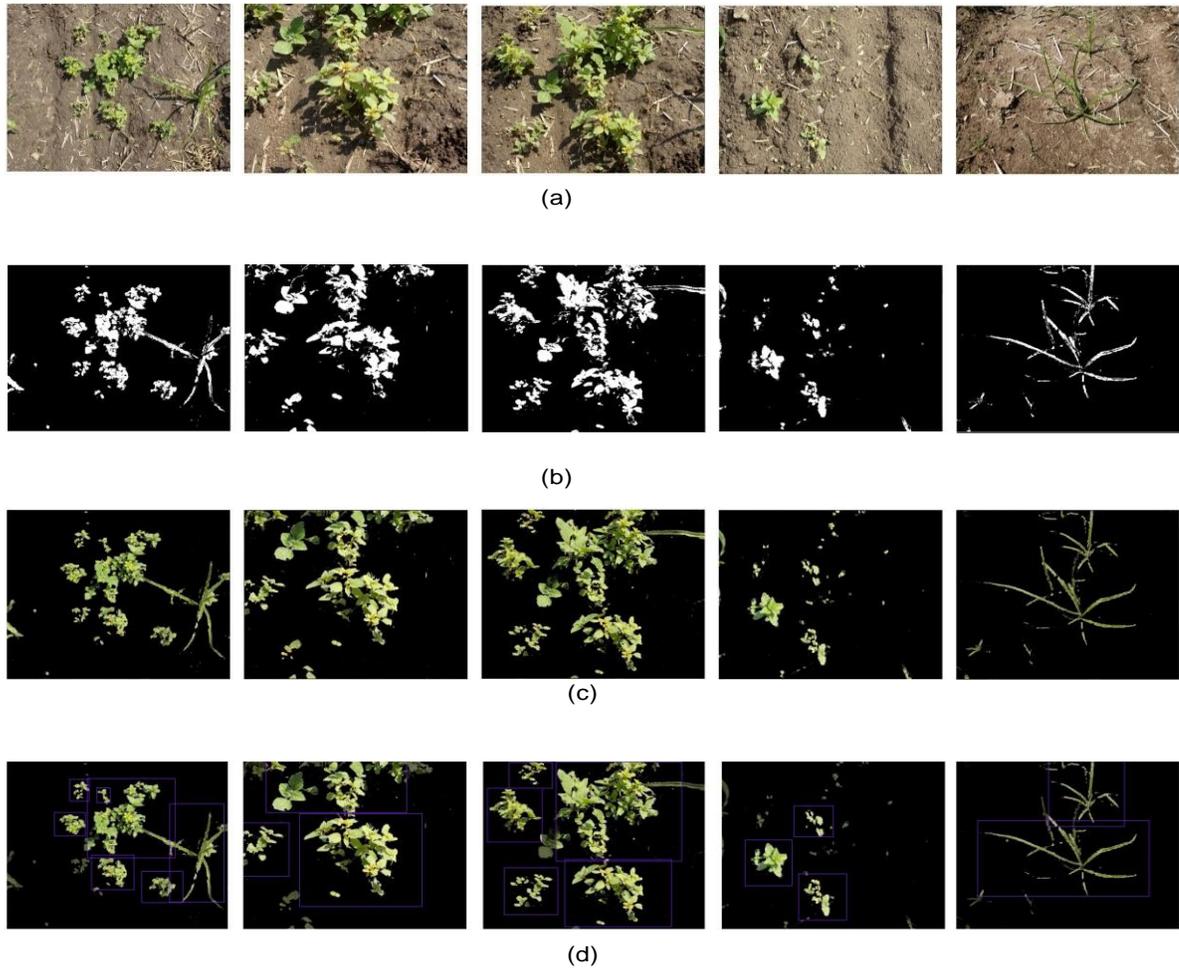


Figura 9. Transformación de algunas muestras de malezas durante el proceso propuesto de etiquetado automático de máscaras. (a) Imágenes RGB originales. (b) Máscaras de la estrategia propuesta. (c) Resultado de la operación AND a nivel de bits entre (a) y (b), es decir, la imagen segmentada con información RGB. (d) Muestra algunas muestras de las etiquetas de segmentación de instancias semiautomáticas propuestas.

3.3.3. Modelo de estimación

Preparación de datos

Como se menciona en el cuadro 3 y se explica en la Sección 3, recopilamos un conjunto de datos de 2131 imágenes, considerando 1028 imágenes de maíz y 1103 imágenes de varios tipos de malezas durante los primeros 40 días del ciclo de vida del cultivo. De este conjunto de datos, recopilamos 3821 instancias, 2000 instancias de malezas y 1821 de maíz a partir de estas imágenes. Sin embargo, para crear más muestras a partir de estos datos, utilizamos el aumento de imagen

realizando rotación (45 y 180 grados de rotación), volteo horizontal, modificación de gamma y adición de ruido gaussiano. Con este aumento de datos, obtuvimos un total de 3507 imágenes, 1650 imágenes de maíz, 1650 de malezas y 207 imágenes aleatorias.

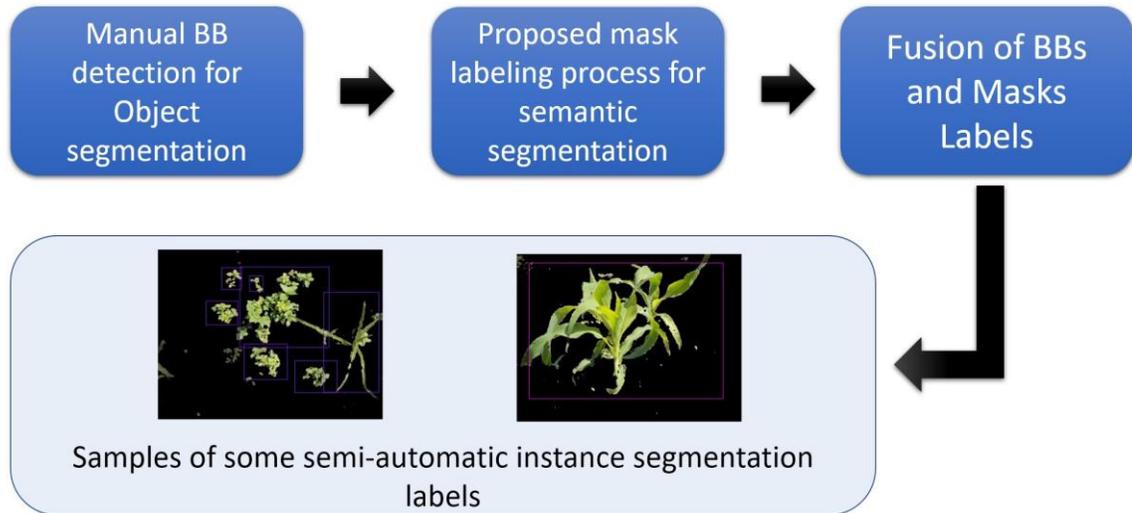


Figura 10. Resumen del proceso de etiquetado de segmentación de instancias semiautomático propuesto.

Las imágenes del conjunto de datos se utilizan para entrenar diferentes arquitecturas, incluyendo RetinaNet (Lin et al., 2020), YOLOv5 (DoCong, 2021) y U-Net (Ronneberger et al., 2015). El entrenamiento de cada modelo de estimación requiere una versión particular del conjunto de datos. Por lo tanto, el cuadro 4 resume las particularidades de cada versión del conjunto de datos de acuerdo con la red correspondiente utilizada en este trabajo.

Cuadro 4. Configuración del conjunto de datos para entrenar diferentes Modelos de Estimación.

Modelo de estimación	Tamaño del dataset aumentado	Tamaño de entrenamiento	Resolución	Tipo etiqueta de
RetinaNet	3507	2805	640x480	BB
YOLOv5	3507	2805	640x480	BB
U-Net	3507	2805	512x512	Proceso automático de

				etiquetado de mascara binaria propuesto
U- Net+YOLOv5	3507	2805	512x512 convertido a 640x480	Proceso semiautomático de segmentación de instancias propuesto

Como se menciona en el cuadro 4, cada experimento consideró su configuración particular de acuerdo con el modelo de estimación y la etiqueta. Entonces, las arquitecturas Retinanet y YOLOv5 requieren etiquetas que contengan información sobre los BBs y las imágenes RGB del conjunto de datos. El U-Net requiere imágenes RGB y etiquetas binarias obtenidas del proceso propuesto de Etiquetado Automático de Máscaras (consulte la Sección 3.2.1). Sin embargo, la arquitectura propuesta procesa el conjunto de datos como se describe en el proceso de Segmentación de Instancias Semiautomático (consulte 3.2.2) e ilustrado en la Fig. 10, y se redimensiona dos veces, una para cada etapa del proceso de etiquetado semiautomático.

Arquitectura propuesta

La arquitectura propuesta se muestra en la Fig. 11. Esta arquitectura se basa en un proceso de dos etapas. Utilizando la arquitectura U-Net, la primera etapa entrena un modelo de segmentación semántica para segmentar la región de las hojas (maíz y malezas) del fondo. La segunda etapa recibe la región segmentada como entrada y realiza la detección de objetos utilizando el detector YOLOv5.

Como se muestra en la Fig. 11, el modelo de segmentación semántica utiliza el proceso de generación de máscaras automáticas propuesto para generar las máscaras RGB de maíz y malezas utilizando la arquitectura U-Net y la operación a nivel de bits entre la imagen RGB y la máscara binaria generada por el U-Net.

En contraste, el modelo detector de objetos utiliza el proceso de etiquetado semiautomático propuesto como entrada del algoritmo YOLOv5.

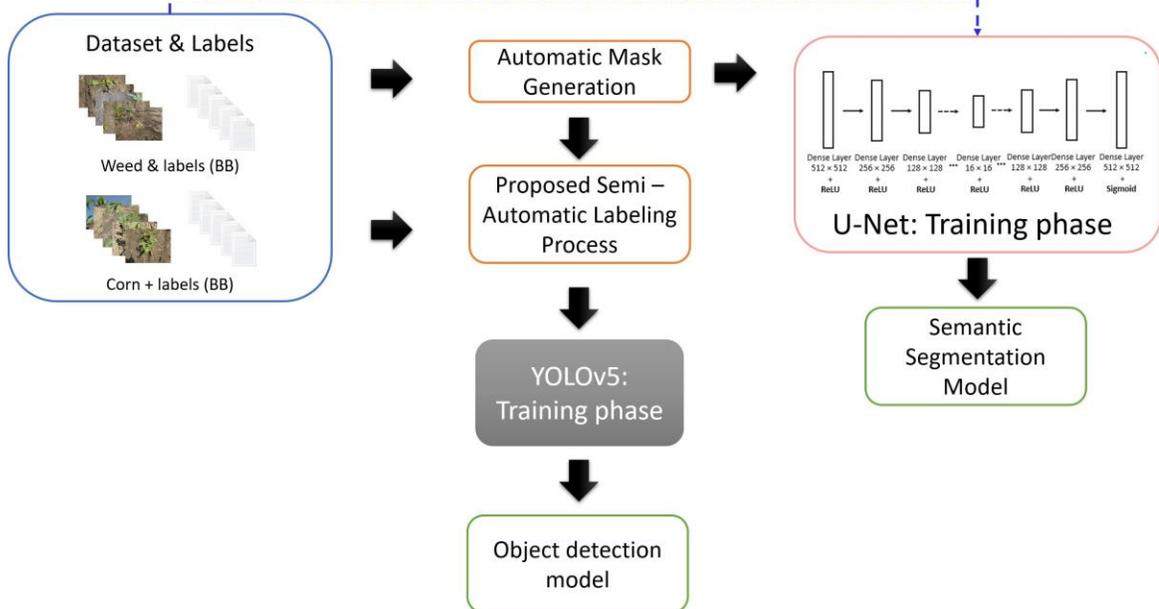


Figura 11. Arquitectura propuesta: Fase de entrenamiento.

Las imágenes y máscaras se redimensionan a una resolución de 512×512 para el modelo U-Net y se redimensionan a una resolución de 640×480 para la etapa de detección de objetos.

Por otro lado, el proceso de inferencia se ilustra en la Fig. 12. Esta figura muestra los resultados de la estrategia propuesta, donde el modelo de segmentación semántica aísla el primer plano del fondo, y el modelo de detección de objetos clasifica cada instancia en una categoría particular, maíz o maleza. Por lo tanto, esta arquitectura se puede ver como un modelo de cuasi segmentación de instancias.

La siguiente sección describe los parámetros de entrenamiento de cada arquitectura y presenta los resultados correspondientes para cada caso.

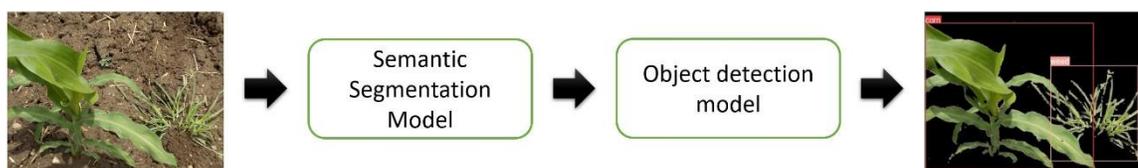


Figura 12. Arquitectura propuesta: Fase de inferencia.

3.4 Resultados experimentales y discusión

Esta sección describe los parámetros de entrenamiento para cada arquitectura y presenta los resultados experimentales de cada modelo evaluado.

Se llevaron a cabo cuatro experimentos, uno para cada arquitectura evaluada. El número de imágenes utilizadas para el entrenamiento y las pruebas se describen en el cuadro 4, en una proporción del 80% del conjunto de datos para entrenamiento y el 20% para pruebas. Todos los experimentos se realizaron utilizando un procesador Intel™ i7-13700K, 128 GB de RAM y una GPU NVIDIA™ GeForce RTX 3090. Todos los modelos se implementan usando Python 3.10 y el marco de trabajo Tensorflow 2.

3.4.1 RetinaNet

En este experimento, RetinaNet (Lin et al., 2020) utiliza un modelo ResNet-50 preentrenado como estructura base. La estructura base ResNet-50 fue preentrenada en el conjunto de datos MS-COCO (Lin et al., 2014). Esta arquitectura de RetinaNet utiliza un tamaño de lote de 8, con 613 pasos por época durante 100 épocas, y Adam como optimizador estocástico. Esta arquitectura se calcula para fines de comparación.

Utilizamos las imágenes RGB y sus correspondientes BB como entrada para entrenar la RetinaNet para lograr un modelo con mAP = 0.843 durante la inferencia. El tamaño del modelo entrenado es de 139 MB y realiza un tiempo promedio de inferencia de 0.061 segundos.

La Fig. 13 muestra algunos ejemplos visuales de predicciones realizadas utilizando el modelo entrenado con la arquitectura RetinaNet. A partir de estos ejemplos, podemos notar un problema de este modelo, en algunos casos, el modelo sobredetecta muestras (ver Fig. 13(b)), mientras que en otros, hay una falta de detección (ver Fig. 13(d)).

3.4.2 Red U-Net

La arquitectura U-Net (Ronneberger et al., 2015) entrenada para la segmentación semántica de hojas (maíz y maleza) considera imágenes RGB como entrada, así como las máscaras binarias generadas por el proceso de generación de máscara automática (ver Sección 3.2.1) como etiquetas. Todas las imágenes y etiquetas se escalan a una resolución de 512 × 512. La U-Net utilizada para el entrenamiento usa la entropía cruzada binaria para clasificar cada píxel, asignando un valor de 0 para los píxeles de fondo y 1 para un píxel correspondiente a una región de hoja. El resto de los parámetros se utilizan como se propone en (Ronneberger et al., 2015). Utilizamos un tamaño de lote de 8, con una tasa de aprendizaje de $1e - 4$, y entrenamiento durante 50 épocas. El 80% de los datos se utiliza para entrenamiento y el 20% para pruebas.

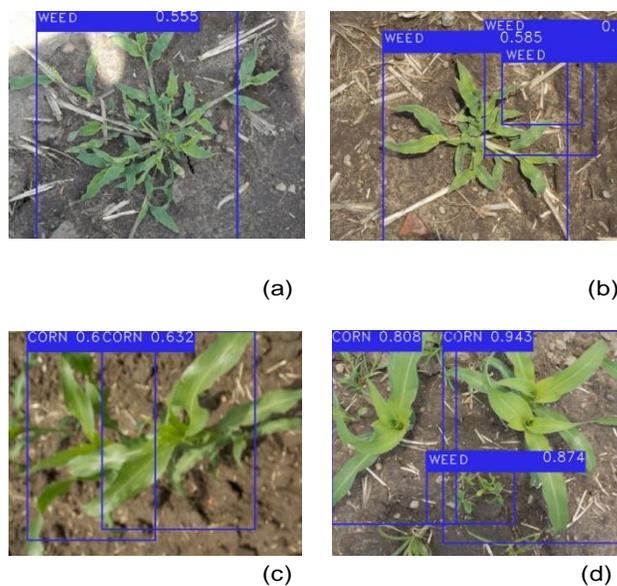


Figura 13. Detección de maíz y malezas utilizando el modelo entrenado con la arquitectura RetinaNet. La primera fila muestra algunos ejemplos de predicciones para malezas, y la segunda fila presenta algunos resultados para predicciones de maíz.

El modelo obtenido a partir de la arquitectura logra las métricas: $accuracy=0.9265$, $recall=0.8792$, $precision=0.8902$ y $IoU=0.7424$. El modelo requiere 14.4 MB para ser almacenado y su tiempo promedio de inferencia es de 0.016 segundos.

La Fig. 14 muestra algunos resultados visuales de las máscaras binarias predichas. En esta imagen, podemos notar que la U-Net es capaz de predecir correctamente las máscaras para el maíz y las malezas.

Sin embargo, el modelo entrenado no puede distinguir entre las máscaras predichas para el maíz de las máscaras para las malezas. Por lo tanto, es necesario implementar un paso adicional para identificar las instancias de maíz y malezas en las máscaras predichas mediante un algoritmo de detección de objetos como YOLO.

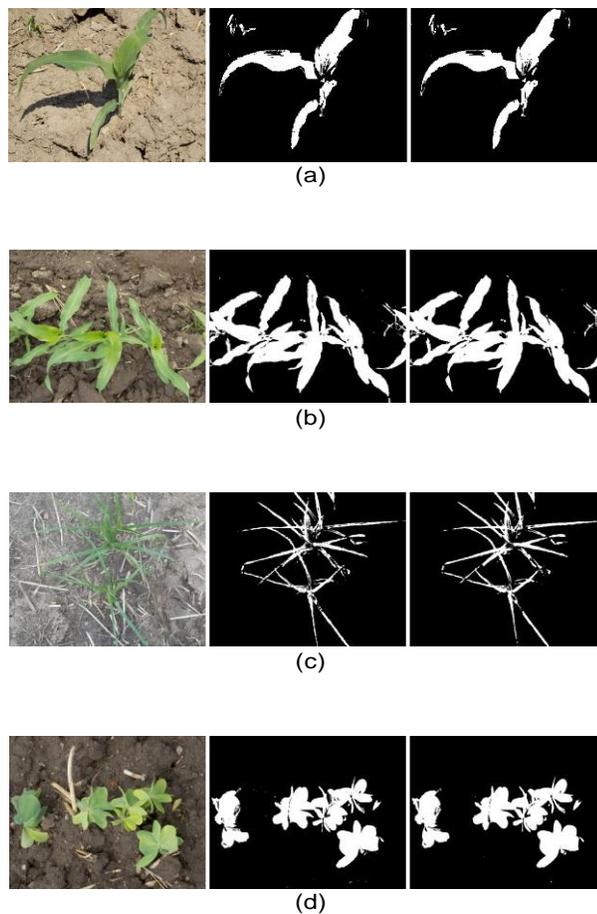


Figura 14. Máscaras predichas utilizando el modelo entrenado con la arquitectura U-Net. La primera columna representa la imagen RGB original; la segunda columna es la etiqueta generada por el generador automático de máscaras propuesto. La tercera columna representa la máscara predicha usando el modelo de U-Net.

Las máscaras permiten aislar el objeto de interés (maíz y malezas) del fondo, lo que reduce la complejidad de la tarea. Entonces, el algoritmo de detección de objetos procesa solo las regiones de interés en lugar de buscar en toda la imagen. Esta es la idea detrás de la arquitectura propuesta de cuasi segmentación de instancias presentada en este trabajo.

3.4.3 YOLO

El detector de objetos YOLOv5 (DoCong, 2021) se utilizó para detectar maíz y malezas considerando el 80% de los datos para el entrenamiento y el 20% para las pruebas. Utilizamos imágenes RGB y BB como se menciona en el cuadro 4. Las imágenes se redimensionan a una resolución de 640 × 480 píxeles. Utilizamos YOLOv5 con pesos preentrenados de MS-COCO (Lin et al., 2014). De manera similar a YOLOv1, YOLOv5 se considera un detector de objetos de una sola etapa, ya que aborda la tarea de detección de objetos como un problema de regresión; es decir, para una imagen dada, la red asigna directamente las probabilidades y la ubicación (coordenadas BB) de un objeto, en lugar de calcular una red de propuestas de región para estimar la ubicación de un objeto y luego predecir cada región. Como resultado, YOLO es un detector de objetos rápido adecuado para la detección en tiempo real.

La Fig. 15 muestra algunos ejemplos de detección de maíz y malezas con el modelo basado en YOLOv5. El algoritmo es capaz de distinguir entre maíz y malezas, logrando un mAP = 0.918 en la detección de maíz y un mAP = 0.871 para la detección de malezas, con un mAP general de 89.45. El modelo requiere 83.5 MB para ser almacenado. Su tiempo de inferencia es de 0.0047 segundos por imagen.

4.4. La arquitectura propuesta de segmentación cuasi-instancial

La arquitectura propuesta de segmentación cuasi-instancial consta de un algoritmo de dos fases. La primera está encargada de la segmentación de hojas utilizando máscaras binarias generadas por una U-Net entrenada con imágenes con una resolución de 512x512. Estas máscaras se combinan con la imagen RGB original para obtener máscaras basadas en RGB. La segunda fase realiza la detección de maíz y malezas en las máscaras basadas en RGB con una resolución de 640

x 480. La Figura 16 muestra algunos resultados de detección de maíz y malezas durante la inferencia utilizando el modelo propuesto. A partir de esta figura, podemos observar que el modelo puede diferenciar el maíz de las malezas, incluso cuando ambas clases están en la misma imagen. El modelo propuesto alcanza un mAP = 0.936 para el caso de detección de maíz, un mAP = 0.909 para el caso de detección de malezas, y un mAP promedio = 0.9225 en general. Este resultado es comparable con los reportados en la literatura. Sin embargo, nuestro modelo puede detectar maíz y malezas en varias etapas de crecimiento, como se muestra en la Fig. 16, ya que nuestro conjunto de datos considera esta información.



Figura 15. Resultados de detección de maíz y malezas utilizando el modelo basado en YOLOv5.

La arquitectura propuesta requiere 0.0196 segundos para la detección durante la inferencia y 94.9 MB para almacenar el modelo. El modelo de segmentación cuasi-instancial propuesto en este trabajo detecta maíz y malezas bajo diferentes

ángulos, vistas y etapas de crecimiento, y determina los límites exactos entre objetos. Esta característica es crucial en caso de que el modelo se integre en un robot eliminador de malezas.

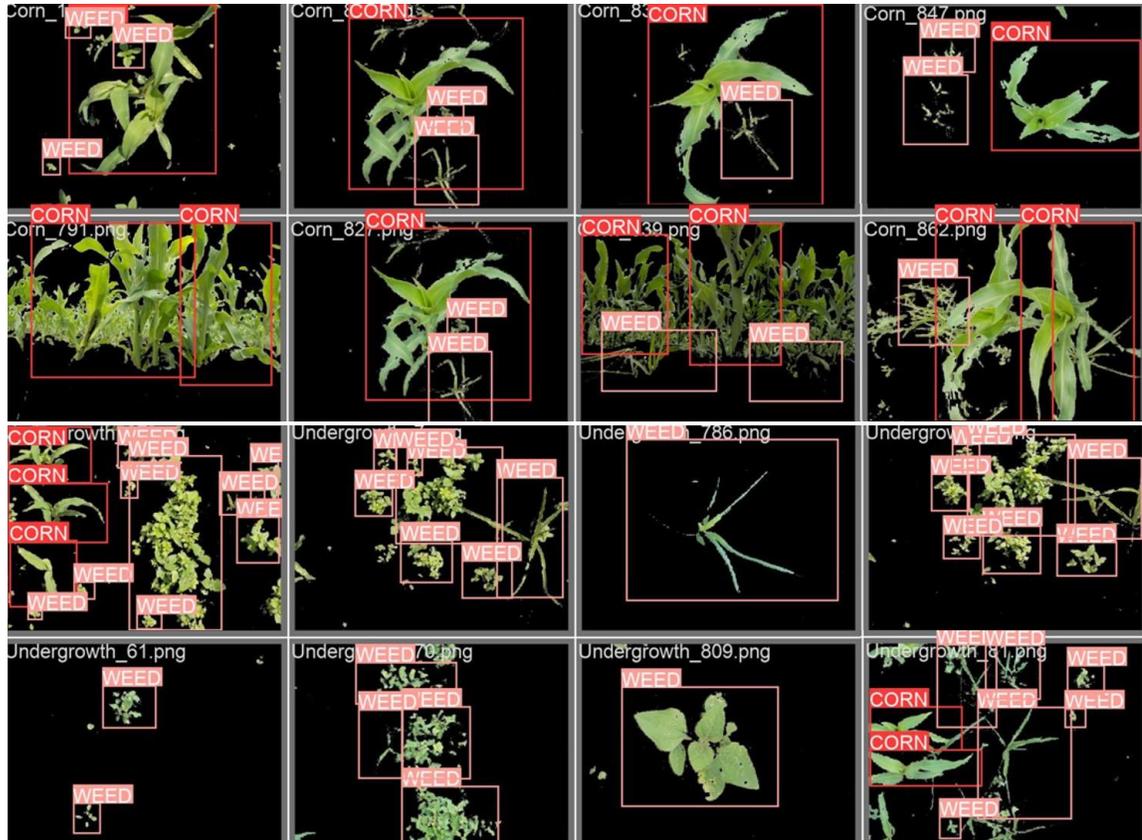


Figura 16. Detección de maíz y malezas utilizando el modelo de la arquitectura propuesta, es decir, la arquitectura de cuasi segmentación de instancias.

3.5 Conclusiones

Este documento propone un conjunto de datos de imágenes de maíz y malezas que considera imágenes desde diferentes ángulos de visión, condiciones climáticas, condiciones de iluminación y etapas de crecimiento. Esta información proporciona una amplia variedad de muestras para generalizar el conocimiento sobre el maíz y las malezas. Además, este trabajo propone un método automático para el etiquetado de máscaras y una arquitectura para la detección de maíz y malezas a nivel de píxel, realizando una tarea de cuasi-segmentación de instancias basada en U-Net y YOLOv5. El modelo propuesto alcanza un mAP = 92.25% con una velocidad de inferencia de 0.0196 y un tamaño de modelo de

94.9MB. Estas métricas prueban el potencial del modelo propuesto para ser integrado en un robot removedor de malezas, como trabajo futuro.

Agradecimientos

Queremos agradecer al CONACYT, la Universidad Autónoma Chapingo, la Universidad Panamericana y la Universidad Veracruzana por su invaluable apoyo durante esta investigación. Además, expresamos nuestra gratitud a Griselda Trinidad Miguel por su apoyo durante la recolección de datos.

Declaración de conflicto de intereses

Los autores declaran que no tienen conflicto de intereses financieros conocidos o relaciones personales que pudieran parecer haber influido en el trabajo reportado en este documento.

3.6 Referencias

Aparna, M., & Radha, D. (2019). Detection of Weed Using Visual Attention Model and SVM Classifier. In *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB)* (pp. 243–253). Springer International Publishing volume 30. doi:[10.1007/978-3-030-00665-5_25](https://doi.org/10.1007/978-3-030-00665-5_25).

Champ, J. et al. (2022). Instance segmentation for the fine detection of crop and weed plants by precision agricultural robots. *Applications in Plant Sciences*, 8 .

Chang, C. L., Xie, B., & Chung, S. C. (2021). Mechanical control with a deep learning method for precise weeding on a farm. *Agriculture*, 11 , 1049. doi:[10.3390/agriculture11111049](https://doi.org/10.3390/agriculture11111049).

DoCong, T. (2021). Evolution of yolo algorithm and yolov5: The state-of-the-art object detection algorithm. In *Information Technology Oulu University of Applied Sciences*, 61. Available at: URL: https://www.theseus.fi/bitstream/handle/10024/452552/Do_Thuan.pdf?isAllowed=y&sequence=2.

Espejo-Garcia, B. et al. (2020). Improving weeds identification with a repository of agricultural pre-trained deep neural networks. *Computers and Electronics in Agriculture*, 175 , 105593. doi:[10.1016/j.compag.2020.105593](https://doi.org/10.1016/j.compag.2020.105593).

Etienne, A. et al. (2021). Deep learning-based object detection system for identifying weeds using uas imagery. *Remote Sensing*, 13 , 5182. doi:[10.3390/rs13245182](https://doi.org/10.3390/rs13245182).

Huang, C., Wu, W., & Lei, Z. (2020). Efficient instance segmentation network. In *2020 IEEE International Conference on Artificial Intelligence and Information Systems (ICAIS)* (pp. 93–101).

- Islam, N. et al. (2021). Early weed detection using image processing and machine learning techniques in an australian chilli farm. *Agriculture*, 11 , 387. doi:[10.3390/agriculture11050387](https://doi.org/10.3390/agriculture11050387).
- Jabir, B., Rabhi, L., & Falih, N. (2021). Rnn- and cnn-based weed detection for crop improvement: An overview. *Foods and Raw Materials*, 9 , 387–396. doi:[10.21603/2308-4057-2021-2-387-396](https://doi.org/10.21603/2308-4057-2021-2-387-396).
- Kamath, R., Balachandra, M., & Prabhu, S. (2020). Paddy crop and weed discrimination: A multiple classifier system approach. *International Journal of Agronomy* , 2020 .
- Khamar, B. S. et al. (2015). Comparative study of skin color detection and segmentation in hsv and ycbcr color space. *Procedia Computer Science*, 57 , 41–48. 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015).
- Lin, T. et al. (2020). Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42 , 318–327
- Lin, T. Y. et al. (2014). Microsoft coco: Common objects in context. In *Computer Vision – ECCV 2014* (pp. 740–755). Springer International Publishing.
- Louargant, M. et al. (2018). Unsupervised classification algorithm for early weed detection in row-crops by combining spatial and spectral information. *Remote Sensing*, 10 , 761. doi:[10.3390/rs10050761](https://doi.org/10.3390/rs10050761).
- Partel, V., Charan, K. S., & Ampatzidis, Y. (2019). Development and evaluation of a low-cost and smart technology for precision weed management utilizing artificial intelligence. *Computers and Electronics in Agriculture*, 157 , 339–350. URL: <https://www.sciencedirect.com/science/article/pii/S0168169918316612>. doi:[10.1016/j.compag.2018.12.048](https://doi.org/10.1016/j.compag.2018.12.048).
- Servicio de Información Agroalimentaria y Pesquera, M. (2021). Producción agrícola. (Date accessed: 9.05.2021). URL: <https://www.gob.mx/siap/acciones-y-programas/produccion-agricola-33119>.
- Quan, L. et al. (2021). An instance segmentation-based method to obtain the leaf age and plant centre of weeds in complex field environments. *Sensors*, 21 .
- Rani, S. et al. (2022). Automated weed detection system in smart farming for developing sustainable agriculture. *International Journal of Environmental Science and Technology* , 19 , 9083–9094.
- Redmon, J. et al. (2016). You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 779–788). IEEE Computer Society.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (pp. 234–241). Springer volume 9351 of LNCS . URL: <http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a> (available on arXiv:1505.04597 [cs.CV]).

SADER (2021). Estima agricultura crecimiento de 2.6 por ciento en la producción de maíz grano en año agrícola 2021. (Date accessed: 04.05.2021). URL: [https://www.gob.mx/agricultura/prensa/estima-agricultura-crecimiento-de-2-6-por-cient o-en-la-produccion-de-maiz-grano-en-ano-agricola-2021](https://www.gob.mx/agricultura/prensa/estima-agricultura-crecimiento-de-2-6-por-cient-o-en-la-produccion-de-maiz-grano-en-ano-agricola-2021).

SEGOB (2020). Dof - diario oficial de la federación. (Date accessed: 31.12.2020). URL: https://www.dof.gob.mx/nota_detalle.php?codigo=5609365&fecha=31/12/2020#gsc.tab=0.

Shirzadifar, A. et al. (2020). Field identification of weed species and glyphosate-resistant weeds using high resolution imagery in early growing season. *Biosystems Engineering*, 200, 200–214. doi:[10.1016/j.biosystemseng.2020.10.001](https://doi.org/10.1016/j.biosystemseng.2020.10.001).

Soltani, N. et al. (2016). Potential corn yield losses from weeds in north america. *Weed Technology*, 30, 979–984. doi:[10.1614/WT-D-16-00046.1](https://doi.org/10.1614/WT-D-16-00046.1).

Subeesh, A. et al. (2022). Deep convolutional neural network models for weed detection in polyhouse grown bell peppers. *Artificial Intelligence in Agriculture*, 6, 47–54. doi:[10.1016/j.aiia.2022.01.002](https://doi.org/10.1016/j.aiia.2022.01.002).

Unal, Z. (2020). Smart farming becomes even smarter with deep learning—a bibliographical analysis. *IEEE Access*, 8, 105587–105609. doi:[10.1109/access.2020.3000175](https://doi.org/10.1109/access.2020.3000175).

Urmashv, B. et al. (2021). Development of a weed detection system using machine learning and neural network algorithms. *Eastern-European Journal of Enterprise Technologies*, 6. doi:[10.15587/1729-4061.2021.246706](https://doi.org/10.15587/1729-4061.2021.246706).

World of Corn, U. (2021). In *United States department of agriculture. National corn growers association*. (Date accessed: 23.5.2021). URL: <https://www.worldofcorn.com/#world-corn-production>.

Wu, Z. et al. (2021). Review of weed detection methods based on computer vision. *Sensors*, 21, 3647. doi:[10.3390/s21113647](https://doi.org/10.3390/s21113647).

Xie, B., Chung, S., & Chang, C. L. (2019). Design and implementation of a modular Alenabled shovel weeder. In *2020 International Symposium on Computer, Consumer and Control (IS3C)* (pp. 236–239). IEEE volume 30. doi:[10.1007/978-3-030-00665-5_25](https://doi.org/10.1007/978-3-030-00665-5_25).

Xiongwei, W., Doyen, S., & Steven, C. (2020). Recent advances in deep learning for object detection. *Neurocomputing*, 396, 39–64.

4. ROBOT PARA NAVEGACIÓN EN CAMPOS DE MAÍZ PARA EL FIELD ROBOT EVENT 2023

D. Ivan Sánchez-Chávez¹; Noé Velázquez-López^{*1}; Mónica E. Berrocal-Aguilar¹;
Guillermo García-Sánchez¹; Alan Hernández-Mercado¹; Omar A.
Avendaño-Lopez¹

¹Universidad Autónoma Chapingo, Carr. Federal México-Texcoco Km 38.5, Texcoco, Z.C. 56230, México.

*email: nvelazquezl@chapingo.mx, teléfono: +52 595 114 3092

Dirección postal: Carr. Federal México-Texcoco Km 38.5, Texcoco, C.P. 56230, México.

Resumen

La navegación en un cultivo de maíz es una tarea crucial para el desarrollo de robots autónomos en la agricultura, con numerosas aplicaciones como la fumigación, el monitoreo del crecimiento y la salud de las plantas, y la detección de malezas y plagas. El Evento de Robots de Campo 2023 continuó desafiando a universidades y otros equipos de investigación para impulsar aún más el desarrollo de algoritmos para robots agrícolas. La Universidad Autónoma Chapingo ha estado desarrollando un robot para diversas tareas agrícolas, con el objetivo de ofrecer una alternativa de bajo costo para trabajar con agricultores mexicanos en el futuro. Para esta edición del FRE, se creó un algoritmo de navegación utilizando un codificador, una IMU, un RPLIDAR y cámaras para recopilar datos para la toma de decisiones. El algoritmo se desarrolló en ROS Melodic, dividiendo la tarea en pasos que se probaron para determinar los movimientos reales del robot. El sistema navega utilizando ROIs (regiones de interés) y el centro de masa para guiar al robot entre las hileras de maíz. Calcula la media de los valores de orientación finales antes de llegar al final de una fila, que se detecta utilizando un RPLIDAR. Para los giros y movimientos en línea recta para llegar a la siguiente fila, se utiliza la orientación como guía. Para detectar plantas para fumigación, se emplean láseres ubicados en cada lado del vehículo. La detección de obstáculos se basa en un modelo entrenado con YOLOv5 y un láser, mientras que la navegación en reversa utiliza una cámara trasera. Durante la competencia, el robot enfrentó desafíos como lidiar con la hierba, el pequeño tamaño de las plantas y la necesidad de utilizar una fuente de energía diferente, lo que afectó su rendimiento.

Palabras clave: visión por computadora, CNN, ROI, navegación autónoma.

ROBOT FOR NAVIGATION IN MAIZE CROPS FOR THE FIELD ROBOT EVENT 2023

D. Ivan Sánchez-Chávez¹; Noé Velázquez-López^{*1}; Mónica E. Berrocal-Aguilar¹; Guillermo García-Sánchez¹; Alan Hernández-Mercado¹; Omar A. Avendaño-Lopez¹

¹Universidad Autónoma Chapingo, Carr. Federal México-Texcoco Km 38.5, Texcoco, Z.C. 56230, México.

*email: nvelazquezl@chapingo.mx, telephone: +52 595 114 3092

postal address: Carr. Federal México-Texcoco Km 38.5, Texcoco, Z.C. 56230, México.

Abstract

Navigation in a maize crop is a crucial task for the development of autonomous robots in agriculture, with numerous applications such as spraying, monitoring plant growth and health, and detecting weeds and pests. The Field Robot Event 2023 continued to challenge universities and other research teams to push the development of algorithms for agricultural robots further. The Universidad Autónoma Chapingo has been developing a robot for various agricultural tasks, aiming to provide a low-cost alternative to work with Mexican farmers in the future. For this edition of the FRE, a navigation algorithm was created using an encoder, an IMU, an RPLIDAR, and cameras to collect data for decision-making. The algorithm was developed in ROS Melodic, dividing the task into steps that were tested to determine the robot's actual movements. The system navigates by using ROIs (regions of interest) and the mass center to guide the robot between maize rows. It calculates the mean of the final orientation values before reaching the end of a row, which is detected using an RPLIDAR. For turns and straight-line movements to reach the next row, the orientation is used as a guide. To detect plants for spraying, lasers located on each side of the vehicle are employed. Obstacle detection relies on a YOLOv5 trained model and a laser, while reverse navigation uses a rear camera. During the competition, the robot faced challenges such as dealing with grass, the small size of the plants, and the need to use a different power source, which affected its performance.

Keywords: machine vision, CNN, ROI, autonomous navigation

4.1 Introducción

El maíz es un cultivo esencial para los agricultores y se cultiva en más de 150 millones de hectáreas en todo el mundo (Kannan, Ismail y Bunawan, 2018). El maíz es un cereal de gran importancia económica y social a nivel mundial porque es fundamental para la alimentación humana, el uso industrial y la alimentación de animales domésticos (Monteiro, Alencar, Souza y Leao, 2021). Tiene relevancia en la seguridad del suministro de alimentos, biodiésel y biocombustibles, exportaciones agrícolas e ingresos, rotación de cultivos, adaptación al cambio climático y fuentes de ingresos para los agricultores. Según el Departamento de Agricultura de los Estados Unidos (USDA), México es el séptimo productor más grande de maíz en todo el mundo (Puente G, 2023).

Actualmente, la demanda agrícola está siendo superada por el creciente número de habitantes en todo el mundo, ya sea debido a la migración de los jóvenes a las grandes ciudades o la disminución de tierras disponibles para el cultivo. Esta necesidad ha despertado un gran interés a nivel mundial en el desarrollo de nuevas tecnologías y avances en el campo de los robots agrícolas para ayudar a lograr una mejor producción de alimentos. Según la FAO (2009) y Calicioglu et al. (2019), un desafío significativo en la agricultura es producir más alimentos como consecuencia del aumento de la población mundial, apuntando a un equilibrio a largo plazo. Por otro lado, Subeesh y Mehta (2021) enfatizan que la agricultura tradicional requiere mucha mano de obra, con limitaciones en las tareas de monitoreo de cultivos. Además, hay una disminución significativa en la mano de obra calificada, razón por la cual los métodos agrícolas tradicionales no son suficientes para lograr la máxima productividad (Bai et al., 2022). Los robots agrícolas presentan una oportunidad para fortalecer los sistemas agroalimentarios al abordar la escasez de mano de obra y reducir las emisiones de CO₂ (Orum et al., 2023).

Muchas operaciones agrícolas requieren que los operadores de maquinaria tengan una gran habilidad para lograr buenas trayectorias y configurar parámetros técnicos en tiempo real (Fujita et al., 2020). Como solución

alternativa, la agricultura moderna introduce robots agrícolas y equipos inteligentes, reemplazando gradualmente las operaciones humanas como la dirección para el desarrollo agrícola futuro (Xie et al., 2022). El interés en los sistemas robóticos agrícolas ha aumentado en los últimos años, promoviendo el desarrollo de vehículos más autónomos e inteligentes en la agricultura. Los robots agrícolas autónomos tienen el potencial de aumentar la eficiencia de la producción agrícola y reducir el consumo de recursos naturales (Khadatkar et al., 2021). En entornos agrícolas, los robots agrícolas se han utilizado para diversas tareas como arar, trasplantar, podar, desmalezar, cosechar, plantar, rociar, fertilizar y más (Mao et al., 2020).

Un aspecto esencial para permitir que un robot trabaje de manera autónoma en el campo es la navegación autónoma. Con este objetivo en mente, se está llevando a cabo investigación en el campo de la navegación. En estos entornos, hay sensores capaces de detectar líneas de cultivo, siendo los más comunes conocidos como LiDAR y cámaras. Aunque LiDAR tiene ciertas ventajas sobre las cámaras debido a las características de los entornos donde deben navegar los robots agrícolas (típicamente entornos abiertos con cambios frecuentes en las condiciones de iluminación), es en este punto donde las cámaras presentan desventajas relacionadas con la variabilidad en los tonos de color que pueden resultar en diferentes niveles de contraste (Nehme et al., 2021).

Para lograr un sistema robótico con las características adecuadas para operar en campos abiertos, es necesario contar con ciertos componentes clave. Estos componentes incluyen: sensores, construcción de mapas y localización, control de movimiento, planificación y navegación. Además del odómetro y LiDAR mencionados anteriormente, los sensores utilizados en el sistema también incluyen una unidad de medición inercial (IMU), que se utiliza para corregir errores de odometría (Feng et al., 2023).

La diferencia de color entre el cultivo y los demás elementos del entorno, como el suelo y el cielo, se puede utilizar para realizar el reconocimiento de las plantas basado en la segmentación de color para identificar las hileras de cultivo.

El robot Voltan utiliza un sistema de Visión por Máquina que segmenta las plantas en tiempo real utilizando un vector de valores de interés en el espacio de color RGB, y luego calcula el punto que utiliza para la orientación usando el centro de masa, para hacer la detección del final de la hilera utiliza un RPLidar, y para medir el desplazamiento del vehículo utiliza un codificador y una IMU que proporciona información sobre su orientación. En este trabajo, se propone un algoritmo que utiliza sensores asequibles y visión por máquina para navegar automáticamente entre las hileras de maíz.

4.2 Materiales y Métodos

4.2.1 Sistema de navegación

Software principal en ROS

Para navegar adecuadamente entre las filas de un campo de maíz, es necesario utilizar diversos sensores que permitan al robot tener información sobre su entorno. En trabajos anteriores se han utilizado RPLidar, IMU, Encoders, como base para los sistemas.

Es importante considerar que la navegación autónoma agrícola es una ingeniería de sistemas compleja, que consta de cuatro tecnologías clave: percepción ambiental, posicionamiento preciso, toma de decisiones y planificación, y control de ejecución (Binbin et al., 2023).

Mun (2019) utiliza una Unidad de Medición Inercial (IMU) para determinar con precisión la posición de una silla de ruedas. Esto se logra midiendo cada eje del acelerómetro y el giroscopio, que luego se alinean con el marco de coordenadas de la silla de ruedas. Los datos se transforman posteriormente utilizando una rotación matricial. Para validar la efectividad de los métodos utilizados para transformar los datos de la IMU y aplicar varios filtros, desarrolló un prototipo utilizando un Arduino Uno, un sensor IMU MPU6050 de 6 grados de libertad y una pantalla OLED. Está programado en el lenguaje de programación basado en C de Arduino.

Zhang et al. (2014) proponen un vehículo autónomo diseñado para operaciones en huertos. Este vehículo está equipado con un codificador en el volante para medir el ángulo de dirección, y un codificador en la rueda motriz, para medir la velocidad angular de las ruedas traseras. Al combinar las lecturas de ambos codificadores, es posible calcular la velocidad del vehículo.

En el desarrollo de un vehículo autónomo, Salmane et al. (2023) utilizaron sensores LiDAR para la detección de objetos en 3D. En este caso, los datos recopilados del LiDAR se integraron con información de una cámara estéreo. Esta integración permitió la medición precisa de distancias a obstáculos y proporcionó una visión precisa del tamaño y volumen de los objetos, lo cual puede ser extremadamente útil al clasificar objetos como automóviles, peatones, etc.

El robot Voltan de la Universidad Autónoma Chapingo tiene como sensores principales una IMU, un encoder, un RPLIDAR y cámaras, y opera con código desarrollado en Robot Operating System (ROS) debido a las ventajas que proporciona, tales como: (1) abstracción de hardware y control de dispositivos de bajo nivel; (2) intercambio de mensajes entre procesos; (3) gestión de paquetes; (4) herramientas para desarrollar, probar y ejecutar código en varios tipos de lenguajes de programación; (5) computación distribuida; (6) reutilización de software; y (7) pruebas rápidas. ROS permite encapsular nodos desarrollados en diferentes lenguajes de programación, facilitando en gran medida el crecimiento incremental del sistema y su prototipado (Saavedra et al., 2023).

La navegación se dividió en diferentes fases que deben realizarse. Primero, el movimiento entre las filas, para esta parte el sistema de visión artificial envía valores para ajustar la trayectoria del vehículo, al mismo tiempo en un nodo ROS diferente, el RPLidar combinado con un codificador detecta las plantas y busca espacios sin plantas para determinar el final de la fila y detener el control del sistema de visión cuando sea necesario. Segundo, el movimiento de la parte final está guiado por el promedio de las últimas 50 orientaciones del robot. La tercera parte es el giro, para esta parte el robot utiliza el giroscopio para realizar un giro

de 90 grados sobre su propio eje calculado usando un promedio de las últimas orientaciones cuando el robot navegaba utilizando el sistema de visión artificial. Cuarto, se mueve en dirección a la siguiente fila midiendo el desplazamiento con un codificador, una vez que se alcanza la distancia deseada, el robot se detiene y realiza otro giro de 90 grados, finalmente navega utilizando el giroscopio y el control de visión artificial comienza de nuevo.

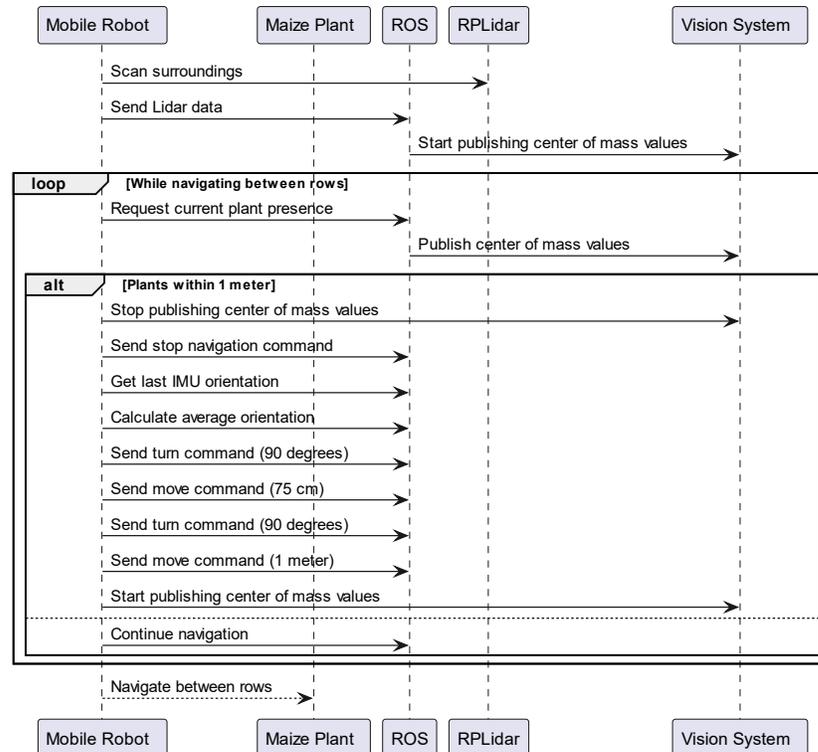


Figura 17. Pasos de la navegación.

Sistema de visión artificial

Para la detección de las plantas, es posible utilizar la visión artificial, las ventajas de este enfoque incluyen el bajo costo de las cámaras y los recursos abiertos como OpenCV. El reconocimiento y detección de hileras de cultivos es una de las tecnologías clave para la navegación automática en el campo, y se ha llevado a cabo mucha investigación por académicos tanto nacionales como internacionales. Según el modelo de procesamiento de imágenes, se pueden clasificar de la siguiente manera: reconocimiento basado en el algoritmo de transformada de Hough (HT), reconocimiento basado en segmentación de color, reconocimiento basado en agrupación de puntos característicos, reconocimiento

basado en métodos de visión estéreo y reconocimiento basado en métodos de aprendizaje profundo (Liu et al., 2023).

Yang et al. (2018) identificaron las líneas centrales de las hileras de cultivos de maíz extrayendo los tallos en la parte inferior de las plantas de maíz. Garcia-Santillan et al. (2018) y Hassanein et al. (2019) utilizaron el método de acumulación de píxeles verdes para extraer hileras de cultivos de maíz e hileras de plantación de colza, respectivamente.

Otro enfoque es el algoritmo de reconocimiento basado en visión estéreo, un algoritmo que utiliza cámaras estéreo para capturar imágenes tridimensionales e identificar hileras de cultivos basándose en la diferencia de altura entre los cultivos y las malezas (Behrje et al., 2018). Dichos algoritmos no son adecuados para el reconocimiento de hileras de cultivos de maíz en plántulas ya que requieren grandes cantidades de cálculo y dependen de la diferencia de altura entre el cultivo y el fondo, como las malezas.

El algoritmo de reconocimiento basado en la agrupación de puntos característicos es un algoritmo para extraer las líneas centrales de las hileras de cultivos utilizando algoritmos de agrupación que agrupan puntos característicos de cultivos similares y cercanos. Jiang et al. (2015) combinaron características geométricas de hileras de cultivos y zonas activas de robots utilizando varias regiones de interés (ROI, por sus siglas en inglés) y extrajeron el centro de masa de las hileras de cultivos mediante métodos de agrupación para obtener líneas centrales de hileras de cultivos. Zhang et al. (2018) utilizaron un algoritmo de agrupación posicional y un método de camino más corto para determinar el conjunto de puntos característicos de la agrupación y extrajeron líneas centrales de hileras de cultivos utilizando regresión lineal.

Para navegar entre las plantas, el sistema principal del robot utiliza la visión artificial; para desarrollar esta parte, fue necesario lograr la segmentación en tiempo real para ignorar los valores del espacio de color que no son importantes y que incluyen el cielo, el suelo y otros elementos. El sistema desarrollado utiliza

el centro de masa y la coordenada en el eje x de la imagen capturada por la cámara; para esto, la cámara está posicionada en el frente y en el centro del robot para que pueda usar los valores de la coordenada x de las ROI para guiar su movimiento. Si el centro del surco está ubicado a la derecha o izquierda de la imagen o video obtenido por una cámara web, el robot corrige para mantener la cámara en el centro de las filas. Fue necesario limitar la visión cuando la cámara no está viendo las filas correctamente, utilizando máscaras en el lado izquierdo y derecho de la imagen analizada y de esta manera no permitir que otras filas entren al cuadro de interés. Para reducir los recursos de cómputo, fue necesario usar una resolución de imagen de 640x480 píxeles. El sistema ha sido desarrollado para ser útil incluso en diferentes condiciones de luz presentes en el exterior en diferentes horas del día. El código se desarrolló en lenguaje C++ con las librerías de OpenCV junto con ROS Melodic.

Encoder

El robot estaba equipado con un encoder modelo E50S8-500-3-T-5, que es un sensor óptico de 5000 pulsos por revolución del eje de la rueda a través de un emisor y receptor infrarrojos. La programación del algoritmo se realizó en el IDE de Arduino, y calculó los grados de avance del codificador (Gr), lo que permite la determinación de la distancia recorrida por el robot. Para ello, se utilizó la resolución del codificador (5000 pulsos por revolución) y, junto con la circunferencia de la rueda, se determinó la distancia recorrida en tiempo real por el robot utilizando las expresiones 1 y 2, respectivamente.

$$Gr = \frac{360}{ppr} * pr \quad (1)$$

Donde: Gr = número de grados de avance del eje del encoder [°], ppr = número total de pulsos por revolución del codificador (5000), [Pulsos/revolución] pr = pulsos registrados por el Arduino, [Pulsos/metro].

$$lr = \frac{p}{360} * Gr \quad (2)$$

Cuadro 5. Diagrama de conexión Encoder-Arduino.

Encoder	Arduino			
	5v	GND	Pin 19	Pin 20
Cable azul		X		
Cable café	X			
Cable blanco			X	
Cable negro				X

Este es un codificador incremental, por lo que, para una mejor captura de datos y prevenir la pérdida de pulsos, el codificador y la IMU se conectaron a un Arduino MEGA, que publica las lecturas en ROS en una computadora a través de comunicación serial.

RPLIDAR

El lidar utilizado para la detección de las plantas funciona conociendo la próxima dirección de giro, esto se escribe en un archivo txt utilizando la misma codificación que la ruta describe en las tareas del FRE 2023 (1R, 3L, etc.), un nodo de ROS busca espacios de más de un metro de largo, por lo que está suscrito a un topic (asunto) que publica las mediciones del codificador, está ubicado en la parte frontal del robot centrado, a una baja altura del nivel del suelo.

IMU (Unidad de Medición Inercial)

Con el continuo desarrollo de la tecnología MEMS (Sistemas Micro-Electro-Mecánicos), un sensor de unidad de medición inercial (IMU) compuesto por un acelerómetro y un giroscopio ha pasado a ser utilizado como una herramienta importante para el posicionamiento de vehículos. La mayoría de los sistemas de navegación modernos utilizan señales de satélite para determinar la posición y están diseñados para compensar con IMUs. Por lo tanto, es un grupo importante de sensores para tener cuando se trabaja con navegación para vehículos.

Las IMU son componentes esenciales para aplicaciones que requieren un seguimiento preciso del movimiento, detección de orientación y conciencia del entorno. Al combinar datos de acelerómetros, giroscopios y magnetómetros, los sensores IMU pueden proporcionar una vista integral del movimiento y la orientación 3D de un objeto en tiempo real. Estos sensores se utilizan ampliamente en robótica para navegación y control, en realidad aumentada para un seguimiento preciso de la cabeza, en drones para un vuelo estable, y en muchas otras aplicaciones donde los datos de movimiento y orientación son críticos (Kurniawan, 2021).

Voltan tiene una IMU MPU6050 que consta de un acelerómetro de 3 ejes y un giroscopio de 3 ejes. Juntos, estos sensores pueden proporcionar la información para determinar el rumbo (heading), cabeceo (pitch) y orientación (orientation) de un objeto. Una unidad de medición inercial (IMU) se puede usar para medir la aceleración y la velocidad angular (Cizmic et al., 2023).

El MPU-6050 cuenta con tres convertidores analógicos a digitales (ADC) de 16 bits para digitalizar las salidas del giroscopio y tres ADC de 16 bits para digitalizar las salidas del acelerómetro. Para un seguimiento preciso de los movimientos rápidos y lentos, las partes cuentan con un rango de escala completo del giroscopio programable por el usuario de ± 250 , ± 500 , ± 1000 y $\pm 2000^\circ/\text{s}$ (grados por segundo, dps por sus siglas en inglés) y un rango de escala completo del acelerómetro programable por el usuario de $\pm 2g$, $\pm 4g$, $\pm 8g$ y $\pm 16g$.

Las mediciones del giroscopio se utilizan para interpretar la actitud y orientación del acelerómetro como un marco de referencia (Sung Sic, Humayun, & Heung-Shik, 2022). Los principales datos utilizados de este sensor fueron la orientación del robot con respecto al plano del suelo, ya que se mencionó que se utilizaba para calcular los giros. Para hacer esto, la IMU está conectada a un Arduino Mega que está conectado a la computadora usando roserial, los datos se publican en un tema con la orientación, a la que puede acceder cualquier otro nodo del sistema ROS.

Sistema eléctrico

Los componentes de potencia eléctrica del robot se seleccionaron teniendo en cuenta que el robot podría trabajar en suelo labrado sin derrapajes excesivos mientras arrastra un abridor de surcos.

El sistema eléctrico en esta configuración incluye los siguientes componentes:

1. Motores DC: Hay dos motores de reducción de engranajes con una relación de transmisión de 1:3 (con catarinas de 9 y 27 dientes) DC de 250W que operan a 12 V cada uno.
2. Controlador de motor Sabertooth 2x60. Cada motor está equipado con un módulo controlador de motor dual Sabertooth 2x60 con una capacidad de 60A. Estos controladores permiten la gestión independiente de la velocidad y dirección del motor.
3. Batería: El sistema está alimentado por una batería de plomo-ácido sellada de 12V de MHB. Tiene una capacidad de 26Ah, que indica su capacidad de almacenamiento de energía.
4. Cableado: Se utiliza cableado para conectar los diversos componentes del sistema eléctrico, permitiendo el flujo de energía desde la batería hasta los motores y controladores.

Este sistema eléctrico está diseñado para proporcionar energía y control a los motores DC, permitiéndoles impulsar el vehículo o la maquinaria a la que están conectados. Los controladores de motor desempeñan un papel crucial en la gestión de la velocidad y la dirección de los motores, mientras que la batería sirve como la fuente de energía principal.

Componentes mecánicos

Para la construcción del robot agrícola multitarea "Voltan", se utilizó aluminio de 0.125 pulgadas (3.175 mm) para el cuerpo y el chasis para lograr un diseño ligero que resultara en un menor consumo de batería, estos diseños se describen en

los modelos de utilidad MX 60106 B (Reyes Amador y Velázquez López, 2021) y MX 58211 B (Reyes Amador y Velázquez López, 2020a). Se colocaron casquillos de acero en el chasis, donde se instalaron 24 rodamientos de bolas para posicionar los ejes de transmisión que transmiten el movimiento generado por los motores.

El movimiento del vehículo es del tipo skid-steer (dirección deslizante), lo que significa que ejecuta giros ajustando las velocidades de los dos lados del robot.

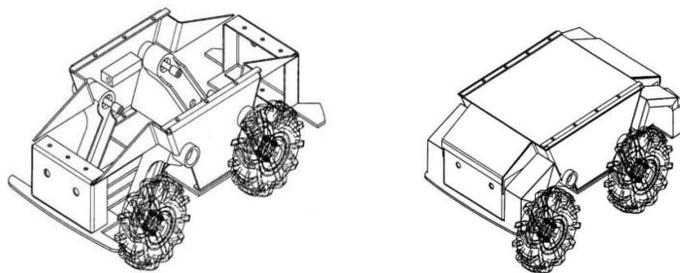


Figura 18. Diagramas de carrocería y chasis.

Para lograr esto, cada uno de los dos motores eléctricos se utiliza para controlar dos ruedas en cada lado.

El sistema de transmisión es por cadena paso 35 con una catarina en la parte superior de 9 dientes y una catarina de 27 dientes en la parte inferior. Este diseño se replica para impulsar las 4 llantas. Esta configuración genera movimiento en llantas agrícolas de tamaño 3.50-4, montadas en rines con un diámetro de 4 pulgadas (101,6 mm).

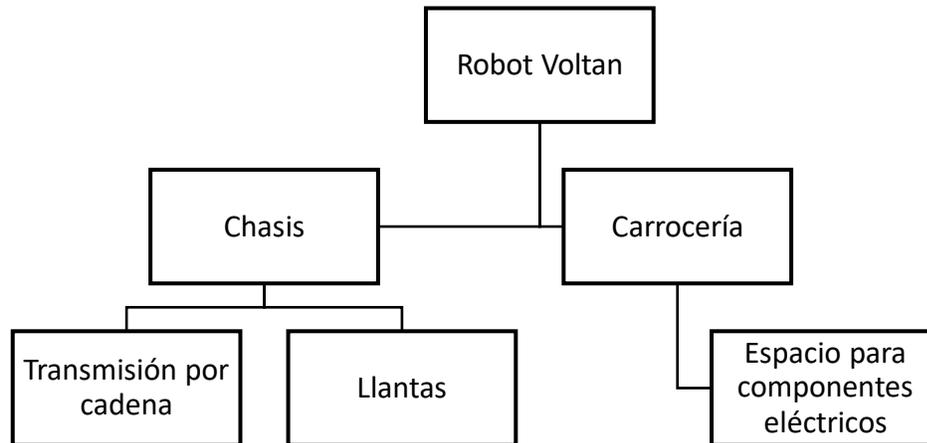


Figura 19. Componentes mecánicos.

Para la conexión entre el chasis y las ruedas, se utilizó un brazo de acero, construido con un perfil tubular rectangular de 1.75 pulgadas x 0.75 pulgadas (44.45 mm x 19.05 mm) para cada rueda.



Figura 20. Neumático agrícola 3.50-4.

Con el objetivo de garantizar un rendimiento adecuado en terrenos irregulares, se diseñó y construyó un sistema de suspensión para vehículos terrestres, sean autónomos o no, el cual se describe en la patente mexicana MX 4369 B (Reyes Amador y Velázquez López, 2020b). Este sistema se caracteriza por el uso de un arreglo triangular compuesto por tres resortes helicoidales operados por tensión. Otro componente importante son los neumáticos, que están diseñados para uso agrícola y tienen un patrón de banda de rodadura especial para un mejor agarre en el suelo.

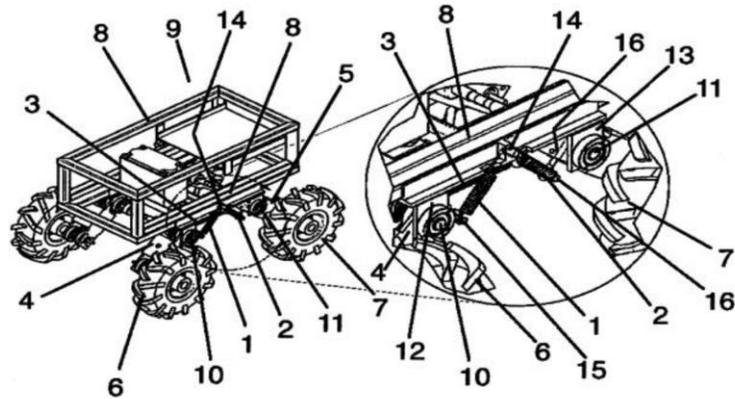


Figura 21. Sistema de suspensión patente MX4369B.

Software de control principal en ROS

Todos los sensores y actuadores se comunican con ROS Melodic en una computadora portátil marca Dell con 8 GB de memoria RAM, 471.4 GB de disco de estado sólido, procesador Intel core i7-4510U de 2.00 GHz x 4, y gráficos Intel HD 4400, que realiza el control. El codificador y la IMU publican sus datos en un topic, lo mismo ocurre con las cámaras que funcionan en nodos especializados para tener la información abierta a los nodos que envían comandos a los motores. Otro nodo se utiliza para la detección del final de la fila, esto utiliza información del RPLIDAR y el codificador, y cuando hay un espacio de más de un metro, publica un mensaje de fin de fila para detener la navegación usando visión artificial. Hay un nodo que verifica el paso en el que se encuentra la navegación y envía comandos para realizar los giros y desplazamientos para entrar en una nueva fila. Este nodo también calcula la media de los valores de orientación para saber cuántos grados necesita el robot para colocarse en una posición perpendicular a las filas. Cuando el sistema mide la distancia necesaria para entrar en la próxima fila, se guía por los valores de la orientación calculada.

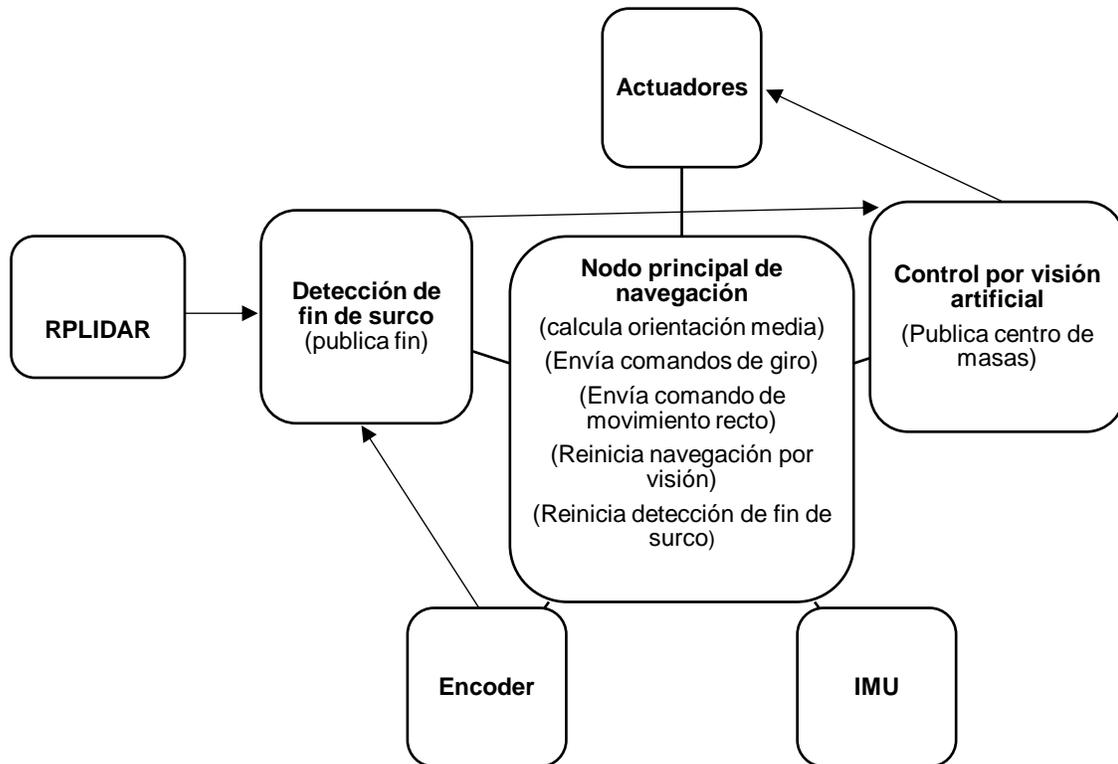


Figura 22. Nodos usados para la navegación de Voltan.

Los nodos principales son:

1. Nodo de navegación. Este nodo sigue la secuencia utilizada para navegar en el campo, calcula la orientación del robot para los giros y movimientos rectos, y reinicia otros nodos cuando no son necesarios. También lee la ruta de un archivo .txt con instrucciones para el próximo giro y número de filas a recorrer.
2. Detección de fin de fila. Detecta las áreas sin plantas y determina si se ha cubierto la distancia de la fila y el lado del próximo giro tiene espacio para los giros.
3. Control de visión artificial. Segmenta las plantas, calcula el centro de masa y publica las coordenadas del punto para guiar el robot.
4. Encoder. Está conectado a un Arduino que publica los datos de distancias recorridas en ROS.
5. IMU. Está conectado a un Arduino que publica los datos en ROS.

6. RPLIDAR. Está conectado a la computadora y publica lecturas del láser en ROS.

4.2.2 Sistema de Riego por Aspersión

Para el sistema de riego localizado, el robot fue equipado con dos sensores Tf-Lidar Plus, colocando un sensor en cada lado posterior del robot. El algoritmo fue programado utilizando Arduino IDE. Se utilizó la biblioteca TF-mini para leer individualmente los sensores. La desventaja de usar estas bibliotecas es que solo se puede leer un sensor por microcontrolador. Por esta razón, se utilizó un microcontrolador esclavo para cada sensor, y la lectura de cada sensor se envió a través de comunicación serial a un microcontrolador maestro donde se realizaba la toma de decisiones.

El cuadro 6 de Mediciones del Esclavo muestra la conexión realizada entre el sensor y el esclavo, donde se puede observar que el tipo de comunicación entre los dos elementos es serial, utilizando el puerto serial número 1.

Cuadro 6. Conexiones del lidar TFmini.

TF-Mini	Arduino (esclavo)			
	5v	GND	Pin 18	Pin 19
Cable rojo	X			
Cable negro		X		
Cable verde RX			X	
Cable azul TX				X

El algoritmo de los microcontroladores esclavos fue programado de manera que cuando un objeto está a una distancia igual o menor de 40 cm, se envía una señal al microcontrolador igual a 1; de lo contrario, se envía un valor igual a 0.

En el microcontrolador maestro, se midió la distancia hacia adelante del vehículo y también recibió señales de los dos esclavos. El diagrama de conexión se muestra como en el cuadro 7.

Cuadro 7. Conexión del controlador Arduino maestro.

		MASTER							
		Rx	Tx	Rx	Tx	GND	5V	20	21
Esclavo 1	Rx		X						
	Tx	X							
Esclavo 2	Rx				X				
	Tx			X					
	Azul					X			
Encoder	Café						X		
	Negro							X	
	Blanco								X

El principio de funcionamiento del algoritmo se muestra en la Figura 23. El algoritmo se divide en dos rutinas, utilizando la primera rutina cuando el robot entra entre filas de plantas y se ejecuta una vez.

Los Arduinos esclavos tienen la función de monitorear los lados del robot para la presencia o ausencia de plantas. Los sensores Lidar detectan objetos una vez que el vehículo comienza a navegar. Cuando un objeto aparece a una distancia igual o menor de 40 cm, se envía un valor de estado de 1 a través de la comunicación serial. Por el contrario, si no hay ningún objeto dentro de ese rango, se envía un valor de 0.

Usando los datos recibidos por el Arduino maestro, se determinan las acciones para activar o desactivar los relés que controlan los aspersores. Antes de ejecutar cualquier acción, se verifica el valor de la bandera para determinar si está en la rutina de entrada de filas o ya dentro de ella.

Si ambos sensores Lidar detectan la presencia de un objeto a ambos lados del sistema de aspersión y la bandera tiene un valor de 1, se ejecuta un retraso de 5 segundos. Este retraso compensa la diferencia de tiempo ya que los sensores y los aspersores no están ubicados en la misma posición. Con esta compensación, hay tiempo para que el aspersor llegue a la planta que el sensor ha monitoreado.

El aspersor se enciende 2 segundos antes de llegar a la planta y continúa rociando durante 2 segundos adicionales. La activación o desactivación de los aspersores depende de la detección o no detección por los sensores Lidar. Cuando se completa esta rutina, indica que el robot está ahora entre filas, por lo que el valor de la bandera se actualiza a 2. Cuando la bandera toma este valor, el algoritmo comienza a ejecutar la rutina en la que el retraso es de 3 segundos, y esta rutina continúa durante toda la navegación.

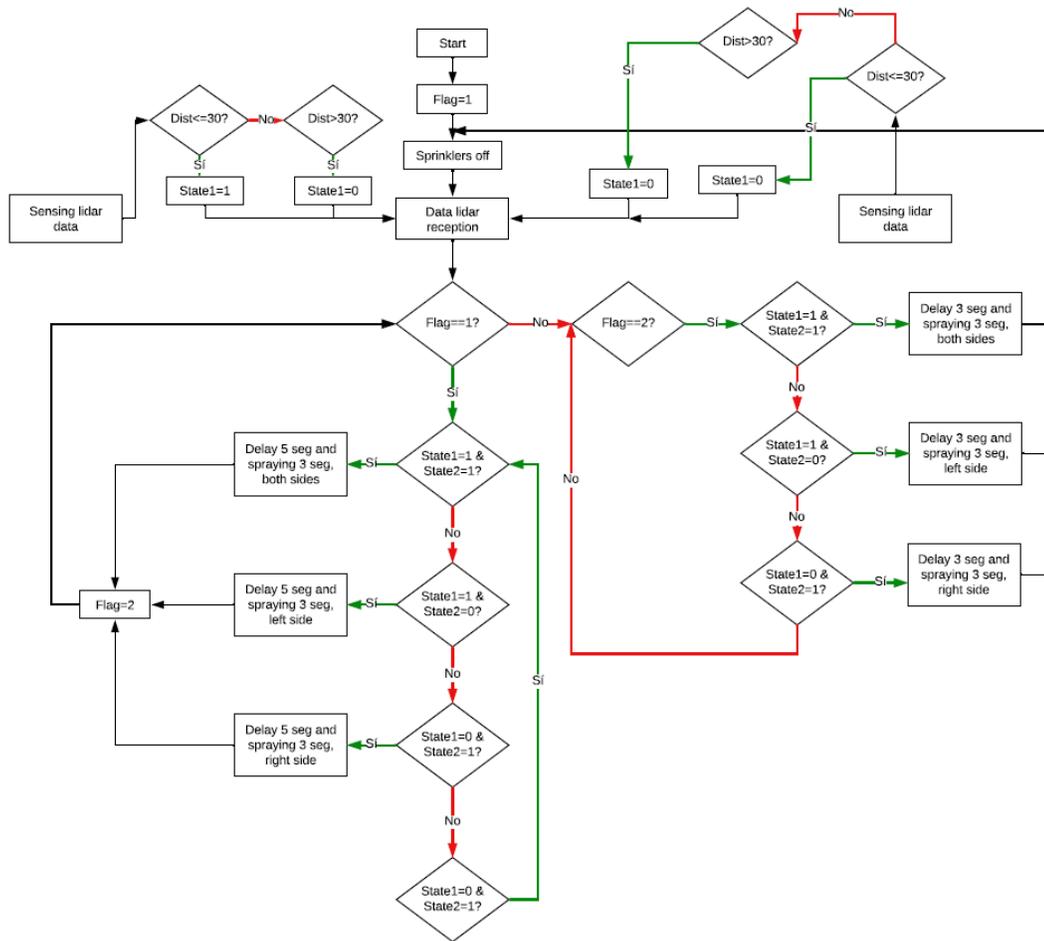


Figura 23. Diagrama de flujo del sistema de aspersión con detección de la presencia de plantas de maíz.

4.2.3 Sistema de Detección de Imágenes

Para detectar los obstáculos representados por imágenes de 3 categorías: humanos, ciervos y otra categoría seleccionada por cada equipo en la competición, utilizamos una arquitectura de Red Neuronal Convoluta, YOLO V5 (You Only Look Once), un modelo popular en visión por computadora. La detección de objetos implica crear características a partir de imágenes de entrada. Estas características se alimentan a través de un sistema de predicción para dibujar cajas alrededor de los objetos y predecir sus clases. El modelo YOLO fue el primer detector de objetos que unió el procedimiento de predecir cajas delimitadoras con etiquetas de clase en una red diferenciable de extremo a extremo.

La red YOLO consta de tres partes principales:

1. Columna vertebral (Backbone): Una red neuronal convolucional que agrega y forma características de imagen en diferentes granularidades.
2. Cuello (Neck): Una serie de capas para mezclar y combinar características de imagen para pasarlas adelante a la predicción.
3. Cabeza (Head): Consume características del cuello y realiza pasos de predicción de caja y clase (Solawetz, 2020).

Para comenzar el entrenamiento, es obligatorio tener un conjunto de datos de imágenes y etiquetar el conjunto de datos. Etiquetamos el conjunto de datos utilizando el software "make sense", donde marcamos en una caja el objeto de interés y colocamos una etiqueta con la categoría de este objeto.

Para la tarea, usamos cuatro categorías: humano, ciervo, cabra y gallo; creamos un conjunto de datos de imágenes con estos objetos, el tamaño de esto es de 1085 imágenes entre las tres categorías, cada foto contiene la etiqueta correspondiente para el objeto que contiene. Este es el conjunto de entrenamiento.

Para el conjunto de validación, separamos 84 imágenes con estas tres categorías y objetos que no corresponden a ninguna de las cuatro categorías.

El procesamiento de imágenes se realizó en Google Collab, utilizando un modelo preexistente de YOLO v5 programado en Python 3. Establecimos un entrenamiento de 100 épocas.

Cuando se completó el entrenamiento, obtuvimos un archivo con los valores correspondientes para cada categoría y, importamos este archivo en un programa principal de Python donde es posible obtener datos de una cámara y hacer las cajas para señalar los objetos que son parte de una de las categorías de interés.

Cuando un objeto aparece en la imagen, el software lo reconoce y trata de clasificarlo en una de las categorías de interés, si el objeto es parte de ellas, aparece en una caja con un texto con el nombre de la categoría a la que pertenece. Durante la tarea, usamos una cámara conectada a una computadora portátil montada en el robot y las señales de detección de objetos fueron las etiquetas y las cajas alrededor de cada objeto.



Figura 24. Detección de una cabra utilizando el código desarrollado.

4.3. Resultados

Esta sección incluye una descripción del rendimiento del robot Voltan en la competencia Field Robot Event 2023 en las 4 tareas diferentes, pero también incluye datos de las pruebas de los pasos utilizados para la navegación en el campo de maíz simulado.

4.3.1 Tarea 1: navegación

Para esta tarea, el robot tuvo problemas con el tamaño de las plantas, que eran demasiado pequeñas para que la cámara y el RPLIDAR las detectaran utilizando las posiciones de los sensores y los códigos programados, por lo que el equipo cambió su ubicación. Otro desafío fue la presencia de césped en el campo, que creó ruido durante la segmentación de las filas de plantas. Las baterías tuvieron

que ser compradas en Europa debido a la restricción de la aerolínea para este tipo de dispositivos, que no podían volar a Eslovenia.

En el momento de la tarea, el robot tuvo problemas para seguir las secciones curvas de las filas debido al ruido en el sistema de segmentación y la fuente de energía diferente, lo que provocó que se desviara del camino. La posición final para la tarea fue el 11^o lugar.

4.3.2 Tarea 2: tratamiento (aspersión) de las plantas

Para abordar el problema con la presencia de césped y árboles, la visión de la cámara fue restringida mediante el uso de máscaras en el código OpenCV, solo utilizó datos de las secciones inferiores izquierda y derecha de la cámara. Pero el uso de solo una batería cambió la energía proporcionada por el controlador a los motores, causando problemas en el movimiento y la corrección de la posición.

Durante la prueba de la tarea 2, el robot no pudo entrar al campo para iniciar la navegación debido a los bajos valores de PWM y la energía de la batería. Como resultado, el resultado final se registró como "DNS" (did not start).

4.3.3 Tarea 3: detección y reconocimiento de posibles obstáculos

Para esta tarea, utilizamos el modelo YOLOv5 que habíamos entrenado. El sistema encontró problemas con el tamaño de las imágenes, lo que lo hacía incapaz de detectar categorías de imágenes a distancia. En consecuencia, el robot alcanzó la 7^a posición.

4.3.4 Tarea 4: obstáculos estáticos y dinámicos.

Para esta tarea, se agregó una cámara de navegación trasera. Funciona utilizando una versión invertida del sistema principal de visión por máquina, permitiendo que el robot retroceda al encontrar obstáculos y luego entre en la siguiente fila. Además, se integró un láser para detectar obstáculos en el campo dentro de una distancia de 40 cm, deteniendo el movimiento del robot para la detección basada en cámara.

A pesar de estas mejoras, el robot continuó experimentando problemas relacionados con valores bajos de PWM y energía de la batería, que afectaron la precisión de sus movimientos. Como resultado, el robot terminó en la 8ª posición.

4.3.5 Pruebas

Giros usando la IMU

Uno de los pasos en el sistema de navegación implica hacer giros de 90 grados. Para probar este aspecto, se midieron los ángulos de giro reales en un piso sólido utilizando una escuadra de metal y un transportador. Las ruedas en un lado del robot se alinearon con marcas de referencia, y el robot luego rotó utilizando la IMU para detenerse después de 90 grados. Se registraron los grados reales de rotación. Este proceso se repitió 20 veces para giros tanto a la izquierda como a la derecha. Los resultados se muestran en el cuadro 8.

Cuadro 8. Ángulos medidos para giros a la izquierda del robot.

Repetición	Ángulo real (°grados)	Ángulo final sensor (°grados)	Ángulo objetivo (°grados)	Ángulo medido por el sensor (°grados)	Error
1	93	90.14	90	90.14	2.86
2	96	180.15	180.14	90.01	5.99
3	93	270.22	270.15	90.07	2.93
4	94	360.29	360.22	90.07	3.93
5	96	450.37	450.29	90.08	5.92
6	95	540.39	540.37	90.02	4.98
7	95	630.47	630.39	90.08	4.92
8	94	720.55	720.47	90.08	3.92
9	94	810.57	810.55	90.02	3.98
10	95	900.61	900.57	90.04	4.96
11	96	990.66	990.61	90.05	5.95
12	95	1080.78	1080.66	90.12	4.88
13	95	1170.87	1170.78	90.09	4.91
14	94	1260.89	1260.87	90.02	3.98
15	96	1351	1530.89	90.11	5.89
16	95	1441.01	1441	90.01	4.99
17	95	1531.1	1531.01	90.09	4.91
18	94	1621.22	1621.1	90.12	3.88
19	96	1711.32	1711.22	90.1	5.9

20	96	1801.41	1801.32	90.09	5.91
Mean	94.85			90.0705	
SD	0.98808693			0.039666372	
MAE					4.7795

Se puede inferir que el error está entre 2 y 6 grados más que el giro deseado, por lo que se tomó en consideración para calcular un ángulo objetivo más adecuado. Utilizando el error promedio, incluimos un ángulo objetivo de 85 grados en el código para los giros a la izquierda. Para las mediciones registradas para los giros a la derecha, los datos se muestran en el cuadro 9.

Cuadro 9. Ángulos medidos para giros a la derecha del robot.

Repetición	Ángulo real (°grados)	Ángulo final sensor (°grados)	Ángulo objetivo (°grados)	Ángulo medido por el sensor (°grados)	Error
1	92	90.03	90	90.03	1.97
2	93	180.11	180.03	90.08	2.92
3	92	270.12	270.11	90.01	1.99
4	93	360.26	360.12	90.14	2.86
5	92	450.31	450.26	90.05	1.95
6	92	540.33	540.31	90.02	1.98
7	93	630.52	630.33	90.19	2.81
8	92	720.62	720.52	90.1	1.9
9	93	810.71	810.62	90.09	2.91
10	93	900.72	900.71	90.01	2.99
11	92	990.78	990.72	90.06	1.94
12	92	1080.82	1080.78	90.04	1.96
13	93	1170.91	1170.82	90.09	2.91
14	94	1260.99	1260.91	90.08	3.92
15	92	1351.05	1350.99	90.06	1.94
16	91	1441.12	1441.05	90.07	0.93
17	93	1531.21	1531.12	90.09	2.91
18	92	1621.29	1621.21	90.08	1.92
19	92	1711.33	1711.29	90.04	1.96
20	94	1801.37	1801.33	90.04	3.96
Mean	92.5			90.0685	
SD	0.76088591			0.04368247	
MAE					2.4315

Se puede inferir que el error está entre 0.9 y 4 grados más que el giro deseado, por lo que se tomó en consideración para calcular un ángulo objetivo más adecuado. Utilizando el error promedio, incluimos un ángulo objetivo de 87 grados en el código para los giros a la izquierda.

Prueba de desplazamiento de fila a fila

Para probar el código para el movimiento recto al final de la fila, se midieron los movimientos al final de la fila y hasta que el robot esté en la nueva fila para reiniciar la navegación utilizando visión artificial. Para que el robot termine en una buena posición final, es necesario que el cálculo del ángulo de orientación promedio que está en la parte final de la fila se haga bien, porque este ángulo es la referencia para terminar en una posición paralela en una nueva fila. Esta parte se dividió en 5 movimientos, como se puede ver en la figura 25.

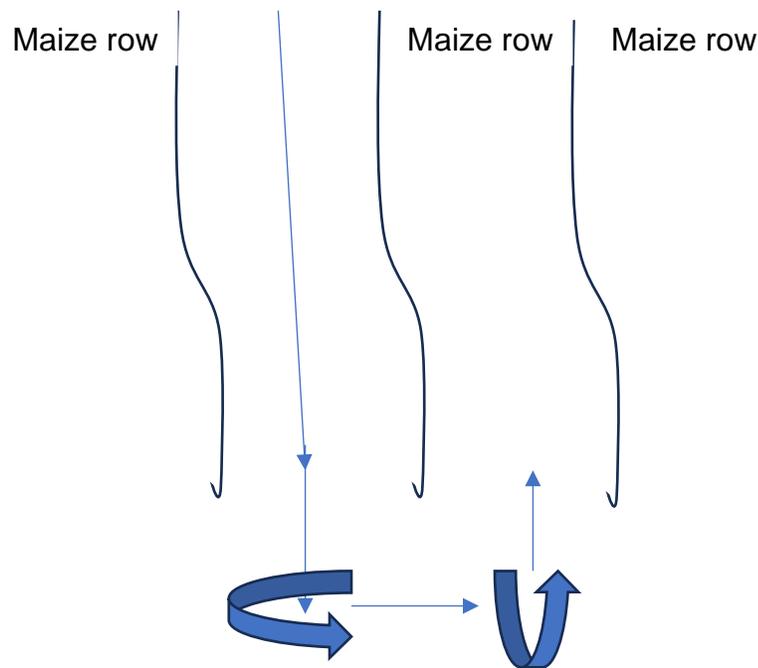


Figura 25. Movimientos del robot para alcanzar la siguiente fila.

Para conocer el desplazamiento real después de cada giro al final de la fila, se construyó un modelo utilizando rectángulos de papel verde que simulaban las plantas, unidos a una cuerda de 5 metros con un espacio de 40 cm entre ellos. El robot tenía el objetivo de navegar utilizando el sistema de visión artificial

durante 5 metros y luego iniciar la rutina de giros, primero se utilizó un desplazamiento de 0.6 metros entre filas.

Cuadro 10. Desplazamientos del robot medidos en movimiento.

repetición	des	dr1	ddg1	dr2	ddg2	dr3
1	5.5	1.06	0.33	0.61	nd	0.52
2	5.2	1.01	0.28	0.57	0.34	0.5
3	5.14	1.01	0.4	0.68	0.26	0.63
4	4.98	1.05	0.27	0.6	0.25	0.55
5	4.3	1.08	0.28	0.59	0.23	0.57
6	4.64	1.1	0.27	0.56	0.23	0.55
Media	4.96	1.051667	0.305	0.601667	0.262	0.553333
SD	0.391578	0.033375	0.04717	0.038909	0.040694	0.041096

Des=desplazamiento entre surcos, dr=desplazamiento recto, ddg=desplazamiento después del giro

A partir de estos resultados, se puede observar que la distancia promedio que el robot se desplaza después de los giros de 90 grados es de 0.3 m para el primero y 0.26 m para el segundo. Utilizamos estos valores para establecer el desplazamiento recto necesario para moverse a las siguientes filas. Repetimos este experimento con una distancia diferente para el dr2.

Cuadro 11. Desplazamientos del robot medidos en movimiento.

repetición	des	ddg1	dr2	ddg2	dr3
1	5.03	0.47	1.6	0.38	0.5
2	5.3	0.38	1.49	0.34	0.49
3	5.27	0.32	1.69	0.27	0.5
4	4.85	nd	1.56	0.35	0.5
5	4.56	0.45	1.63	0.4	0.6
6	5.25	0.29	1.46	0.41	0.51
7	4.62	0.42	1.54	nd	0.53
Media	4.982857	0.388333	1.567143	0.358333333	0.518571

SD	0.312181	0.358608	0.071949	0.07994	0.051153
----	----------	----------	----------	---------	----------

Se repitió lo mismo una tercera vez, pero solo considerando los últimos 5 movimientos.

Cuadro 12. Desplazamientos del robot medidos en movimiento.

repetición	ddg1	dr2	ddg2	dr3
1	0.22	0.55	0.37	0.51
2	0.28	0.48	0.3	0.53
3	0.34	0.6	0.33	0.57
4	0.37	0.53	0.23	0.52
5	0.28	0.54	0.4	0.5
Media	0.298	0.54	0.326	0.526
SD	0.058481	0.043012	0.065803	0.027019

Utilizando el número total de ddg1 y ddg2, obtuvimos los desplazamientos promedio. Esta es una mejor representación de los movimientos reales del robot.

Cuadro 13. Promedio de ddg1 y ddg2.

	Ddg1	Ddg2
mean	0.33235294	0.318125
SD	0.07163531	0.06554579

Prueba de detección de plantas

En esta prueba se utilizó el mismo modelo que emplea papel para representar las plantas, se anotaron las plantas a las que se aplicó agua. La distancia entre las líneas con plantas era de 0.8 m. En el cuadro 14 se puede ver que el sistema detectó todos los papeles mientras el robot estaba navegando.



Figura 26. Voltan con el aspersor para la competencia.

Cuadro 14. Resultados de la detección con láseres.

Número de papeles	Papeles rociados	Porcentaje de éxito
36	36	100 %

Como se puede ver, los láseres funcionaron bien con una distancia de detección de 0.4 m.

Métricas de CNN

Después del entrenamiento de la red YOLOv5 utilizada, los resultados fueron como se muestra en el cuadro 15.

Cuadro 15. Métricas del modelo YOLOv5 entrenado.

Clase	imágenes	instancias	P	R	mAP50	mAP50-95
Todas	84	93	0.46	0.49	0.523	0.352
Persona	84	27	0.967	1	0.995	0.682
Cabra	84	21	0.0231	0.022	0.187	0.143
Ciervo	84	16	0.152	0.25	0.207	0.142
Gallo	84	29	0.696	0.69	0.702	0.442

Los resultados muestran que la CNN tiene problemas con las cabras y los ciervos, con una precisión baja, la persona puede ser detectada mejor que el

resto de las clases y el gallo fue una mejor opción para proponer para la clase. Bueno, estos problemas provienen del bajo número de imágenes utilizadas para entrenar la red.

4.4 Conclusión

El sistema de visión por computadora para la navegación es una buena alternativa económica al uso de RPLIDAR, pero necesita funcionar en una computadora más potente. La inclusión del IMU para guiar los movimientos rectos funciona bien.

La navegación en la prueba real fue deficiente y presentó desafíos a considerar para futuras ediciones de la competencia y trabajo en un campo agrícola real. El uso de máscaras para limitar el marco fue una buena adición al sistema.

Las mediciones de los desplazamientos reales muestran una posible pérdida de datos del codificador.

El sistema, tal como está en este momento, no es confiable y necesita incluir al menos un codificador más y una versión actualizada de ROS.

Los algoritmos deben incluir la combinación de los sensores (RPLIDAR, cámara, encoder e IMU) en la detección del final de la fila y la navegación entre plantas, el cálculo de la velocidad de corrección y los valores de corriente podrían ser una gran adición al sistema que es necesario incluir en futuras versiones del sistema.

Agradecimientos

Agradecimiento a la Universidad Autónoma Chapingo, su comunidad y autoridades, que ha hecho posible el desarrollo del vehículo Voltan y también la participación del equipo Chapingo en el FRE 2023.

4.5 Referencias

Bai, Y., Zhang, B., Xu, N., Zhou, J., Shi, J., & Diao, Z. (2023). Vision-based navigation and guidance for agricultural autonomous vehicles and robots: A review. *Computers and Electronics in Agriculture*, 205, 107584.

Behrje, U., Himstedt, M., & Maehle, E. (2018). An autonomous forklift with 3D time-of-flight camera-based localization and navigation. In: 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV), 1739-1746. doi:10.1109/ICARCV.2018.8581085.

Calicioglu, O., Flammini, A., Bracco, S., Bellù, L., & Sims, R. (2019). The Future Challenges of Food and Agriculture: An Integrated Analysis of Trends and Solutions. *Sustainability*, 11(1), 222. Doi: <https://doi.org/10.3390/su11010222>

Cizmic, D., Hoelbling, D., Baranyi, R., Breiteneder, R., & Grechenig, T. (2023). Smart Boxing Glove "RD a": IMU Combined with Force Sensor for Highly Accurate Technique and Target Recognition Using Machine Learning. *Applied Sciences*, 13(16), 1-16. doi:<https://doi.org/10.3390/app13169073>

Food and Agriculture Organization (FAO). (2009). La agricultura mundial en la perspectiva del año 2050. *Cómo alimentar al mundo en 2050*, 4. Disponible en: https://www.fao.org/fileadmin/templates/wsfs/docs/Issues_papers/Issues_papers_SP/La_agricultura_mundial.pdf

Feng, X., Liang, W. J., Chen, H. Z., Liu, X. Y., & Yan, F. (2023). Autonomous Localization and Navigation for Agricultural Robots in Greenhouse. *Wireless Personal Communications*, 1-15. Doi: <https://doi.org/10.1007/s11277-023-10531-z>

Fujita, S., Emaru, T., Ravankar, A. A., & Kobayashi, Y. (2020). Development of Robust Ridge Detection Method and Control System for Autonomous Navigation of Mobile Robot in Agricultural Farm. In *Symposium on Robot Design, Dynamics and Control* (pp. 16-23). Cham: Springer International Publishing.

Garcia-Santillan, I., Guerrero, J.M., Montalvo, M., et al. (2018). Curved and straight crop row detection by accumulation of green pixels from images in maize fields. *Precision Agriculture*, 19(1), 18–41.

Hassanein, M., Khedr, M., & El-Sheimy, N. (2019). Crop row detection procedure using low-cost UAV imagery system. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 349–356.

Jiang, G., Wang, Z., & Liu, H. (2015). Automatic detection of crop rows based on multi-ROIs. *Expert Systems with Applications*, 42(5), 2429–2441.

Khadatkar, A., Mathur, S. M., Dubey, K., & BhusanaBabu, V. (2021). Development of embedded automatic transplanting system in seedling transplanters for precision agriculture. *Artificial Intelligence in Agriculture*, 5, 175-184. Doi: <https://doi.org/10.1016/j.aiia.2021.08.001>

Kannan, M., Ismail, I., & Bunawan, H. (2018). Maize Dwarf Mosaic Virus: From Genome to Disease Management. *Viruses*, 10(9). doi:10.3390/v10090492

Kurniawan, A. (2021). IMU Sensor: Accelerometer and Gyroscope. In: *Beginning Arduino Nano 33 IoT*. Apress, Berkeley, CA. doi:https://doi.org/10.1007/978-1-4842-6446-1_3

Liu, X., Qi, J., Zhang, W., Bao, Z., Wang, K., & Li, N. (2023). Recognition method of maize crop rows at the seedling stage based on MS-ERFNet model. *Computers and Electronics in Agriculture*, 211, 107964. <https://doi.org/10.1016/j.compag.2023.107964>

Mao, S., Li, Y., Ma, Y., Zhang, B., Zhou, J., & Wang, K. (2020). Automatic cucumber recognition algorithm for harvesting robots in the natural environment using deep learning and multi-feature fusion. *Computers and Electronics in Agriculture*, 170, 105254. Doi: <https://doi.org/10.1016/j.compag.2020.105254>

Monteiro, N., Alencar, E., Souza, N., & Leao, T. (2021). Ozonized water in the preconditioning of corn seeds: physiological quality and field performance. *Ozone Science and Engineering*, 43(5), 436-450. doi:10.1080/01919512.2020.1836472

Mun, D.B. (2019). Odometry error reduction in wheelchair using more than one sensor. University Tunku Abdul Rahman. <http://eprints.utar.edu.my/3457/1/ME-2019-1301582-1.pdf>

Nehme, H., Aubry, C., Solatges, T., Savatier, X., Rossi, R., & Boutteau, R. (2021). Lidar-based structure tracking for agricultural robots: Application to autonomous navigation in vineyards. *Journal of Intelligent & Robotic Systems*, 103, 1-16. Doi: <https://doi.org/10.1007/s10846-021-01519-7>

Orum, J., Wubale, T., Marcus, S., Harold, A., Veldhuisen, B., & Hildrands, H. (2023). Optimal use of agricultural robot in arable crop rotation: A case study from the Netherlands. *Smart Agricultural Technology*, 5, 1-8. doi:<https://doi.org/10.1016/j.atech.2023.100261>

Puente G, A. (2023). Secretaría de Agricultura y Desarrollo Rural outlook 2023: <https://www.cmdrs.gob.mx/sites/default/files/cmdrs/sesion/2023/01/13/4698/materiales/cmdrs-vf-puente.pdf>

Reyes Amador, A., & Velázquez López, N. (2020a). Modelo industrial de carrocería para vehículo de uso agrícola. MX 58211 B. Instituto Mexicano de la Propiedad Industrial.

Reyes Amador, A., & Velázquez López, N. (2020b). Sistema de suspensión para vehículos terrestres autónomos o no autónomos. MX 4369 B. Instituto Mexicano de la Propiedad Industrial.

Reyes Amador, A., & Velázquez López, N. (2021). Modelo industrial de chasis para vehículo de cuatro ruedas. MX 60106 B. Instituto Mexicano de la Propiedad Industrial.

Saavedra Sueldo, C., Perez Colo, I., De Paula, M., Villar, S. A., & Acosta, G. G. (2023). ROS-based architecture for fast digital twin development of smart manufacturing robotized systems. *Annals of Operations Research*, 322(1), 75-99. Doi: <https://doi.org/10.1007/s10479-022-04759-4>

Salmane, P.H., Rivera Velázquez, J.M., Khoudour, L., Mai, N.A.M., Duthon, P., Crouzil, A., Saint Pierre, G., & Velastin, S.A. (2023). 3D Object Detection for Self-Driving Cars Using Video and LiDAR: An Ablation Study. *Sensors*, 23, 3223. <https://doi.org/10.3390/s23063223>

- Solawetz, J. (2020). What is YOLOv5? A Guide for Beginners. Roboflow. <https://blog.roboflow.com/yolov5-improvements-and-evaluation/>
- Subeesh, A., & Mehta, C. R. (2021). Automation and digitization of agriculture using artificial intelligence and internet of things. *Artificial Intelligence in Agriculture*, 5, 278-291. Doi: <https://doi.org/10.1016/j.aiia.2021.11.004>
- Sung Sic, Y., Humayun, K., & Heung-Shik, L. (2022). Multiple IMU sensor fusion using resolution refinement method to reduce quantization error. *Journal of Mechanical Science and Technology*, 37(1), 163-168. doi:10.1007/s12206-022-1216-1
- Xie, B., Jin, Y., Faheem, M., Gao, W., Liu, J., Jiang, H., ... Li, Y. (2023). Research progress of autonomous navigation technology for multi-agricultural scenes. *Computers and Electronics in Agriculture*, 211, 107963.
- Xie, D., Chen, L., Liu, L., Chen, L., & Wang, H. (2022). Actuators and sensors for application in agricultural robots: A review. *Machines*, 10(10), 913. Doi: <https://doi.org/10.3390/machines10100913>
- Yang, S., Mei, S., & Zhang, Y. (2018). Detection of maize navigation centerline based on machine vision. *IFAC-PapersOnLine*, 51(17), 570–575. <https://doi.org/10.1016/j.ifacol.2018.08.140>.
- Zhan, J.; Maeta, S.; Bergerman, M.; Singh, S. (2014). Mapping Orchards for Autonomous Navigation. 2014 Montreal, Quebec Canada July 13 – July 16, 2014. <https://doi:10.13031/aim.20141838567>
- Zhang, X., Li, X., Zhang, B., et al. (2018). Automated robust crop-row detection in maize fields based on position clustering algorithm and shortest path method. *Computers and Electronics in Agriculture*, 154, 165–175.

5. APENDICES

5.1 Descripción de las tareas del Field Robot Event 2023

5.1.1 Tarea 1: Navegación

Descripción general

En esta tarea, los robots navegan de manera autónoma a través de un campo real de maíz. El giro debe seguir las filas adyacentes desde la pista 1 a la 5. Al salir de la pista 5, el robot debe seguir un patrón de giro particular dado. Esta tarea se trata de la precisión, la suavidad y la velocidad de la operación de navegación entre las filas. Dentro de tres minutos, el robot navega entre las filas. El objetivo es cubrir la mayor distancia posible recorrida. Se encuentra un ejemplo de campo y patrón de conducción en la figura 27.

Las primeras 3 pistas no tienen huecos entre filas para facilitar el inicio a los robots. El resto del campo (pista 4 a 11) tiene huecos entre filas, a veces incluso en ambos lados. En la última parte (después de la pista 5) el robot tiene que seguir un patrón de giro y fila en particular. El patrón puede ser: S - 1L - 1R - 3L - 2L - 2R - F.

Se colocan piedras y guijarros al azar a lo largo del camino. Por lo tanto, se requiere una altura de despeje del suelo de la máquina. Para facilitar el trabajo a los sensores, no habrá huecos en las entradas y salidas de las filas. Los extremos o comienzos de las filas pueden no estar en la misma línea. El perímetro del campo estará quizás indicado por una valla o zanja o similar.

Reglas para robots

Cada robot debe iniciar después de una indicación de inicio (señal acústica) en 1 min. El tiempo máximo disponible para la carrera es de 3 min.

Distribución de puntos

La distancia recorrida siguiendo el camino dado durante la duración de la tarea se mide. (Tan pronto como el robot abandone el camino especificado, se detendrá la medición de la distancia.) La distancia final se calculará incluyendo especialmente un factor de bonificación cuando se alcance el final del campo en

menos de 3 min. La distancia final, incluyendo un factor de bonificación, se calcula como:

$$S_{\text{final}} [\text{m}] = S_{\text{corrected}} [\text{m}] * 3 [\text{min}] / t_{\text{measured}} [\text{min}]$$

La distancia corregida incluye la distancia recorrida y los valores de penalización. La distancia recorrida, los valores de penalización y el tiempo de rendimiento son medidos por los oficiales del jurado.

El daño a las plantas de cultivo por el robot resultará en una penalización del 2% de la distancia total de la longitud de la fila en metros por planta dañada. (Ejemplo: 10 filas x 10 m = 100 m de distancia máxima, significa una penalización de 2 m por planta dañada.)

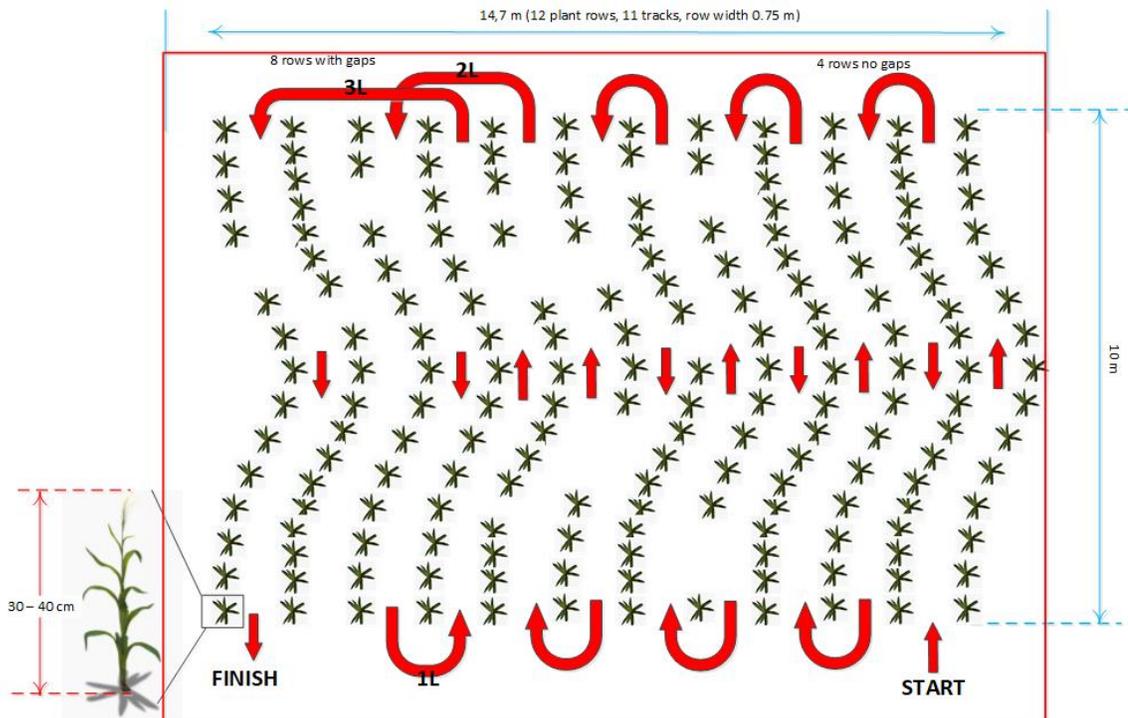


Figura 27. Concepto de estructura del campo para la tarea de navegación (ejemplo); Pista 1 a 3 sin huecos, pista 4 a 11 con huecos. Después de la pista 7 en navegación con patrón 2L (segundo a la izquierda), 1L (uno a la izquierda) y 3L (tercero a la izquierda) como ejemplo. Los perímetros tienen 2 m de ancho.

5.1.2 Tarea 2: Tratamiento (aspersión) de las plantas

Descripción general

Para esta tarea, los robots navegan de manera autónoma a través de un campo real de maíz, como en la primera parte de la tarea 1, pero omitiendo cada segunda fila. Además, los robots deben tratar las plantas cuando estén presentes y detener el tratamiento donde no haya plantas.

En condiciones óptimas, los robots deberían comenzar a asperjar las plantas a su izquierda y derecha. Si faltan plantas en un lado, el robot debe detener la aspersión en ese lado hasta que no detecte nuevas plantas. Además, el robot debe detener la aspersión cuando llegue al final de la fila y comience el procedimiento de giro. En el mundo real, los robots podrían quedarse sin agua o incluso no tener la capacidad de pulverizar. Por lo tanto, los robots deben estar equipados con dos indicadores / luces brillantes, que se encienden y apagan según la presencia de plantas. Para el concurso, el medio de aspersión será agua.

Reglas para robots

Cada robot debe iniciar después de una indicación de inicio (señal acústica) en 1 min. El tiempo máximo disponible para la carrera es de 3 min.

Distribución de puntos

Habrán 10 áreas en el campo donde faltarán plantas aproximadamente 1 m de filas vacías, distribuidas en uno o posiblemente ambos lados. Entonces, en total, faltarán 10 m de plantas. Los puntos se otorgarán según el número de regiones vacías detectadas con éxito (Smissing), con un peso establecido en 1, y la distancia total recorrida (Stravelled).

$$S_{\text{final}} [\text{m}] = 10 * \text{peso} * S_{\text{missing}} [\text{m}] + S_{\text{travelled}} [\text{m}]$$

Se otorgará un factor de bonificación a los robots que realmente asperjen las plantas. Para evaluar esto, se colocará papel sensible al agua (WSP) en tablas de madera secas y estas tablas de madera se colocarán en el suelo donde falten

las plantas(*1). Los WSP se colocarán uno cada 10 cm y, según el número de WSP mojados, el peso se establecerá de la siguiente manera:

The percentage of all WPS that are dry *2	Weight
80% or more	2

1 La tabla de madera con WPS se colocará cerca de la última planta. Si este primer WSP está mojado, se aplicarán los pesos. Si el robot indica el área con plantas faltantes, pero no logra pulverizar (ya que probablemente se quedó sin agua) y los WSP están todos secos, el peso se establecerá en 1.

2 Preste atención a la deriva de las filas anteriores, ya que podría arruinar los resultados.

60% to 79%	1.75
40% to 59%	1.5
Less than 40%	1.25
0 %, correct indication	1
0 %, false indication	0

where, if needed, the percentage will be rounded to the first integer value.

El daño a las plantas de cultivo por el robot resultará en una penalización del 2% de la distancia total de la longitud de la fila en metros por planta dañada. (Ejemplo: 10 filas x 10 m = 100 m de distancia máxima, significa una penalización de 2 m por planta dañada.)

En caso de que el robot no detecte las plantas en medio de la fila donde no faltan plantas, evidente ya que el robot dejará de pulverizar y / o el indicador de luz se apagará, esto resultará en puntos de penalización como en el caso de plantas dañadas (por ejemplo, una planta dañada por cada acción positiva falsa) 3*.

3* Algunos robots pueden ser bastante precisos para detectar dónde están las plantas y dónde solo hay (pequeñas porciones) de hojas presentes. Esto podría resultar en una situación constante de aspersión de encendido / apagado en

áreas donde están presentes las plantas. Para evitar puntos de penalización, se aconseja a los equipos mantener la aspersion / indicador encendido hasta la siguiente planta (si está presente).

5.1. 3 Tarea 3: detección y reconocimiento de posibles obstáculos

Descripción general

Para tener éxito en la siguiente tarea, se evaluará y probará qué tan buenos son los robots en reconocer posibles nuevos obstáculos. Por lo tanto, un conjunto de imágenes (una tras otra) se colocará frente al robot, cada una de ellas perteneciente a uno de los tres grupos: un ciervo, un humano y algo más. Los robots deben tener un indicador acústico y / o visual que permita al jurado / público saber qué ve el robot: un humano, un ciervo o desconocido.

Para hacer la competencia lo más justa posible, cada uno de los equipos proporcionará 3 imágenes antes del inicio de las tareas. Luego se hará una selección aleatoria donde una de estas 3 imágenes se usará en el conjunto final (así que si hay 15 equipos compitiendo, el conjunto constará de 15 imágenes). Todas las imágenes se imprimirán (cada una) en una hoja blanca de papel A3 y se colocarán en orden aleatorio frente al robot a una distancia de 1.5 m. Los robots tendrán entonces un marco de tiempo de 5 segundos para hacer una detección, reconocerla e indicar lo que ven.

Para esta tarea, los robots se colocarán al principio del campo entre las dos filas de plantas, pero no se moverán / conducirán durante esta tarea. En cambio, los robots se centrarán en lo que se coloca frente a ellos y harán una clasificación. Solo se puede hacer una clasificación por obstáculo y no se puede cambiar (solo cuenta la primera).

Reglas para robots

Cada robot debe iniciar la detección después de una indicación de inicio (señal acústica) en 1 min. El tiempo máximo disponible para la detección es de 5 segundos por obstáculo. Puede haber una ventana de hasta 10 segundos para cambiar las imágenes por el jurado (primero quitando la imagen anterior y

colocando la nueva). Una vez que se realiza una detección, no puede cambiar en los siguientes 5 segundos. Solo cuenta la primera detección.

Distribución de puntos

El jurado evalúa la detección y clasificación durante la carrera:

Detected object and right category (true positive)	5 points
Detected object wrong category (false positive)	-5 points

Imágenes de ejemplo para la tarea 3:

<https://fieldrobot.nl/event/index.php/2023/05/05/task-3-example-images/>



Figura 28. Imágenes de ejemplo para la tarea 3.

5.1.4 Tarea 4: Obstáculos estáticos y dinámicos

Descripción general

Esta tarea trata sobre la seguridad. Los robots conducirán a través del campo como en la tarea 1, pero sin omitir las filas. Además, tendrán que detectar obstáculos estáticos y dinámicos que puedan / se presentarán en su camino. Si el obstáculo es dinámico, el robot se detendrá y emitirá un sonido acústico y / o una indicación visual (por ejemplo, destellos brillantes) y el obstáculo se moverá en 5 segundos, lo que significa que el robot puede continuar conduciendo. Si el obstáculo es estático, no se moverá, y el robot debe retroceder y continuar en la siguiente fila.

Como en la tarea 3, los robots podrían encontrar un obstáculo, pero ahora mientras conducen. Para esta tarea habrá 5 obstáculos en el campo; 3 de ellos dinámicos y 2 estáticos. Para los obstáculos dinámicos, se colocará y quitará una imagen de un humano una vez que el robot dé un indicador claro (señal de audio

o visual). Si no se hace ninguna indicación, el obstáculo permanecerá en su camino y el robot tendrá que retroceder y continuar en la siguiente fila y esto contará como un falso positivo. Para obstáculos estáticos, se colocará una imagen de un ciervo en el camino del robot y no se quitará, donde el robot debe dar una indicación clara de lo que reconoció. Si el robot da una indicación incorrecta, se contará como un falso positivo.

A diferencia de la tarea anterior, las imágenes de los obstáculos se proporcionarán a los equipos antes del evento.

El daño a las plantas de cultivo por el robot resultará en una penalización del 2% de la distancia total de la longitud de la fila en metros por planta dañada. (Ejemplo: 10 filas x 10 m = 100 m de distancia máxima, significa una penalización de 2 m por planta dañada.)

Reglas para robots

Cada robot debe iniciar después de una indicación de inicio (señal acústica) en 1 min. El tiempo máximo disponible es de 5 min.

Distribución de puntos

El jurado evalúa la detección y clasificación durante la carrera:

Path travelled	x*0.5 points
Successful detection of a static / dynamic obstacle (true positive)	10 points
Unsuccessful detection of a static / dynamic obstacle (false positive)	-10 points

5.2 Imágenes de las tareas del FRE 2023.



Figura 29. Participantes del FRE 2023.



Figura 30. Terreno de pruebas de la competencia.



Figura 31. Zona de pruebas de navegación.



Figura 32. Tarea 2 aspersion.



Figura 33. Revisión de los jueces tarea 2.



Figura 34. Tarea 3 reconocimiento de imágenes.



Figura 35. Ejemplo de imágenes usadas en la tarea 3.



Figura 36. Tarea 4 navegación con obstáculos.



Figura 37. Campo del concurso durante tarea 4.

CERTIFICATE

OF PARTICIPATION

Field Robot Event, 12th to 15th of June 2023,
Pivola, Slovenia

WE HEREBY CONFIRM THAT

David Ivan Sánchez Chávez

has successfully attended the 20th International Field Robot
Event held in Slovenia at the Faculty of Agriculture and Life
Sciences, University of Maribor, from 12th to 15th of June 2023

Jurij Rakun, PhD
EVENT ORGANIZER



Prof. Branko Kramberger, PhD
DEAN

Figura 38. Certificado de participación FRE 2023.