

**Universidad Tecnológica Nacional**  
**Proyecto Final**

---

**Utilización de Inteligencia Artificial para la  
estimación de la actividad forrajera de las  
hormigas cortadoras de hojas.**

---

*Autores:*

- QUIROGA, Agustin Antonio
- VINZON STHELE, Eric Alan David

*Director:*

- MAGGIOLINI, Lucas

*Co-Director:*

- SABATTINI, Julian Alberto

*Proyecto final presentado para cumplimentar los requisitos académicos  
para acceder al título de Ingeniero electrónico*

*en la*

**Facultad Regional Paraná**

Julio de 2025

## **Declaración de autoría:**

Nosotros declaramos que el Proyecto Final “Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.” y el trabajo realizado son propios. Declaramos:

- Este trabajo fue realizado en su totalidad, o principalmente, para acceder al título de grado de Ingeniero electrónico, en la Universidad Tecnológica Nacional, Regional Paraná.
- Se establece claramente que el desarrollo realizado y el informe que lo acompaña no han sido previamente utilizados para acceder a otro título de grado o pre-grado.
- Siempre que se ha utilizado trabajo de otros autores, el mismo ha sido correctamente citado. El resto del trabajo es de autoría propia.
- Se ha indicado y agradecido correctamente a todos aquellos que han colaborado con el presente trabajo.
- Cuando el trabajo forma parte de un trabajo de mayores dimensiones donde han participado otras personas, se ha indicado claramente el alcance del trabajo realizado.

Firmas:

- *QUIROGA, Agustin Antonio*
- *VINZON STHELE, Eric Alan David*

Fecha: Julio de 2025

## Agradecimientos:

Agradecemos a nuestro director y co-director, Maggiolini Lucas y Sabattini Julián que nos aconsejaron y ayudaron durante la realización de nuestra tesis. A la Facultad Regional Paraná por darnos el espacio y todos los conocimientos necesarios para llegar a la entrega del proyecto y a nuestras familias, parejas y amigos que nos acompañaron durante todo el camino, empujando a nuestro lado sin descanso.

Quiroga, Agustín Antonio  
Vinzón Sthele, Eric Alan David

Universidad Tecnológica Nacional

*Abstract*

Facultad Regional Paraná

Ingeniero en Electrónica

**Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas**

QUIROGA, Agustin Antonio

VINZON STHELE, Eric Alan David

**Abstract**

An automated system was developed to achieve tracking and analysis of the foraging behavior of leaf-cutting ants. This allowed obtaining statistical information about the accumulated area of transported leaves, detection frequency, and trajectories followed by the ants.

For this purpose, convolutional neural networks (CNNs) implemented with YOLOv8 were used, which due to their optimized performance for detection, was selected over other methods. Tools such as Python, PyTorch, and OpenCV were employed. Additionally, image datasets were created from video recordings in a controlled environment, later expanded and annotated using Labelme and Roboflow, with the latter selected as the main framework.. Custom tracking algorithms were applied to overcome challenges related to load irregularity and erratic ant movements.

We implemented metrics such as Successful Tracking Rate and Trajectory Completeness Rate that measured our algorithm's score. In low-activity environments, we obtained 83.7 % and 95.3 %, respectively. In high-activity environments, the results were 52.3 % and 95.3 %. Finally, a graphical interface was implemented with CustomTKinter that facilitates interaction, parameters configuration and visualization of results.

## **Keywords**

YOLO, Detection, Ecology, Foraging, Segmentation, Tracking, Artificial Intelligence

## **Resumen:**

Se desarrolló un sistema automatizado con el fin de lograr el seguimiento y análisis del comportamiento forrajero de hormigas cortadoras de hojas. El objetivo central fue obtener información estadística sobre el área acumulada de las hojas transportadas, la frecuencia de detecciones y las trayectorias seguidas por las hormigas.

Para ello, se utilizaron redes neuronales convolucionales (CNN) implementadas con YOLOv8, que, por su rendimiento optimizado para la detección, se seleccionó por encima de otros métodos. Se emplearon herramientas como Python, PyTorch y OpenCV. Además se implementaron datasets de imágenes recopiladas a partir de grabaciones de video en un entorno controlado que luego fue ampliado y anotado con Labelme y Roboflow, siendo este último el framework seleccionado. Se aplicaron algoritmos personalizados de seguimiento para superar desafíos relacionados con la irregularidad de las cargas y los movimientos erráticos de las hormigas.

Implementamos métricas como la Tasa de Seguimiento Exitoso y Tasa de Completitud de Trayectoria que midieron la puntuación de nuestro algoritmo. En entornos de baja actividad, obtuvimos 83,7% y 95,3%, respectivamente. En entornos de alta actividad, los resultados fueron 52,3 % y 95,3 %. Finalmente, se puso en práctica una interfaz gráfica con CustomTKinter que facilita la interacción, configuración de parámetros y visualización de los resultados.

## **Palabras Clave:**

YOLO, Detección, Ecología, Forrajeo, Segmentación, Seguimiento, Inteligencia Artificial

# Índice

<b>Abstract .....</b>	4
<b>Keywords .....</b>	5
<b>Resumen:.....</b>	6
<b>Palabras Clave:.....</b>	6
<b>Capítulo 1: Introducción .....</b>	11
1.2 Marco teórico.....	13
1.3 Objetivos .....	14
<b>Capítulo 2: Desarrollo teórico.....</b>	15
2.1 Imagen .....	15
2.2 Procesamiento digital de imágenes .....	16
2.3 Transformaciones de intensidad y filtrado espacial.....	17
2.4 Transformaciones de intensidad y filtrado espacial: conceptos elementales.....	18
2.5 Métodos en el dominio espacial.....	19
2.6 Kernel.....	19
2.7 Procesamiento de imágenes a color.....	21
2.7.1 Fundamentos del color.....	21
2.7.2 Convirtiendo Colores de RGB a HSI .....	26
2.7.3 Segmentación de color.....	26
2.8 Imágenes binarias .....	28
2.8.1 Máscara .....	28
2.8.2 Segmentación.....	29
2.9 Visión artificial .....	31
2.9.1 Algoritmo de aprendizaje.....	32
2.9.2 Tensor.....	33
2.9.3 Redes neuronales.....	34
2.9.4 Deep Learning .....	36
2.9.5 Redes neuronales convolucionales.....	37
2.9.6 Componentes de las Redes Neuronales Convolucionales. ....	38
2.10 Seguimiento de objetos .....	39
<b>Capítulo 3: Materiales y métodos.....</b>	41
3.1 Materiales.....	41
3.1.1 PyTorch .....	41
3.1.2 OpenCV .....	42

3.1.3 You Only Look Once (YOLO).....	43
3.1.4 Labelme .....	44
3.1.5 Roboflow .....	44
3.1.6 Anaconda.....	44
3.1.7 TKinter (Tk/Tcl) .....	44
3.1.8 CustomTKinter .....	45
3.2 Métodos .....	45
3.2.1 Investigación .....	45
3.2.2 Recopilación de imágenes .....	46
3.2.3 Anotación del dataset.....	48
3.2.4 Entrenamiento.....	50
3.2.5 Entrenamiento del modelo propio.....	54
3.2.6 Evaluación del modelo .....	57
3.2.7 Seguimiento multi-objeto.....	64
3.2.8 Cálculo del área .....	77
3.2.9 Interfaz gráfica .....	80
3.2.10 Etapa de validación.....	87
Capítulo 4: Límites de implementación .....	91
4.1 Limitaciones del dominio de aplicación.....	91
4.2 Desafíos en la consistencia de identificadores .....	92
4.3 Criterios de validación para análisis estadístico.....	92
Capítulo 5: Resultados .....	94
5.1 Entorno de menor actividad .....	95
5.2 Entorno de mayor actividad .....	100
5.3 Análisis comparativo entre entornos .....	105
5.4 Resultados en vídeos de larga duración.....	106
Capítulo 6: Mejoras Futuras y Líneas de Investigación .....	108
6.1 Implementación en Tiempo Real .....	108
6.2 Robustez ante Oclusiones Complejas .....	108
6.3 Adaptación a Entornos Naturales .....	109
Capítulo 7: Discusión .....	110
Capítulo 8: Conclusión .....	112
Capítulo 9: Bibliografía.....	113

## Índice de figuras

Figura 1 - Aplicación del kernel	20
Figura 2 - Ejemplo de aplicación de un kernel. (a) Imagen original. (b) valores del kernel (c) imagen de salida.	21
Figura 3 - Rango visible del espectro electromagnético.	22
Figura 4 - Diagrama de cromaticidad del CIE	25
Figura 5 - Segmentación de color que detecta colores verdes. a) Imagen original. b) Imagen recortada.	28
Figura 6 - Ejemplo de aplicación de una máscara a una imagen a color. (a) Imagen original. (b) imagen producida por un umbral del canal verde de (a). (c) máscara binaria de (b)	29
Figura 7 - Ejemplo de aplicación de una máscara a una imagen a color. (a) Imagen original. (b) imagen binaria producida por el modelo de segmentación de (a).	31
Figura 8 - Diagrama del funcionamiento de un algoritmo de aprendizaje básico.	33
Figura 9 - Esquema de una red neuronal multicapa. Xn las entradas, Zn las salidas y las capas intermedias son las capas ocultas (yn).	35
Figura 10 - Ilustración de un modelo de aprendizaje profundo.	37
Figura 11 - Diagrama de flujo básico de una red neuronal convolucional. (Pro Deep Learning with TensorFlow p. 179)	39
Figura 12 - Representación gráfica de una convolución entre un tensor y un filtro.	42
Figura 13 - División de celdas realizada por YOLO	43
Figura 14 - Imagen original e imagen escalada a 640 x 640	47
Figura 15 - Anotación segmentada de la carga en la herramienta Labelme	48
Figura 16 - Anotación segmentada de la carga en la herramienta Roboflow.	49
Figura 17 - Comparación de los parámetros de YOLOv8 con sus versiones anteriores. a) Curva de parámetros-mAP(Precisión mediana). b) Curva de velocidad(ms/img)-mAP.	53
Figura 18 - Métricas del modelo en detección y segmentación entrenado en 100 épocas.	56
Figura 19 - Intersección sobre unión	60
Figura 20 - Matriz de confusión del Conjunto base adicional	61
Figura 21 - Gráfica de comparación de métricas de los modelos entrenados.	64
Figura 22 - Visualización de una detección y comparación positiva.	68
Figura 23 - Diagrama de flujo del algoritmo de seguimiento.	72
Figura 24 - Diagrama de flujo del algoritmo de filtrado.	76
Figura 25 - Interfaz gráfica realizada con CustomTKinter (Sin vídeo activa)	81
Figura 26 - Interfaz gráfica realizada con CustomTKinter (Con vídeo activa)	82
Figura 27 - Secuencia de botones.	83
Figura 28 - Pestaña de configuración de parámetros.	84
Figura 29 - Entorno del video utilizado	96
Figura 30 - Comparación de cantidad de hojas por intervalo	97
Figura 31 - Gráfico de la distribución por intervalo en parámetros permisivos.	98
Figura 32 - Gráfico de la distribución por intervalo en parámetros semi-permisivos	99
Figura 33 - Gráfico de la distribución por intervalo en parámetros estrictos.	100
Figura 34 - Entorno del segundo vídeo utilizado	101
Figura 35 - Comparación de cantidad de hojas por intervalo	102
Figura 36 - Gráfico de la distribución por intervalo en parámetros permisivos.	103
Figura 37 - Gráfico de la distribución por intervalo en parámetros semi-permisivos	104
Figura 38 - Gráfico de la distribución por intervalo en parámetros estrictos.	105
Figura 39 - Cantidad de hojas por intervalo.	106
Figura 40 - Área total por intervalo.	107

## **Índice de tablas**

Tabla 1 – Métricas de los modelos pre-entrenados de YOLO .....	52
Tabla 2 - Parámetros de entrenamiento.....	55
Tabla 3 - Tabla comparativa con métricas de los modelos entrenados. ....	63
Tabla 4 - Comparación de tiempos de inferencia. ....	64
Tabla 5 - Comprobación de comportamiento ante el cambio de parámetros.....	69
Tabla 6 - Definición de parámetros de prueba.....	95
Tabla 7 - Resultados parámetros permisivos. ....	98
Tabla 8 - Resultados parámetros semi-permisivos.....	98
Tabla 9 - Resultados parámetros estrictos .....	99
Tabla 10 - Resultados parámetros permisivos. ....	103
Tabla 11 - Resultados parámetros semi-permisivos.....	103
Tabla 12 - Resultados parámetros estrictos. ....	104

## **Dedicado a:**

Mis papás Mario y Laura, mis hermanos Joaquín y Manuel y a mi novia Elena, que siempre estuvieron ahí para apoyarme en los momentos donde todo se hacía cuesta arriba.

Quiroga, Agustín

Mis padres, a mis hermanos, a mi pareja y a colegas varios, quienes nos brindaron su apoyo y motivación incondicional, desde el conocimiento y la posición que cada uno ocupó, impulsandonos a lo largo de todo este proceso.

Vinzón Sthele, Eric Alan David

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	1 – Introducción

## Capítulo 1: Introducción

La sostenibilidad de los sistemas agrícolas modernos no solo depende del uso eficiente de recursos productivos, sino también del conocimiento profundo de las interacciones ecológicas que se desarrollan en el entorno natural donde se insertan los cultivos. En particular, los agroecosistemas son espacios complejos donde conviven especies vegetales, microorganismos, fauna silvestre y especies domesticadas, generando dinámicas que pueden ser beneficiosas o perjudiciales para la productividad agrícola. Dentro de estas interacciones, los insectos juegan un papel central, ya sea como agentes polinizadores, controladores biológicos o como plagas que pueden reducir significativamente el rendimiento económico del productor.

Entre las especies de insectos que generan impactos relevantes en el entorno agrícola, las hormigas cortadoras de hojas (HCH) —particularmente aquellas pertenecientes a los géneros *Atta* y *Acromyrmex*— constituyen uno de los principales desafíos para el manejo de cultivos en vastas regiones de América Latina [1]. Son consideradas plagas agrícolas debido a su comportamiento social altamente organizado y su notable capacidad para recolectar grandes volúmenes de material vegetal fresco. Dicho material no es consumido directamente, sino que es transportado hacia sus nidos subterráneos para cultivar un hongo específico (principalmente del género *Leucoagaricus*), del cual se alimentan simbóticamente [2] [3]. A diferencia de otras plagas que pueden afectar únicamente ciertas partes del ciclo de vida del cultivo o especies vegetales específicas, las HCH pueden atacar una amplia variedad de plantas, tanto cultivadas como silvestres, y durante diferentes etapas de su crecimiento, lo que las convierte en una amenaza persistente y económicamente dañina [1]. Además de su rol como plagas, es importante reconocer que las HCH también son ingenieras del ecosistema, modificando la estructura del suelo y los ciclos de nutrientes [2].

Cuantificar de manera precisa y eficiente la actividad forrajera de las HCH en los agroecosistemas ha resultado tradicionalmente complejo. Las metodologías convencionales para estimar el material transportado o la intensidad del forrajeo se

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	1 – Introducción

basan en observaciones manuales directas, muestreos destructivos en campo y estimaciones visuales del área defoliada o del tamaño de las cargas transportadas por las obreras, principalmente individuos de castas medias y grandes, que se especializan en las tareas de corte y transporte de fragmentos de hojas [2]. Estos métodos, si bien útiles, a menudo implican un margen considerable de error, son intensivos en mano de obra, difíciles de estandarizar y presentan una baja replicabilidad en distintos contextos o a lo largo del tiempo.

En función de lo anterior se propuso una solución innovadora que aborda estas limitaciones mediante el diseño y desarrollo de un software especializado. Este sistema utiliza herramientas de inteligencia artificial (IA), específicamente modelos de aprendizaje profundo (deep learning) dentro del campo del machine learning (ML), para el procesamiento y análisis avanzado de imágenes digitales. El objetivo central es estimar automáticamente el área foliar de los fragmentos de hojas transportadas por las hormigas a partir de registros visuales a través de videos. Esta propuesta busca generar una base cuantitativa sólida y automatizada para comprender mejor el comportamiento forrajero de estas especies, su impacto real en términos de biomasa removida.

Tomando como antecedente el trabajo de tesis de Sturniolo, Francisco Daniel [7] sobre el monitoreo y la conducta general de las HCH, nuestra propuesta actual consiste en una herramienta diseñada para mejorar la estimación de la actividad de forrajera. El foco principal de nuestro proyecto es la cuantificación del área de la carga acarreada por estos insectos, que permite obtener datos estadísticos detallados sobre el área acumulada de las hojas transportadas.

El desarrollo de este software se pudo dar con ayuda de una exhaustiva investigación del marco teórico. El aprendizaje profundo, y en particular las Redes Neuronales Convolucionales (CNNs), son esenciales para resolver esta problemática. Las CNNs son arquitecturas de redes neuronales inspiradas en el córtex visual de los primates, diseñadas explícitamente para procesar datos con una topología de rejilla, como las imágenes [4]. Una CNN típica se compone de una secuencia de capas: a) convolucionales que aplican un conjunto de filtros aprendibles para extraer

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	1 – Introducción

características jerárquicas; b) agrupación o también llamada pooling; y c) capas completamente conectadas al final de la red, que utilizan las características de alto nivel extraídas para realizar tareas de clasificación o regresión (por ejemplo, identificar una hormiga y segmentar la hoja que transporta) [9]. El "aprendizaje" en estos modelos se logra mediante la optimización de los parámetros de los filtros y las conexiones (pesos) a través de un algoritmo llamado retropropagación (backpropagation), utilizando grandes conjuntos de datos de imágenes previamente etiquetadas. Este proceso permite al modelo aprender patrones complejos directamente de los datos visuales, mejorando su capacidad para realizar tareas específicas con alta precisión [4] [9].

Finalmente, es imperativo asegurar que el uso de estas tecnologías no cause daños ni altere el comportamiento natural de las hormigas. Para ello se implementaron medidas para minimizar la interacción directa con las colonias, utilizando cámaras para capturar imágenes de forma no invasiva. Estas consideraciones éticas son cruciales y se alinean con estudios recientes sobre la aplicación responsable de la IA en la investigación biológica y la conservación [13].

## 1.2 Marco teórico

El procesamiento y análisis digital de imágenes es un área de la ingeniería que se encarga de manipular y analizar la información contenida en una imagen digital. El procesamiento digital de imágenes consiste en aplicar técnicas a las imágenes digitales con el objetivo de mejorar su calidad. Sin embargo, el análisis de imágenes se enfoca en facilitar la búsqueda e interpretación de información contenida en esas imágenes.

Dentro del campo del procesamiento de imágenes se encuentran diversas técnicas que involucran operaciones cuyo origen es una imagen y cuyo resultado final es una imagen procesada. El valor de cada píxel en la imagen de salida puede depender de: a) valor de la imagen de entrada, b) valores de sus vecinos o, c) de todos los puntos de la imagen de entrada. Por otro lado, el análisis de imágenes toma

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	1 – Introducción

como entrada una imagen generando una de salida, que en la mayoría de los casos se relaciona estadísticamente con la de entrada.

Para el desarrollo de este proyecto, las redes neuronales fueron una pieza fundamental, por ser una herramienta poderosa en el procesamiento y análisis digital de imágenes. Las mismas pueden aprender a reconocer patrones y características en las imágenes, mejorar la calidad de las mismas, realizar tareas de restauración, comprimir las y segmentarlas en diferentes regiones. Además, estas redes también pueden extraer información estadística y facilitar la interpretación de los datos contenidos en las imágenes.

### 1.3 Objetivos

El desarrollo de una herramienta para obtener información útil sobre el comportamiento forrajero de las hormigas cortadoras de hojas, dando como resultado datos estadísticos sobre: a) el área acumulada de las hojas transportadas, b) la frecuencia de detecciones y c) las trayectorias seguidas por las hormigas para estudios ecológicos.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

## Capítulo 2: Desarrollo teórico

En este capítulo se expondrán aquellos conceptos que están intrínsecamente relacionados con el procesamiento digital de imágenes. Es importante denotar los diferentes tipos de técnicas y análisis que pueden ser aplicados a una imagen para lograr los resultados deseados. También abordaremos la visión artificial en conjunto con las técnicas de detección y segmentación de objetos.

### 2.1 Imagen

Una imagen puede ser definida como “la reproducción de un objeto formada por la convergencia de los rayos luminosos que, procedentes de él, atraviesan una lente o aparato óptico, y que puede ser proyectada en una pantalla” [112]. De este modo, podemos decir que las imágenes son una representación de la realidad, y en nuestro caso, tomaremos como que estas representaciones pueden ser percibidas visualmente por las personas.

Desde una perspectiva matemática rigurosa, una imagen puede definirse como una función bidimensional  $f(x,y)$ , donde  $x$  e  $y$  representan las coordenadas espaciales de un plano, y la amplitud de  $f$ , en cualquier par de coordenadas  $(x,y)$  corresponde a la intensidad o nivel de gris de la imagen en ese punto específico. Esta representación matemática constituye la base conceptual fundamental para todo el procesamiento digital de imágenes moderno [8].

Cuando  $x$ ,  $y$  y los valores de intensidad de  $f$  son cantidades finitas y discretas, la función resultante se denomina imagen digital. Esta definición matemática, ampliamente establecida en la literatura especializada por González y Woods, representa el punto de partida para todas las operaciones de procesamiento digital subsecuentes [8].

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

## 2.2 Procesamiento digital de imágenes

El campo del procesamiento digital de imágenes se define fundamentalmente como el estudio de la manipulación y análisis de imágenes digitales a través de un procesador. Esta disciplina, como señalan González y Woods, "refiere a aquel estudio que se realiza a diferentes tipos de imágenes por medio de un procesador. Cada componente de estas imágenes digitales posee una naturaleza discreta, con un lugar y valor determinado" [8], lo que implica que cada punto en la imagen tiene coordenadas y una intensidad específica. A estos componentes individuales se les conoce por varios nombres, incluyendo "elementos de imagen, pels y pixeles, siendo esta última la más utilizada" [8]. En esencia, el procesamiento digital de imágenes toma una imagen de entrada, compuesta por estos píxeles discretos, y produce una versión transformada o modificada de la misma. Esta transformación se basa en el hecho de que "una imagen digital es una imagen  $f(x, y)$  que se ha discretizado tanto en coordenadas espaciales como en intensidad" [8], lo que permite su representación y manipulación mediante estructuras matemáticas. De hecho, "una imagen digital puede ser considerada como una matriz cuyos índices de fila y columna identifican un punto en la imagen y el valor correspondiente del elemento de la matriz" [8], facilitando el uso de herramientas del álgebra lineal para el análisis y procesamiento. Su representación matemática sería

$$f(x, y) = \begin{bmatrix} f(0, 0) & f(0, 1) & \dots & f(0, N-1) \\ f(1, 0) & f(1, 1) & \dots & f(1, N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1, 0) & f(M-1, 1) & \dots & f(M-1, N-1) \end{bmatrix}$$

En el contexto del procesamiento digital de imágenes, la representación matricial de una imagen es fundamental. Podemos observar que  $M \times N$  representa el tamaño de la imagen, donde  $M$  indica el número de filas (altura) y  $N$  el número de columnas (ancho) de la matriz, o de la imagen. Por convención, "el origen de una imagen se encuentra en la esquina superior izquierda, es decir en  $f(0,0)$ , esta convención tiene sus raíces en la forma en que muchas pantallas de visualización,

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

como los monitores de TV construyen la imagen, comenzando desde la esquina superior izquierda y barriendo hacia la derecha, fila por fila" [8]. Es a esta representación matricial de la imagen a la que se aplican diversos tipos de procesamiento. Por lo tanto, es crucial considerar el nivel de intensidad de cada pixel de la imagen. Inicialmente, este nivel de intensidad se representa mediante escalas de grises, denotadas por L, donde el valor de L estará dado por

$$L = 2^k$$

Donde "k representa el número de bits que se necesitan para representar un valor de L determinado. En la práctica, se asume que los niveles discretos están igualmente espaciados y son integradores en el rango de [0, L-1]" [8]. Este rango incluye L niveles de gris, desde el negro (0) hasta el blanco (L-1).

El número total de bits necesarios para almacenar una imagen digital de tamaño M x N, donde cada píxel tiene un nivel de gris representado por k bits (lo que permite L = 2<sup>k</sup> niveles de gris), se calcula como:

$$b = M \times N \times k$$

### 2.3 Transformaciones de intensidad y filtrado espacial

En el ámbito del procesamiento digital de imágenes, el término dominio espacial se refiere al plano de la imagen y, por extensión, a los métodos de procesamiento que operan directamente sobre la disposición de los píxeles. A diferencia de otras técnicas que pueden involucrar una transformación de la imagen al dominio de la frecuencia, obteniendo la transformada inversa para devolver el resultado al dominio espacial, el procesamiento en el dominio espacial opera inherentemente sobre la representación espacial de la imagen [8].

Dentro de este dominio, las principales categorías del procesamiento en el dominio espacial son la transformación de intensidad y el filtrado espacial [8]. Es crucial entender la distinción entre ellas: "la transformación de intensidad opera solamente en un pixel, ajustando su valor independientemente de sus vecinos (aunque su posición pueda ser un factor en algunas transformaciones), mientras que

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

el filtrado espacial realiza operaciones en los vecinos de cada pixel en una imagen" [8], considerando el contexto local para determinar el nuevo valor del pixel central.

Estos métodos de procesamiento en el dominio espacial ofrecen una notable flexibilidad y versatilidad al interactuar directamente con los píxeles de una imagen. Permiten, así, realizar ajustes precisos en los niveles de intensidad para realzar contrastes o modificar el rango dinámico, realzar características específicas, eliminar ruido y mejorar la calidad visual en general [8] mediante operaciones de suavizado y mejora con el fin de facilitar la interpretación de la imagen. Adicionalmente, la capacidad de aplicar operaciones locales que consideran la información contextual de los píxeles vecinos resulta beneficioso para tareas como la detección de bordes, la segmentación [8] y otros análisis detallados de la estructura de la imagen, ya que se evalúan las relaciones espaciales entre elementos.

## 2.4 Transformaciones de intensidad y filtrado espacial: conceptos elementales.

El procesamiento en el dominio espacial se describe de manera fundamental a través de la siguiente expresión: $g(x,y) = T[f(x,y)]$

Donde "  $f(x,y)$  es una imagen de entrada,  $g(x,y)$  es la imagen de salida, y  $T$  es un operador sobre  $f$  definido sobre una vecindad del punto  $(x,y)$  " [8]. Es relevante destacar que la vecindad más elemental considerada es de tamaño  $1 \times 1$ . Cuando la vecindad es de este tamaño mínimo, la operación se realiza únicamente sobre el pixel en la ubicación  $(x,y)$ , y  $g$  dependerá solamente del valor de  $f$  en un solo punto  $(x,y)$ . En tal escenario, el operador  $T$  se vuelve un operador de intensidad, y la función de transformación toma la forma

$$s = T(r)$$

donde  $r$  representa la intensidad del pixel en la imagen original y  $s$  es la intensidad transformada en la imagen de salida. [8]

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

## 2.5 Métodos en el dominio espacial

Las transformaciones de intensidad, también conocidas como operaciones puntuales, representan una de las técnicas más fundamentales y conceptualmente sencillas en el procesamiento digital de imágenes [8], siendo consideradas el tipo más simple de operación de procesamiento de imágenes [115]. La característica distintiva de estas operaciones reside en que modifican el valor de intensidad de un píxel de la imagen de salida basándose solamente en el valor de intensidad del píxel correspondiente en la imagen de entrada, independientemente de los valores de los píxeles vecinos. Esta relación se describe mediante una función de transformación, donde cada nivel de gris de entrada se mapea a un nuevo nivel de gris de salida [8] [115].

## 2.6 Kernel

Un componente esencial en el proceso de filtrado espacial es el kernel. Gonzalez y Woods lo definen como una “matriz bidimensional”. Este elemento es conocido por varios sinónimos en el campo, incluyendo máscara, matriz de convolución o filtro [8]. El kernel actúa como una plantilla que se utiliza para aplicar una operación específica de vecindario a cada píxel de la imagen original.

El mecanismo de operación del kernel en el filtrado espacial implica posicionar su centro coincidente con el píxel actual que se está procesando en la imagen de entrada. Una vez posicionado, se realiza una operación matemática entre los elementos del kernel y los píxeles de la imagen que se encuentran bajo él. Típicamente, esta operación es una suma ponderada, donde los valores del filtro, también llamados coeficientes del kernel, se multiplican por los valores de píxel correspondientes con los que el filtro está superpuesto, y estos productos se suman para determinar el valor del píxel correspondiente en la imagen de salida [8], tal como se puede observar en la siguiente figura.

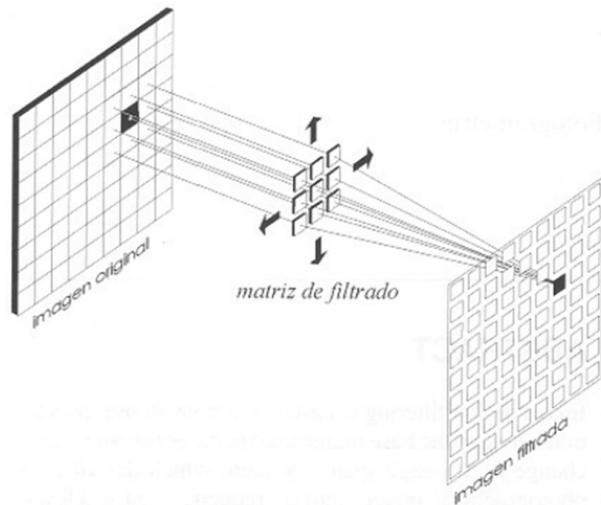


Figura 1 - Aplicación del kernel [118].

La forma y los valores del kernel determinan cómo se realizará la operación de procesamiento en cada píxel. Esto significa que los coeficientes numéricos dentro de la matriz del kernel definen la naturaleza exacta del filtrado que se aplicará a la imagen. Generalmente, el kernel tiene un tamaño pequeño, con el píxel central siendo el píxel que se está procesando, siendo comunes tamaños de 3x3 o 5x5. La simplicidad y eficiencia de esta operación la convierten en una herramienta poderosa en el dominio espacial [8].

En la figura 2 se puede observar un ejemplo de la aplicación del Kernel sobre una imagen obtenida.

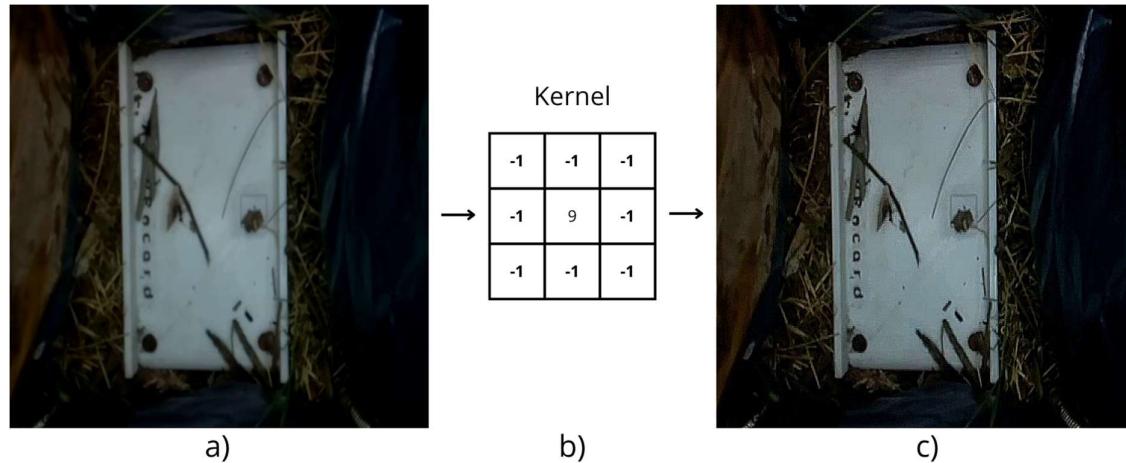


Figura 2 - Ejemplo de aplicación de un kernel. (a) Imagen original. (b) valores del kernel (c) imagen de salida.

## 2.7 Procesamiento de imágenes a color

El empleo del color en el procesamiento de imágenes es un aspecto fundamental que responde a factores significativos. En primer lugar, el color se erige como un “descriptor potente que a menudo simplifica la identificación y extracción de objetos” [8] en una escena. Su rica información contribuye a una representación más completa y detallada de las características presentes en una imagen, lo cual es esencial para diversas tareas de análisis y visión por computadora. En segundo lugar, la capacidad del sistema visual humano para distinguir miles de tonos de color supera con creces nuestra habilidad para discernir solamente alrededor de dos docenas de tonos de gris. Esta superior discriminación cromática permite una interpretación y apreciación más precisa de las imágenes por parte de los observadores humanos. [8]

### 2.7.1 Fundamentos del color

La percepción del color en los objetos, tanto para los seres humanos como para algunos otros animales, está intrínsecamente ligada a la luz reflejada por dicho objeto. Tal como se ilustra típicamente en las figuras que representan el espectro electromagnético, la luz visible para el ojo humano constituye una banda

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

relativamente estrecha de frecuencias en el espectro electromagnético [8], como podemos observar en la figura 3. Dentro de este rango visible, diferentes longitudes de onda corresponden a diferentes colores.

Cuando un objeto refleja luz que está balanceada en todas las longitudes de onda visibles, este objeto aparece en blanco para el observador. Por el contrario, si un objeto favorece la reflexión en un rango limitado del espectro visible, al absorber predominantemente otras longitudes de onda y reflejar solo ciertas frecuencias, exhibe ciertas tonalidades de color [8] que son percibidas por el observador. La forma en que los objetos interactúan con la luz (absorbiendo, reflejando o transmitiendo selectivamente ciertas longitudes de onda) es, por lo tanto, el principio subyacente a la manifestación del color.

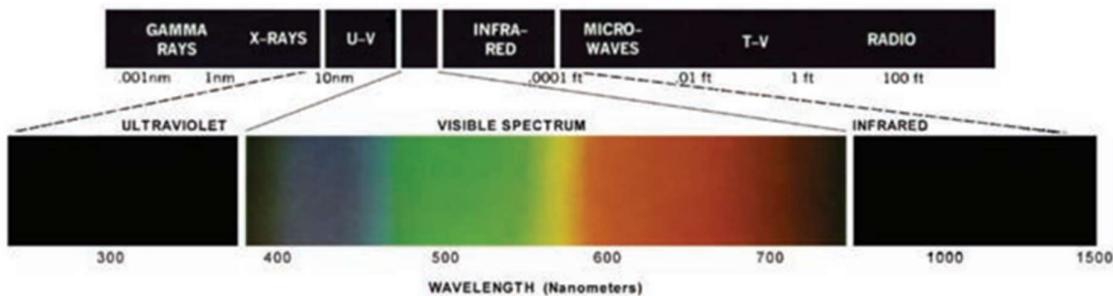


Figura 3 - Rango visible del espectro electromagnético [8].

En el contexto de la captura y representación de imágenes digitales, es una práctica común que los dispositivos detecten el color mediante tres componentes primarios: rojo (R), verde (G) y azul (B). Estos tres canales forman el espacio de color **RGB**, que es un modelo conveniente para la captura de color y su representación y despliegue en dispositivos de hardware, como monitores y cámaras [8]. Sin embargo, a pesar de que los valores de los píxeles se almacenan y manipulan inicialmente en el espacio RGB en los sistemas computacionales, estamos interesados en utilizar las

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

propiedades de los modelos de color para facilitar tareas de procesamiento de imágenes, como la detección de regiones [8].

Aunque el espacio de color RGB posee una estructura matemática cartesiana, su principal limitación para ciertas tareas de procesamiento radica en que “no describe el color en términos de la forma en que los seres humanos lo perciben y describen” [8]. Esto significa que la variación en cada uno de los componentes R, G o B no se corresponde directamente con atributos perceptivos intuitivos como la tonalidad o la pureza del color. Como resultado, realizar procesamientos basados en características perceptuales puede ser problemático en el espacio RGB [8], ya que los ajustes directos sobre los componentes R, G o B pueden no generar los cambios percibidos esperados en el color.

A partir de esto, se identifican características perceptualmente más relevantes para distinguir los colores que las componentes de color primarias. las características usadas para distinguir un color particular son el brillo, matiz (hue) y saturación [8].

- **Brillo:** Este atributo incorpora la noción de intensidad. Se refiere a la luminancia o cuán claro u oscuro aparece un color [8].
- **Matiz (Hue):** El matiz es un atributo asociado con la longitud de onda dominante en una mezcla de longitudes de onda de luz. Así, matiz es el color dominante según lo percibido por un observador. Es lo que comúnmente identificamos como el "tipo" de color (rojo, verde, azul, amarillo, etc.) [8].
- **Saturación:** La saturación se refiere a la pureza de un color. Indica cuán diluida está una luz de color puro por luz blanca. La saturación es inversamente proporcional a la cantidad de luz blanca mezclada con un matiz. Un color de alta saturación es vívido, mientras que uno de baja saturación tiende a ser grisáceo o pastel [8].

Conjuntamente, matiz y saturación, tomados en conjunto, se llaman crominancia. Es por esto por lo que, un color puede ser caracterizado de manera más intuitiva y

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

perceptual por su crominancia y brillo, separando la información de color (tono y pureza) de la información de intensidad. La descripción precisa de un color también puede involucrar sus **coeficientes tricromáticos**, que se definen en el contexto de la estandarización del color, como en los diagramas de cromaticidad [8].

$$x = \frac{X}{X + Y + Z}$$

$$y = \frac{Y}{X + Y + Z}$$

$$z = \frac{Z}{X + Y + Z}$$

Entonces,

$$x + y + z = 1$$

Otra forma de determinar el color es usando el diagrama de cromaticidad, impuesto por el CIE (Commission Internationale de l'Eclairage— Comisión Internacional de la Iluminación), que muestra la composición como una función de  $x$  (rojo) e  $y$  (verde). Para cualquier valor de  $x$  e  $y$ , el valor correspondiente de  $z$ (azul) es obtenido de  $z = 1 - (x + y)$ .[8]

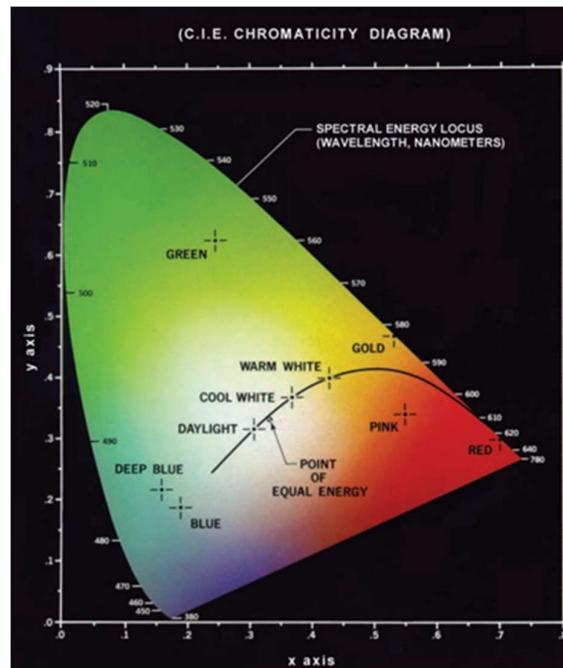


Figura 4 - Diagrama de cromaticidad del CIE [8].

Los dispositivos de medición de color emplean los ejes CIE para determinar la longitud de onda dominante y la saturación de cualquier color. Cuando se mezclan dos colores, se genera un nuevo punto en el diagrama, ubicado a lo largo de una línea recta que conecta los dos colores originales.[8]

El modelo de color HSI, anteriormente descripto, es una herramienta muy útil al momento de desarrollar algoritmos de procesamiento de imágenes basados en descripciones de color que son naturales e intuitivas para los seres humanos, quienes, después de todo, son los desarrolladores y usuarios de estos algoritmos.

Esta versatilidad del modelo HSI, en el procesamiento digital de imágenes, es un aspecto clave para obtener resultados más cercanos a la percepción humana, lo que mejora la calidad y comprensión de las imágenes procesadas.

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

## 2.7.2 Convirtiendo Colores de RGB a HSI

Dada una imagen en formato de color RGB, la componente H de cada pixel RGB es obtenida usando la ecuación [8]

$$H = \begin{cases} \theta, & \text{si } B \leq G \\ 360 - \theta, & \text{si } B > 0 \end{cases}$$

Con

$$\theta = \left\{ \frac{\frac{1}{2}[(R - G) + (R - B)]}{[(R - G) + (R - B)(G - B)]^{\frac{1}{2}}} \right\}$$

El componente de saturación será

$$S = 1 - \frac{3}{(R + G + B)} [(R, G, B)]$$

Finalmente, el componente de intensidad estará dado por

$$I = \frac{1}{3} (R + G + B)$$

## 2.7.3 Segmentación de color

El remarcado de un rango específico de colores en una imagen es una técnica muy útil en el procesamiento de imágenes a color, particularmente muy útil al momento de separar objetos de sus alrededores [8]. Este método, conocido como rebanado de color (color slicing) o umbralización de color, tiene como “idea básica uno de dos propósitos: (1) mostrar el color de interés para así hacerlo destacar del fondo o (2) usar la región definida por el color como una máscara para un procesamiento más profundo” [8].

Una de las formas más simples de segmentar una imagen a color consiste en asignar colores fuera de un rango de interés a un color neutro y poco destacado. Para implementar esta técnica, si el rango de color de interés se define como los colores que están encerrados por un cubo en un espacio de color apropiado, con este cubo

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

centrado en un color prototípico que tiene componentes ( $a_1, a_2, \dots, a_n$ ) y un ancho W [8], las transformaciones necesarias para lograr el remarcado se definen de la siguiente manera:

$$s_i = \begin{cases} 0,5, & \text{&si } \left[ |r_j - a_j| > \frac{W}{2} \right]_{1 \leq j \leq n} \\ r_i, & \text{en caso contrario } i = 1,2, \dots, n \end{cases}$$

Estas transformaciones resaltan los colores alrededor del prototipo, al forzar que todos los demás colores se ubiquen en el punto medio del espacio de color de referencia. [8]

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

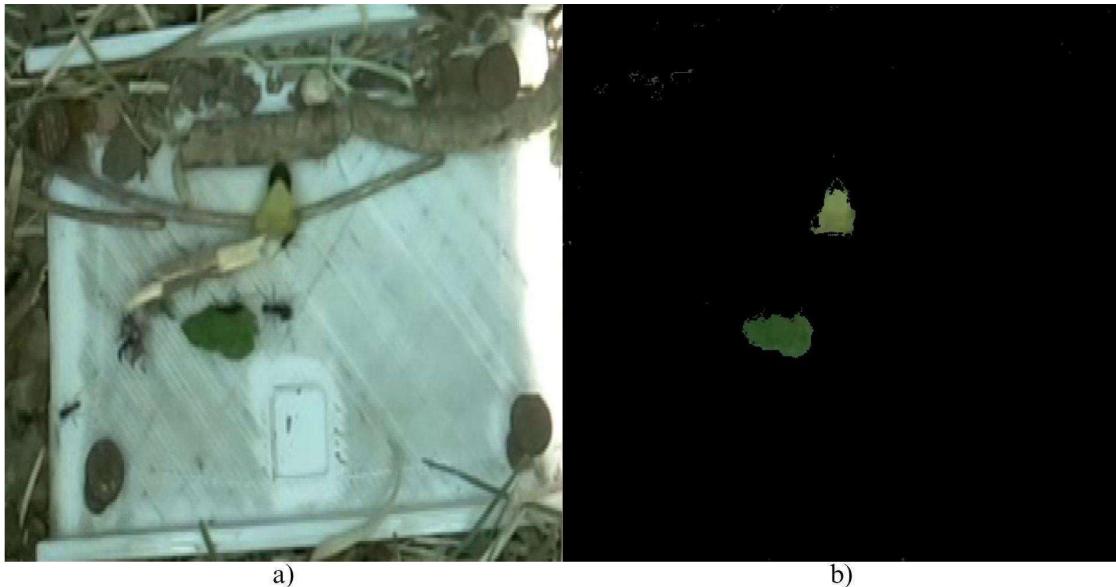


Figura 5 - Segmentación de color que detecta colores verdes. a) Imagen original. b) Imagen recortada.

## 2.8 Imágenes binarias

### 2.8.1 Máscara

Las máscaras binarias constituyen una herramienta fundamental en el procesamiento digital de imágenes, permitiendo la selección precisa de regiones de interés (ROI) para su análisis o modificación. Estas estructuras, definidas como matrices bidimensionales de valores 0 y 1, actúan como filtros espaciales que destacan u ocultan porciones específicas de una imagen, ya sea en escala de grises o a color. Su implementación facilita operaciones como la medición de atributos locales, la eliminación de fondos no deseados o la preparación de datos para algoritmos avanzados de visión artificial. [114]

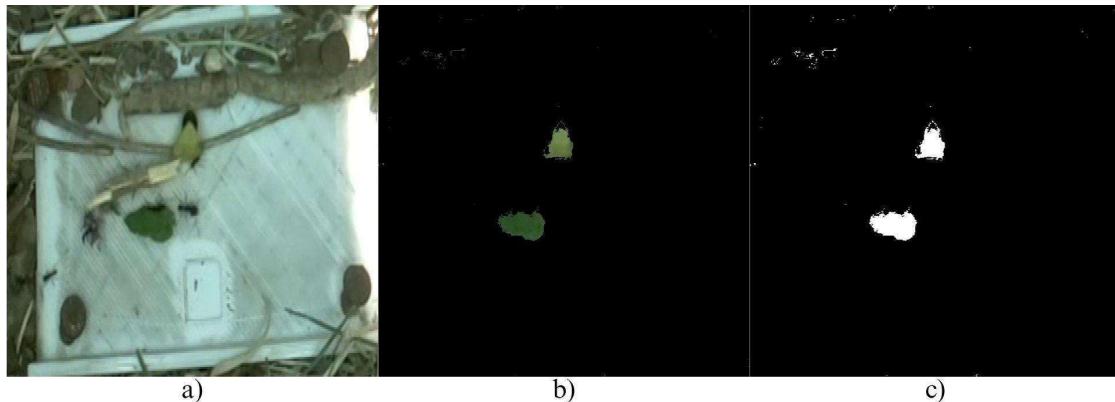


Figura 6 - Ejemplo de aplicación de una máscara a una imagen a color. (a) Imagen original. (b) imagen producida por un umbral del canal verde de (a). (c ) máscara binaria de (b).

La figura 6 muestra la operación de aplicar una imagen binaria a una imagen a color para separar diferentes secciones de la imagen para aislarlas y medirlas de manera más precisa [63].

La máscara binaria se puede utilizar en una superposición o canal alfa para seleccionar los píxeles que se mostrarán, o la máscara puede usarse para modificar la imagen almacenada. Esto se puede lograr multiplicando la imagen en escala de grises por la imagen binaria, con la convención de que los valores de la imagen binaria son 0 (OFF) o 1 (ON) en cada píxel [113]. Alternativamente, se obtiene el mismo resultado combinando las imágenes en escala de grises y binaria para conservar el valor más brillante. Por ejemplo, si la máscara es blanca (valor = 255) para los píxeles de fondo y negra (valor = 0) para los píxeles de primer plano, entonces los valores de píxeles más brillantes en cada ubicación borran todos los píxeles de fondo y conservan el valor en escala de grises o en color para los píxeles de primer plano.

## 2.8.2 Segmentación

En el contexto del procesamiento digital de imágenes, antes de realizar mediciones, especialmente aquellas que involucran automatización, es esencial llevar

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

a cabo un proceso llamado segmentación. Este procedimiento implica la identificación y separación de las regiones de interés, como objetos o estructuras, del fondo circundante en una imagen [97].

La segmentación de objetos representa uno de los pilares fundamentales en el procesamiento digital de imágenes y la visión por computador, constituyendo el proceso mediante el cual se divide una imagen digital en múltiples segmentos o regiones que corresponden a objetos o partes significativas de una escena. Esta técnica busca simplificar y transformar la representación de una imagen en algo más significativo y fácil de analizar, permitiendo localizar objetos y límites específicos dentro del contenido visual [70]. La importancia de esta disciplina radica en su capacidad para extraer información estructural de datos visuales aparentemente, transformando la percepción computacional de manera similar a como el sistema visual humano organiza la información retinal en estructuras coherentes.

La segmentación se define como la asignación de etiquetas a píxeles agrupados por características visuales o geométricas comunes. Según González y Woods , este proceso subdivide una imagen en regiones u objetos, donde el nivel de detalle depende de los requisitos específicos de la aplicación. Los métodos se clasifican en dos categorías principales: basados en discontinuidades (detección de bordes) y basados en similitudes (crecimiento de regiones) [8]

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico



*Figura 7 - Ejemplo de aplicación de una máscara a una imagen a color. (a) Imagen original. (b) imagen binaria producida por el modelo de segmentación de (a).*

Este proceso no solo tiene un impacto positivo en la precisión de nuestras mediciones, sino que también optimiza la eficiencia de nuestros algoritmos de procesamiento. Al enfocar los recursos computacionales en las partes relevantes de la imagen, obtenemos mediciones más confiables y ahorramos tiempo y recursos en comparación con un enfoque más amplio [15].

## 2.9 Visión artificial

La visión artificial tiene como objetivo dotar a los sistemas computacionales de la capacidad de adquirir, procesar, analizar y comprender imágenes del mundo real, produciendo información significativa sobre ellas [8]. Esta disciplina se encuentra estrechamente relacionada con el procesamiento digital de imágenes y ha experimentado un gran avance gracias al uso de técnicas de deep learning que provee herramientas específicas para automatizar la extracción de características mediante arquitecturas neuronales [98]. Esta sinergia se manifiesta en aplicaciones

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

como la segmentación semántica, donde las redes convolucionales (CNN) aprenden a asignar etiquetas a píxeles mediante el análisis de patrones locales y globales en tensores de entrada. Los métodos clásicos de visión artificial (detección de bordes) requieren ingeniería manual de características, mientras que el deep learning automatiza este proceso mediante el aprendizaje basado en datos [98].

### 2.9.1 Algoritmo de aprendizaje

Un algoritmo de aprendizaje automático es aquel que puede aprender a partir de datos. Según Mitchell, una definición formal y ampliamente aceptada es:

“Se dice que un programa de computadora aprende a partir de la experiencia E con respecto a alguna clase de tareas T y medida de rendimiento P, si su rendimiento en las tareas en T, medido por P, mejora con la experiencia E” [10].

Para tener un problema de aprendizaje bien definido, deben identificarse tres componentes clave:

- La clase de tareas (T): qué tipo de actividades debe realizar el sistema.
- La medida de rendimiento (P): cómo se evalúa el desempeño del sistema en dichas tareas.
- La fuente de experiencia (E): los datos o ejemplos utilizados para que el sistema aprenda [16].

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

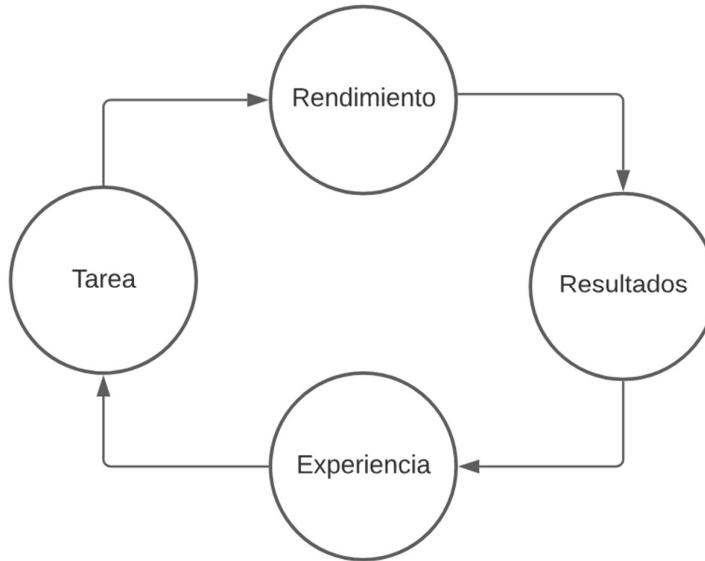


Figura 8 - Diagrama del funcionamiento de un algoritmo de aprendizaje básico.

### 2.9.2 Tensor

Un tensor es una generalización matemática que permite representar datos en múltiples dimensiones, extendiendo los conceptos de escalares, vectores y matrices. En términos básicos, se puede entender como un arreglo multidimensional cuya cantidad de dimensiones define su rango o orden. Por ejemplo, un escalar corresponde a un tensor de rango cero, un vector a uno de rango uno, una matriz a uno de rango dos, y así sucesivamente [11].

En el contexto del aprendizaje profundo, los tensores son la estructura fundamental para almacenar y procesar información dentro de las redes neuronales. Su uso permite organizar eficientemente grandes volúmenes de datos, facilitando operaciones matemáticas complejas como convoluciones, transformaciones lineales y funciones no lineales aplicadas a lo largo de sus diferentes dimensiones [9].

Un ejemplo común es la representación de una imagen en color. Una imagen RGB se almacena normalmente en un tensor tridimensional, donde las tres dimensiones corresponden al ancho, al alto y a los canales de color (rojo, verde y azul). Esto permite manejar de forma integrada toda la información visual contenida en la imagen. Cuando se trabaja con lotes de imágenes -como suele ser el caso en

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

el entrenamiento de redes neuronales convolucionales-, se utiliza un tensor de cuatro dimensiones. La primera dimensión representa el número de imágenes en el lote, la segunda corresponde a los canales de color, y las tercera y cuarta indican la posición espacial de cada píxel en las direcciones horizontal y vertical respectivamente [11].

Esta forma de organización no solo mejora la eficiencia computacional, sino que también facilita la implementación de operaciones paralelas en hardware especializado como GPUs, lo cual es clave para el desempeño de los modelos modernos de inteligencia artificial.

### 2.9.3 Redes neuronales.

Una red neuronal, o más específicamente una red neuronal artificial (ANN), es un modelo computacional inspirado en la estructura y funcionamiento de los sistemas nerviosos biológicos. Estas redes son un enfoque clave dentro del aprendizaje automático, una técnica que permite a los sistemas informáticos mejorar con la experiencia y los datos. Las redes neuronales artificiales se construyen a partir de un gran número de unidades computacionales simples, o "neuronas artificiales", que se interconectan para lograr un comportamiento inteligente [9].

La arquitectura básica de una red neuronal feedforward, como el perceptrón multicapa (MLP), consiste en una capa de entrada, una o más capas ocultas y una capa de salida. Cada conexión entre neuronas posee un peso asociado, que se ajusta durante el proceso de entrenamiento. Las neuronas reciben entradas ponderadas y aplican una función de activación no lineal para producir una salida. La propagación hacia adelante en estas redes implica el paso de la información desde la capa de entrada, a través de las capas ocultas, hasta la capa de salida, generando un resultado [9].

Las redes neuronales se han aplicado con éxito en diversas áreas, incluyendo el procesamiento de imágenes, reconocimiento de voz, y procesamiento del lenguaje natural. En el procesamiento de imágenes, se utilizan para tareas como la clasificación y el reconocimiento de objetos [9].

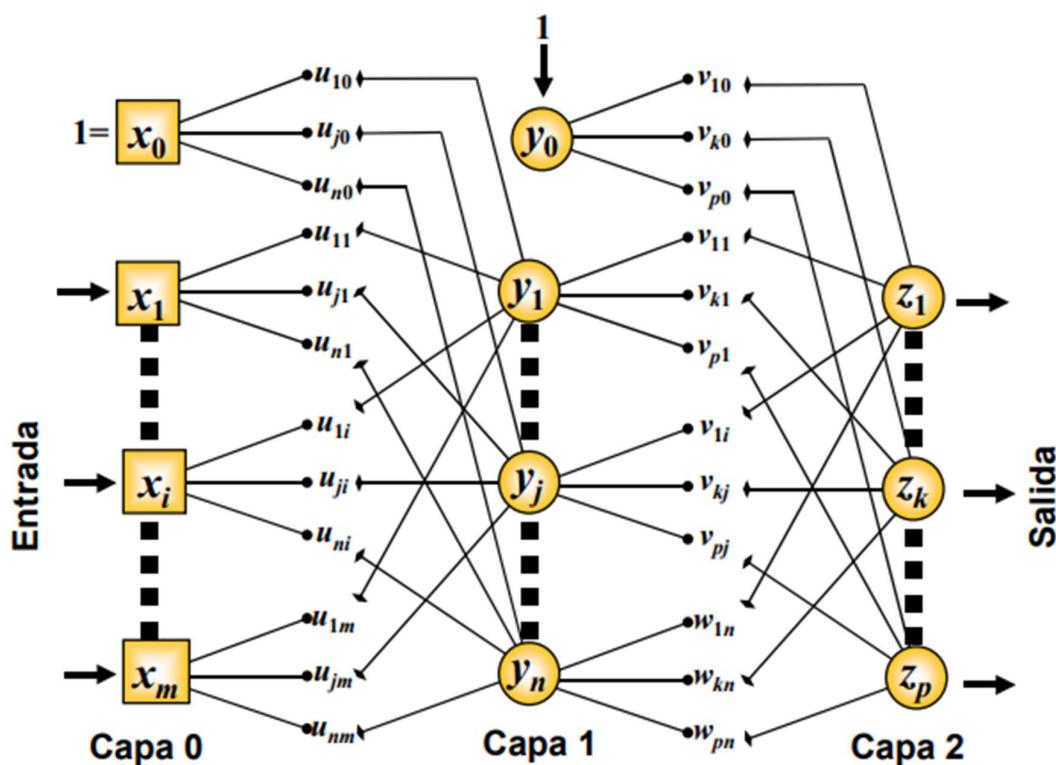


Figura 9 - Esquema de una red neuronal multicapa.  $X_n$  las entradas,  $Z_n$  las salidas y las capas intermedias son las capas ocultas ( $y_n$ ) [116].

Las redes neuronales poseen una notable capacidad de adaptación a diversos tipos de datos y tareas, lo cual logran ajustando sus parámetros internos, como pesos y sesgos, a través de diferentes paradigmas de aprendizaje. El aprendizaje supervisado, un enfoque común, implica entrenar la red con conjuntos de datos que incluyen tanto las entradas como las salidas deseadas, permitiendo a la red aprender a mapear entradas a salidas correctas. En contraste, el aprendizaje no supervisado se centra en encontrar patrones ocultos o estructuras dentro de datos no etiquetados, como la agrupación o la reducción de dimensionalidad. Existe también el aprendizaje por refuerzo, donde un agente aprende a tomar decisiones interactuando con un entorno para maximizar una señal de recompensa [9].

Para la aplicación desarrollada en este proyecto, se emplea un enfoque de aprendizaje supervisado, fundamentado en técnicas de aprendizaje profundo.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

Comparadas con otros métodos de aprendizaje automático, las redes neuronales presentan varias ventajas, incluyendo su habilidad para modelar relaciones complejas y no lineales en los datos, así como su escalabilidad a grandes conjuntos de datos y arquitecturas profundas. Su capacidad de generalización a datos no vistos también es una característica destacada [9]. Sin embargo, también enfrentan desafíos, como el elevado costo computacional requerido para su entrenamiento, la dificultad en la interpretabilidad de sus decisiones ("caja negra"), y la susceptibilidad al sobreajuste, especialmente en modelos muy complejos o con datos limitados.

#### 2.9.4 Deep Learning

El aprendizaje profundo (deep learning) representa un subcampo del aprendizaje automático que permite a los sistemas computacionales aprender representaciones de datos con múltiples niveles de abstracción. La idea central radica en que el sistema aprende a comprender el mundo a través de una jerarquía de conceptos, donde representaciones más abstractas se calculan en términos de representaciones menos abstractas, basándose en la información recolectada de los datos. Este enfoque de aprendizaje jerárquico elimina la necesidad de programar manualmente reglas o especificar explícitamente todo el conocimiento necesario para una tarea [9].

La estructura jerárquica permite que los modelos de aprendizaje profundo construyan conceptos complejos combinando o componiendo conceptos más simples. Si visualizamos esta composición como capas apiladas, donde cada capa aprende a representar los datos en un nivel de abstracción más alto que la capa anterior, obtenemos una arquitectura "profunda" con múltiples capas de procesamiento no lineal [9] [11]. Esta profundidad es una característica distintiva que permite a los modelos de aprendizaje profundo aprender patrones y relaciones altamente complejos y abstractos directamente de los datos [11] [12]. Este proceso



de aprendizaje de representaciones jerárquicas a menudo se realiza de manera automática a partir de grandes volúmenes de datos.

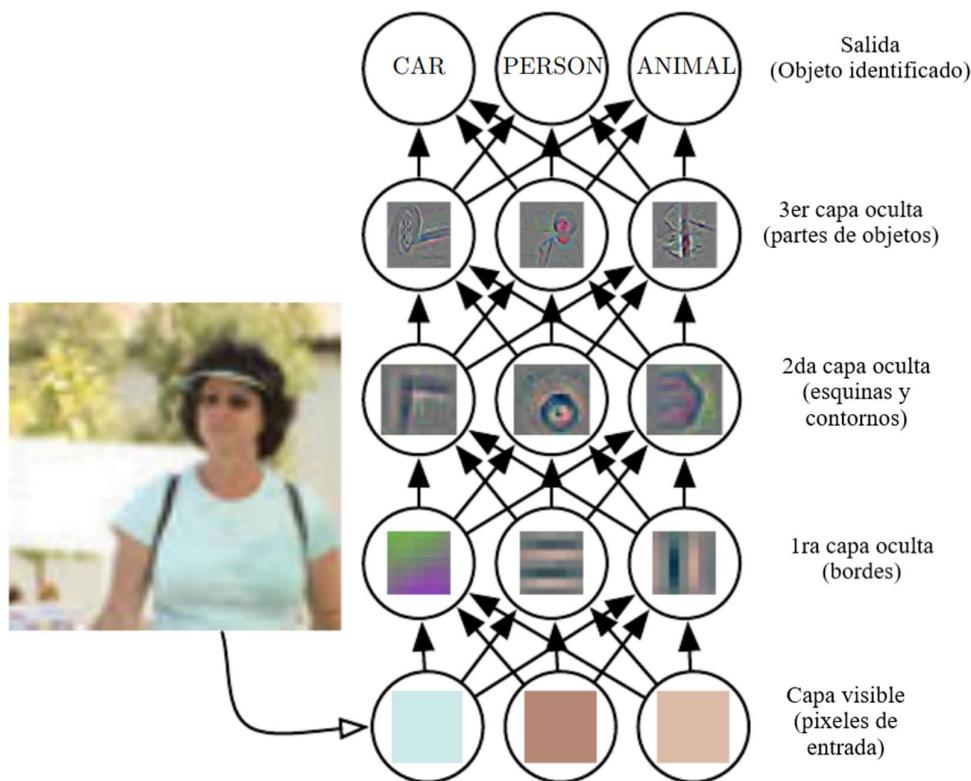


Figura 10 - Ilustración de un modelo de aprendizaje profundo [9].

### 2.9.5 Redes neuronales convolucionales

Las Redes Neuronales Convolucionales (CNNs) constituyen una clase especializada de redes neuronales diseñada para procesar datos con una topología similar a una cuadrícula, como las imágenes. Su arquitectura está inspirada en la organización del sistema visual de los animales, emulando la forma en que la corteza visual procesa la información visual jerárquicamente. Las CNNs han demostrado ser particularmente efectivas en tareas de análisis visual debido a su capacidad para identificar automáticamente patrones y características relevantes de los datos de entrada [11].

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

## 2.9.6 Componentes de las Redes Neuronales Convolucionales.

Los siguientes son los componentes típicos de una red neuronal convolucional:

- “Capa de entrada: Esta capa contiene la intensidad de píxeles de la imagen. Por ejemplo, una imagen de entrada con un ancho de 64, alto de 64 y una profundidad de 3 canales de color (Rojo, Verde y Azul, o RGB) tendría dimensiones de entrada de 64 x 64 x 3.
- Capa de convolución: Esta capa toma imágenes de las capas precedentes y las convoluciona, se entiende por convolución a la operación matemática fundamental que aplica un kernel (matriz de pesos pequeños) sobre una imagen de entrada mediante desplazamientos locales. Este proceso calcula el producto punto entre los valores del kernel y los píxeles subyacentes de la región de la imagen que cubre, generando mapas de características que resaltan patrones espaciales como bordes, texturas o formas [99][100]. El número de mapas de características de salida es igual al número especificado de filtros. Hasta ahora, en general, las CNNs han utilizado principalmente filtros 2D, aunque recientemente se han introducido filtros de convolución 3D.
- Función de activación: Para las CNN, las funciones de activación suelen ser ReLU (Rectified Linear Unit). La dimensión de salida es la misma que la entrada después de pasar por las capas de activación ReLU. La capa ReLU agrega no linealidad a la red y, al mismo tiempo, proporciona gradientes no saturantes para entradas netas positivas.
- Capa de agrupación (Pooling): Esta capa reduce la resolución de los mapas de activación 2D a lo largo de las dimensiones de altura y ancho. La profundidad o el número de mapas de activación no se ve comprometida y permanece igual.
- Capas completamente conectadas: Contienen neuronas tradicionales que reciben conjuntos diferentes de pesos de las capas precedentes; no hay compartición de pesos entre ellas, como es típico en las operaciones de



convolución. Cada neurona en esta capa estará conectada o bien a todas las neuronas en la capa anterior o a todas las salidas coordinadas en los mapas de salida a través de pesos separados. Para la clasificación, las neuronas de salida de clase reciben entradas de las últimas capas completamente conectadas” [11].

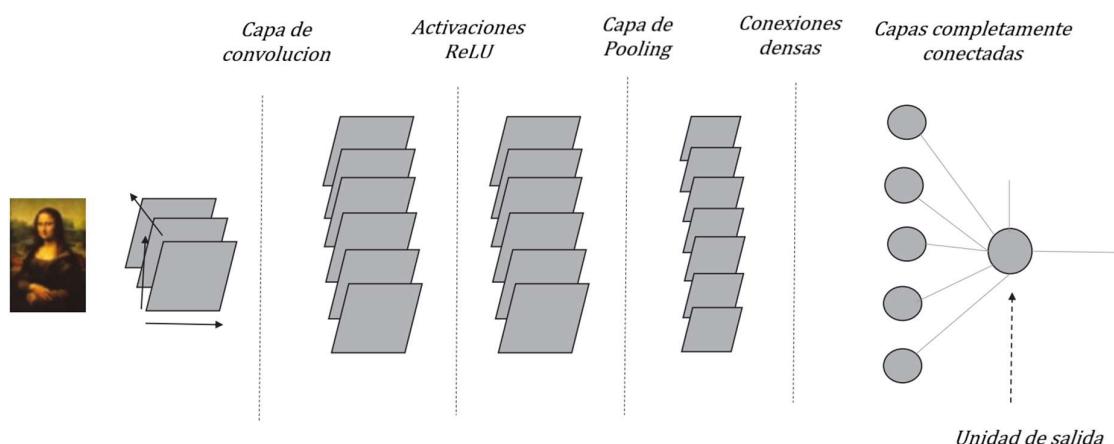


Figura 11 - Diagrama de flujo básico de una red neuronal convolucional [11].

## 2.10 Seguimiento de objetos

El seguimiento de objetos constituye una tarea fundamental en visión por computadora que implica la localización y rastreo continuo de entidades específicas a través de secuencias de imágenes [8]. Este proceso presenta desafíos significativos debido a la variabilidad inherente en las condiciones de captura, incluyendo cambios de iluminación, occlusiones parciales, transformaciones geométricas y variaciones en la apariencia del objeto objetivo [12]. La estabilidad frente a estas condiciones dinámicas representa uno de los principales criterios de evaluación para cualquier sistema de seguimiento efectivo.

Las arquitecturas basadas en redes neuronales convolucionales han revolucionado el campo del seguimiento visual debido a su capacidad para aprender representaciones jerárquicas de características visuales [9]. Estas redes logran capturar tanto patrones de bajo nivel como características semánticas complejas

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	2 - Desarrollo teórico

mediante la aplicación sucesiva de operaciones convolucionales, permitiendo una representación robusta del objeto bajo diversas condiciones ambientales [8].

En términos de implementación práctica, las plataformas de aprendizaje profundo como TensorFlow y PyTorch han facilitado significativamente el desarrollo de sistemas de seguimiento [13]. Estas herramientas proporcionan implementaciones optimizadas de arquitecturas de CNN especializadas, así como modelos pre-entrenados que pueden adaptarse mediante técnicas de transferencia de aprendizaje para aplicaciones específicas [9]. La disponibilidad de estos recursos ha democratizado el acceso a tecnologías de seguimiento avanzadas, permitiendo su implementación en una amplia gama de aplicaciones industriales y de investigación.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

## Capítulo 3: Materiales y métodos

### 3.1 Materiales

En el presente capítulo se describen las tecnologías utilizadas que desempeñarán un rol fundamental en el desarrollo de nuestro proyecto.

Python es un lenguaje de programación interpretado, interactivo y orientado a objetos que se ha elegido para el desarrollo de una aplicación debido a su sintaxis clara y legible, una amplia comunidad, librerías especializadas, escalabilidad y flexibilidad, integración con otras herramientas y un ecosistema robusto. Python incorpora módulos, excepciones, tipificación dinámica, tipos de datos dinámicos de muy alto nivel y clases. Admite múltiples paradigmas de programación más allá de la programación orientada a objetos, como la programación procedural y funcional.

Estas características permiten desarrollar aplicaciones escalables y lo convierten en el lenguaje predilecto para la mayoría de tareas asociadas a data-scientists, modelos de aprendizaje automático, procesamiento de imágenes, etc. Además, Python posee varias bibliotecas y herramientas que se pueden utilizar para aplicar cada una de estas arquitecturas de redes neuronales convolucionales (CNN). La herramientas que usamos para el desarrollo fueron:

#### 3.1.1 PyTorch

PyTorch [117] es una biblioteca popular de Python para el desarrollo de redes neuronales. Proporciona una plataforma flexible y eficiente para construir y entrenar modelos de redes neuronales, incluidas las CNN. PyTorch es conocido por su facilidad de uso y su capacidad para implementar de manera eficiente modelos complejos. Es una opción popular para la investigación y desarrollo en el campo del aprendizaje automático y el procesamiento de imágenes.

Una de las características distintivas de PyTorch es su arquitectura dinámica de gráficos computacionales. Esto permite que los desarrolladores definan y

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

modifiquen el grafo de cálculo durante el tiempo de ejecución, lo que facilita la depuración y la experimentación interactiva.

En términos de aplicaciones de visión por computadora, PyTorch es ampliamente utilizado [94] [95] [96] y admite la implementación de variantes de CNN, como Mask CNN y RCNN. Además, PyTorch ha sido ampliamente adoptado por la comunidad de investigación debido a su flexibilidad y la facilidad para implementar nuevos enfoques de aprendizaje profundo.

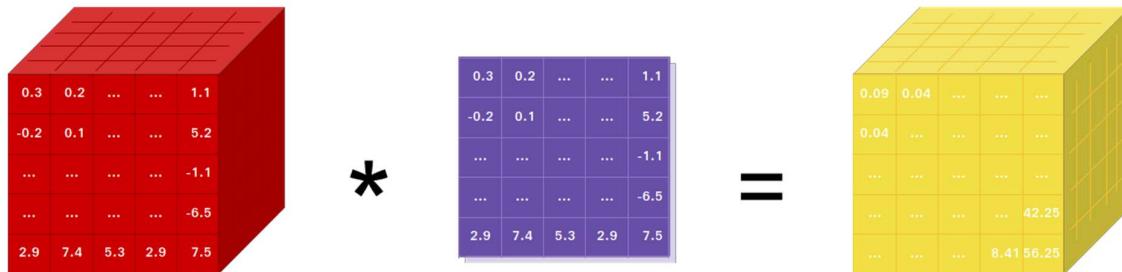


Figura 12 - Representación gráfica de una convolución entre un tensor y un filtro [117].

### 3.1.2 OpenCV

OpenCV es una biblioteca de visión por computadora que se utiliza para realizar operaciones de procesamiento de imágenes, detección de objetos y segmentación en imágenes. Proporciona diversas herramientas y algoritmos eficientes para trabajar con imágenes y es compatible con la implementación de objetos detectores, incluidas las variantes de las CNN [64].

Una de las principales ventajas de OpenCV es su amplia disponibilidad y portabilidad en múltiples plataformas. Puede utilizarse con diversos lenguajes de programación, incluidos C++, Python y Java, lo que facilita su integración en una amplia gama de aplicaciones [64]. En el contexto de las redes neuronales convolucionales, OpenCV se puede utilizar para mejorar y optimizar la detección de objetos, mejorar la precisión de las máscaras generadas por modelos como Mask R-CNN y brindar herramientas para realizar tareas de segmentación semántica en imágenes. Al combinar las capacidades de OpenCV con las redes neuronales



convolucionales, los desarrolladores pueden construir aplicaciones más avanzadas y precisas en el campo de la visión por computadora.

### 3.1.3 You Only Look Once (YOLO)

YOLO es un algoritmo de detección de objetos en tiempo real que utiliza una única pasada a través de la red para detectar los objetos. A diferencia de otros enfoques que dividen la detección en múltiples etapas, YOLO divide la imagen en una cuadrícula y predice las cajas delimitadoras y las probabilidades de cada clase en cada celda. Se desarrollará en profundidad en la sección 3.2.4.

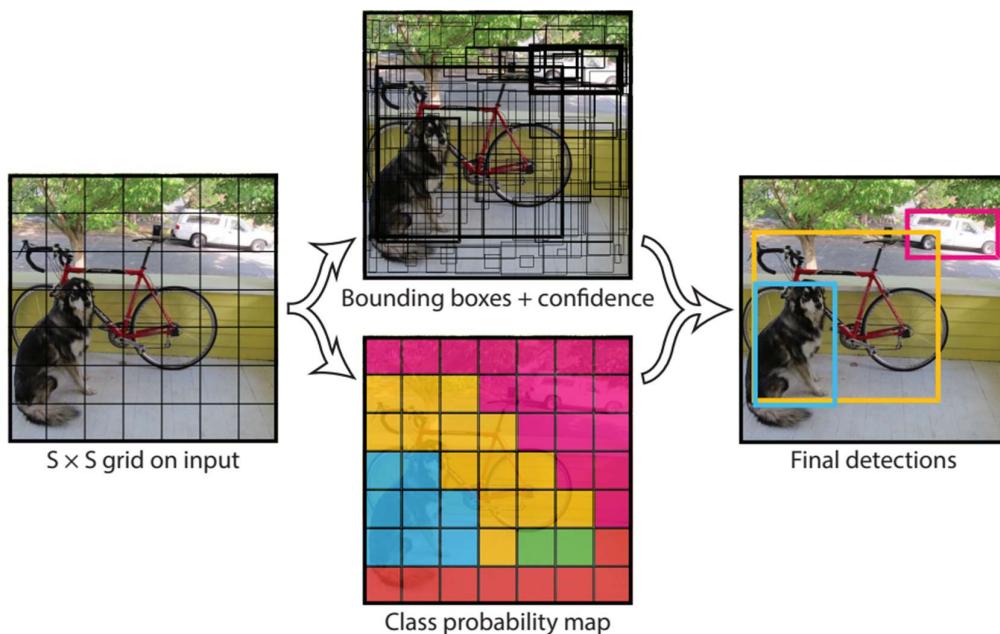


Figura 13 - División de celdas realizada por YOLO [105].

YOLO es conocido por su velocidad y puede detectar múltiples objetos en tiempo real. A diferencia de otros enfoques que requieren múltiples pasadas por la imagen.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

### 3.1.4 Labelme

Labelme permite a un usuario enseñar a una computadora acerca de una imagen utilizando diversas formas de anotación, que incluyen formas (como polígonos, rectángulos, líneas y puntos), etiquetas, grupos, descripciones de texto y combinaciones de estos elementos. Esta anotación humana permite que una computadora adquiera la inteligencia necesaria para realizar una tarea que los humanos desean que la computadora realice como inteligencia artificial (IA).

### 3.1.5 Roboflow

Roboflow es una plataforma diseñada para facilitar el desarrollo y la implementación de modelos de visión por computadora. Su principal función es simplificar el proceso de manejo de datos, permitiendo a los usuarios etiquetar, preprocesar y aumentar conjuntos de imágenes. Una de sus ventajas es su capacidad para exportar los datos preparados en formatos compatibles con frameworks populares como TensorFlow, PyTorch y YOLO.

### 3.1.6 Anaconda

Es la aplicación líder en creación de entornos de desarrollo independientes. Ya sea que deseas construir modelos de ciencia de datos/aprendizaje automático, implementar tu trabajo en producción o gestionar de manera segura a un equipo de ingenieros. Esto permite trabajar instalando diferentes paquetes de diferentes versiones en cada entorno sin perjudicar a los demás entornos de trabajo. (<https://docs.anaconda.com/>)

### 3.1.7 Tkinter (Tk/Tcl)

Tk/Tcl ha sido durante mucho tiempo una parte integral de Python. Proporciona un conjunto de herramientas robusto e independiente de la plataforma para administrar ventanas. tkinter es un conjunto de envoltorios que implementan los

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

widgets Tk como clases de Python. Las principales virtudes de tkinter son que es rápido y que generalmente viene incluido con Python. Aunque su documentación estándar es débil, se dispone de buen material, que incluye: referencias, tutoriales, un libro y otros. tkinter su aspecto obsoleto se ha mejorado enormemente en Tk 8.5.

### 3.1.8 CustomTKinter

CustomTkinter es una biblioteca de interfaz de usuario (UI) para aplicaciones de escritorio en Python, basada en Tkinter, que ofrece widgets modernos y completamente personalizables. Con CustomTkinter, podremos lograr una apariencia coherente y actualizada.

## 3.2 Métodos

### 3.2.1 Investigación

La primera actividad al abordar el proyecto fue la investigación de las tecnologías disponibles para detección de objetos. Tras un análisis comparativo, identificamos dos arquitecturas prominentes: YOLOv8 (You Only Look Once, versión 8) y Mask R-CNN. Optamos por implementar YOLOv8 debido a sus ventajas distintivas. Por un lado, su arquitectura optimizada ofrece mejoras significativas en relación con la velocidad y precisión [14]. Esto tiene una consecuencia positiva sobre su capacidad mejorada para la detección de objetos en diversas tareas de visión artificial [15], y su implementación unificada que integra detección, segmentación y clasificación en un solo framework [14]. Mientras que Mask R-CNN ofrece una segmentación de instancias más detallada, YOLOv8 proporciona un rendimiento superior en tiempo real con una precisión competitiva según análisis comparativos recientes [16] [17], aspecto fundamental para aplicaciones de seguimiento de objetos como es el caso de las hormigas cortadoras. Adicionalmente, la extensa documentación y herramientas de optimización disponibles para YOLOv8 facilitaron

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

considerablemente el desarrollo del proyecto [18]. Esta elección tecnológica sentó las bases para el desarrollo de nuestro software en sus etapas iniciales.

### 3.2.2 Recopilación de imágenes

La fase inicial de creación del dataset es crítica para el entrenamiento de modelos de detección de objetos. Este proceso requiere recopilar una amplia variedad de imágenes que capturen hormigas con diversos tipos de carga, bajo diferentes condiciones de iluminación y en distintos entornos. Tal diversidad es esencial para desarrollar modelos capaces de generalizar en condiciones variables.

Para nuestro estudio, obtuvimos las imágenes a partir de grabaciones de vídeo de 30 minutos que documentan el comportamiento de hormigas cortadoras de hojas en un entorno controlado. Estos especímenes fueron registrados mientras se desplazaban con o sin carga sobre una superficie blanca homogénea, facilitando así el procesamiento posterior de las imágenes [20] . Este material fue proporcionado por el Laboratorio de Ecología de Hormigas de la Facultad de Ciencias Agropecuarias (UNER) dirigido por el Dr. Julián A. Sabattini [21].

El conjunto de datos inicial constaba de 185 imágenes para entrenamiento, posteriormente ampliado con 202 imágenes adicionales, alcanzando un total de 387 muestras. Todas las imágenes fueron redimensionadas a una resolución uniforme de  $640 \times 640$  píxeles, siguiendo las prácticas recomendadas para el entrenamiento con arquitecturas basadas en YOLO [22]. Esta estandarización de resolución no solo optimizó la calidad visual, sino que también incrementó la eficiencia computacional durante la fase de entrenamiento.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos



Figura 14 - Imagen original e imagen escalada a 640 x 640

Para fortalecer aún más el dataset, incorporamos 248 imágenes adicionales, elevando el total a 635. Aplicamos técnicas de aumento de datos (data augmentation), mediante la plataforma RoboFlow, para expandir artificialmente la diversidad del conjunto de entrenamiento [24]. Las transformaciones implementadas incluyeron rotaciones ( $\pm 90^\circ$  y  $\pm 15^\circ$ ) y diferentes tipos de zoom, llevando a la cantidad total de 774 imágenes. Estas técnicas fueron fundamentales para mejorar la generalización del modelo: las rotaciones potencian la invarianza a la orientación, mientras que las modificaciones de iluminación prepararon al sistema para adaptarse a diferentes condiciones ambientales [23]. Posteriormente se analizaron 6.388 imágenes a partir del trabajo realizado por el Laboratorio de Ecología de Hormigas (FCA-UNER). Como resultado de eso, 13.895 hormigas con carga de vegetal de *Acromyrmex lundi* para la Mesopotamia Argentina fueron incorporadas al estudio, logrando así un dataset ([https://universe.roboflow.com/proyecto-hormiga/acromyrmex\\_lundi](https://universe.roboflow.com/proyecto-hormiga/acromyrmex_lundi)) definitivo de mayor tamaño y robusto para así poder analizar diferentes escenarios.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

### 3.2.3 Anotación del dataset

Este proceso define explícitamente qué elementos debe identificar el algoritmo en las imágenes, estableciendo así el fundamento para el posterior entrenamiento supervisado. Para nuestro proyecto de detección de hormigas y su carga, implementamos una metodología de anotación en dos fases.

Inicialmente, empleamos el software Labelme [24], una herramienta de anotación de código abierto que permite la delimitación precisa de objetos mediante polígonos personalizables [25]. Esta aproximación poligonal, en contraste con las tradicionales cajas delimitadoras rectangulares, resultó especialmente adecuada para capturar la morfología irregular tanto de las hormigas como de sus cargas, mejorando así la precisión de las anotaciones [26].



Figura 15 - Anotación segmentada de la carga en la herramienta Labelme

En una segunda fase, migramos el proceso a la plataforma RoboFlow, un entorno especializado que integra capacidades avanzadas de gestión de datasets para visión artificial. Esta herramienta no solo facilitó la estandarización de nuestras anotaciones, sino que también automatizó la implementación de técnicas de aumento de datos y preprocesamiento, generando transformaciones controladas que expandieron significativamente la diversidad del conjunto de entrenamiento [19] [28].

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

Debido a la incorporación de las 6.400 imágenes adicionales, se obtuvieron anotaciones de aproximadamente 12.000 cargas transportadas por hormigas.

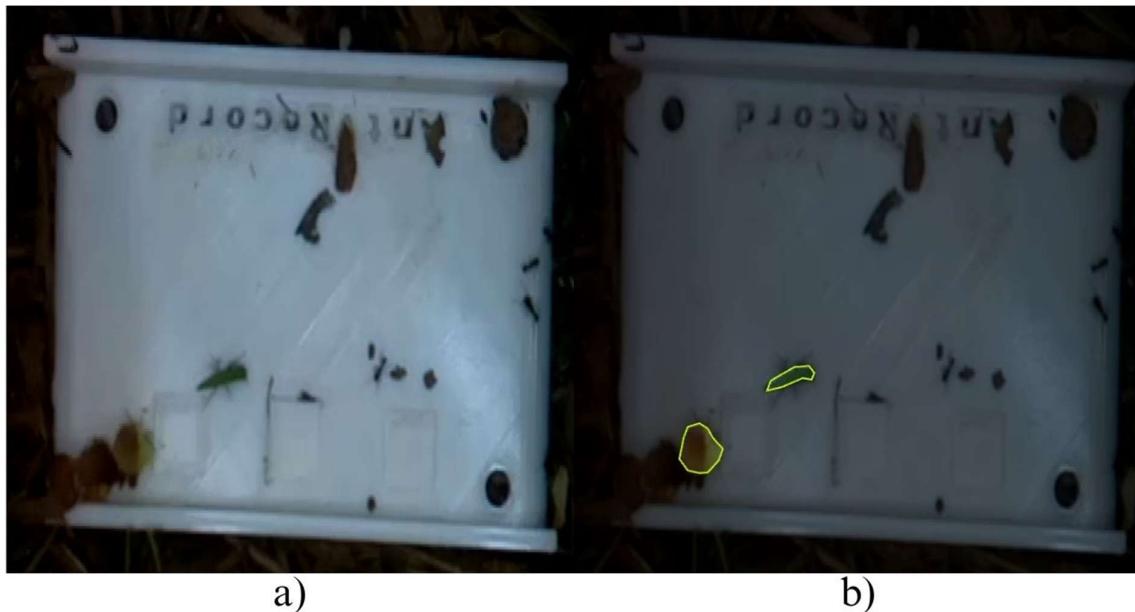


Figura 16 - Anotación segmentada de la carga en la herramienta Roboflow.

Un desafío particular en nuestro estudio fue la naturaleza dinámica y altamente variable de la carga transportada por las hormigas. A diferencia de objetos rígidos con formas predecibles, la carga de estos insectos presenta una geometría inconsistente y una orientación cambiante durante su desplazamiento [27]. Para abordar esta complejidad, desarrollamos un protocolo de anotación continua, donde cada instancia de carga fue etiquetada a lo largo de toda su trayectoria visible, incorporando así las variaciones posicionales en el entrenamiento del modelo. Este enfoque está respaldado por investigaciones recientes que demuestran que la anotación exhaustiva de objetos en diferentes estados mejora significativamente la robustez de los modelos de detección [29] [30].

La combinación de Labelme para anotaciones geométricamente precisas y RoboFlow para la gestión y aumento del dataset nos permitió construir un conjunto de entrenamiento diverso y representativo. Esta estrategia dual capacitó a nuestro modelo para aprender patrones generalizables, reduciendo el riesgo de sobreajuste

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

y mejorando su adaptabilidad a la variabilidad inherente en aplicaciones de monitoreo de insectos en entornos reales.

### 3.2.4 Entrenamiento

Para la implementación de nuestro sistema de detección, seleccionamos YOLOv8, desarrollado por Ultralytics [18]. Esta elección específica se fundamenta en tres ventajas críticas que ofrece YOLOv8 para nuestro caso de estudio:

1. Rendimiento optimizado en la detección de objetos pequeños y con occlusiones parciales, característicos de las cargas de hormigas en entornos naturales [31] [35];
2. Capacidad mejorada para la segmentación de instancias con contornos complejos e irregulares, esencial para la morfología variable de las cargas transportadas [33];
3. Eficiencia computacional superior para procesamiento en tiempo real, requisito fundamental para el análisis de comportamiento dinámico de las hormigas [34].

YOLOv8 se distribuye en cinco variantes escalables que ofrecen diferentes equilibrios entre precisión y velocidad de procesamiento. Cada variante —nano, pequeño, mediano, grande y extra grande— está optimizada para distintos escenarios de implementación, desde dispositivos con recursos limitados hasta sistemas de alto rendimiento [33] [18]. Esta flexibilidad resultó particularmente valiosa para nuestro proyecto, permitiéndonos experimentar con diferentes configuraciones hasta identificar la óptima para nuestras restricciones específicas de hardware y requisitos de precisión.

A diferencia de sus predecesores, YOLOv8 implementa una arquitectura de cuello (neck) mejorada con Path Aggregation Network (PANet) y características de atención espacial [33] [34], que aumentan significativamente la capacidad del modelo para distinguir hormigas individuales en escenarios de alta densidad poblacional. Adicionalmente, incorpora mejoras en el proceso de aprendizaje mediante Mosaic

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

Augmentation y Copy-Paste Augmentation [32], técnicas que resultan especialmente beneficiosas para nuestro conjunto de datos limitado de hormigas.

La detección de objetos mediante YOLOv8 identifica con precisión la ubicación y clasificación de entidades en imágenes o secuencias de vídeo. El resultado comprende cuadros delimitadores que encierran las cargas, junto con etiquetas de clase y puntuaciones de confianza para cada instancia. Estas puntuaciones facilitan la implementación de umbrales adaptativos según nuestros requerimientos específicos de precisión.

La segmentación de instancias, funcionalidad crítica para nuestro análisis, proporciona una delineación precisa del contorno de cada carga. Esta capacidad genera máscaras de píxeles que definen exactamente la morfología de cada entidad, permitiéndonos diferenciar con precisión las diversas configuraciones de carga que transportan las hormigas.

YOLOv8 incorpora modelos pre-entrenados con pesos inicializados mediante entrenamiento en MS COCO [5], un dataset de referencia que comprende 80 categorías semánticas distribuidas en más de 330,000 imágenes. Esta transferencia de conocimiento desde un entrenamiento generalizado a nuestro dominio específico de carga de HCH potencia significativamente la convergencia del modelo, requiriendo sustancialmente menos datos de entrenamiento para alcanzar un rendimiento óptimo, ventaja crucial considerando la naturaleza especializada de nuestro dataset.

Los resultados comparativos de evaluación, realizadas utilizando el dataset de MS COCO [37] entre las diferentes variantes de YOLOv8 se detallan en la Tabla 1, evidenciando el compromiso entre precisión y velocidad de procesamiento que permitió tomar la decisión sobre el sistema a implementar.

Modelo	tamaño (pixels)	mAPbox 50-95	mAPmask 50-95	Velocidad CPU ONNX (ms)	Velocidad A100 TensorRT (ms)
YOLOv8n-seg	640	36.7	30.5	96.1	1.21

	UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica		3 - Materiales y Métodos

YOLOv8s-seg	640	44.6	36.8	155.7	1.47
YOLOv8m-seg	640	49.9	40.8	317.0	2.18
YOLOv8l-seg	640	52.3	42.6	572.4	2.79
YOLOv8x-seg	640	53.4	43.4	712.1	4.02

Tabla 1 – Métricas de los modelos pre-entrenados de YOLO

Donde:

- Tamaño de imagen: Indica las dimensiones de entrada (altura × anchura) a las que se escalan las imágenes durante el procesamiento. Este parámetro establece un equilibrio entre calidad y rendimiento: resoluciones mayores preservan detalles finos de las hormigas y sus cargas, mientras que resoluciones menores optimizan la velocidad de procesamiento.
- mAPbox (Mean Average Precision - Boxes): Cuantifica la precisión media del modelo en la detección de objetos mediante cuadros delimitadores. Integra tanto la precisión (proporción de detecciones correctas) como la exhaustividad (proporción de objetos detectados) a través de múltiples umbrales de confianza. En el contexto de MS COCO, se calcula promediando valores de AP para umbrales de IoU de 0.5 a 0.95 [37]. Un valor más elevado indica mayor precisión en la localización de hormigas y sus cargas.
- mAPmask (Mean Average Precision - Masks): Evalúa la precisión de segmentación a nivel de píxel mediante máscaras de instancia. A diferencia del mAPbox, que solo considera localización aproximada, el mAPmask mide la exactitud con que el modelo delinea el contorno preciso de cada objeto, permitiendo capturar la morfología exacta de las hormigas y sus cargas.
- Velocidad CPU ONNX: Registra la latencia de inferencia (en milisegundos) cuando el modelo se ejecuta mediante ONNX (Open Neural Network Exchange) en CPU [38]. Este valor indica cuánto tarda el modelo en procesar



una sola imagen en entornos sin aceleración por GPU, aspecto relevante para implementaciones en dispositivos con recursos computacionales limitados.

- Velocidad A100 TensorRT: Representa el tiempo de inferencia (en milisegundos) utilizando TensorRT, una biblioteca optimizada para GPUs NVIDIA A100 [39]. Esta métrica ilustra el rendimiento potencial del modelo en infraestructura de alto rendimiento, relevante para procesamiento en tiempo real de secuencias de video.

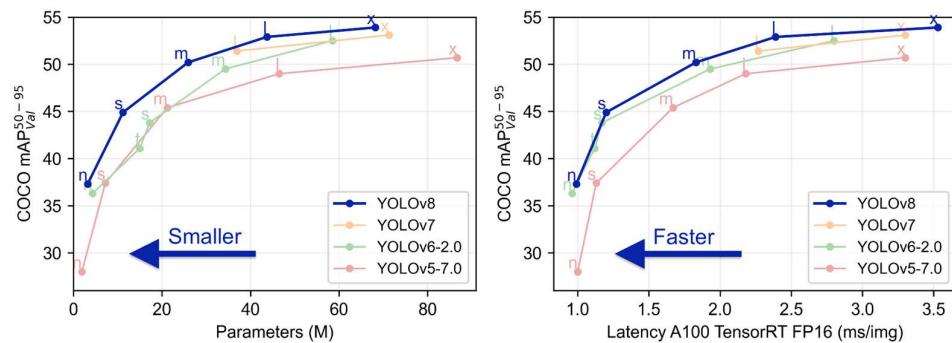


Figura 17 - Comparación de los parámetros de YOLOv8 con sus versiones anteriores. a) Curva de parámetros-mAP(Precisión mediana). b) Curva de velocidad(ms/img)-mAP [18].

### 3.2.4.1 Relevancia de los pesos pre-entrenados

Los pesos y sesgos pre-entrenados en MS COCO constituyen una transferencia de conocimiento fundamental para nuestro modelo especializado en detección de carga de HCH. Estos pesos encapsulan representaciones aprendidas sobre características visuales generalizables (bordes, texturas, formas) y patrones espaciales complejos que son transferibles entre dominios visuales.

La implementación de transfer learning [40] mediante estos pesos pre-entrenados aporta tres ventajas: (1) reducción significativa del volumen de datos de entrenamiento requerido, (2) aceleración en la convergencia durante el entrenamiento, disminuyendo el tiempo necesario para alcanzar un rendimiento óptimo [41] y (3) mejora en la generalización del modelo, permitiéndole adaptarse más efectivamente a variaciones [41].

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

Estos pesos pre-entrenados constituyen, por tanto, la base sobre la cual nuestro modelo construye un conocimiento específico para la detección y segmentación de la carga de HCH, optimizando tanto el proceso de entrenamiento como el rendimiento final del sistema.

### 3.2.5 Entrenamiento del modelo propio.

Para la implementación del sistema de detección y seguimiento de hormigas, seleccionamos YOLOv8m-seg, variante de tamaño mediano de YOLOv8 especializada en segmentación de instancias. Esta elección se fundamenta en su óptimo equilibrio entre precisión y eficiencia computacional, como observamos en la Tabla 1, característica esencial para el análisis de comportamiento de cargas transportadas por hormigas donde se requiere detección precisa de estructuras pequeñas e irregulares.

El proceso de entrenamiento se estructuró mediante la creación de cuatro variantes del conjunto de datos, cada una representando diferentes estrategias de preparación de los datos para optimizar el aprendizaje del modelo:

1. **Conjunto base:** Compuesto por las 185 imágenes originales de hormigas con sus anotaciones correspondientes, estableciendo la línea base de rendimiento.
2. **Conjunto base adicional:** Conjunto base con adición de 202 imágenes para lograr un total de 372.
3. **Conjunto base aumentado:** Implementación de técnicas de aumento de datos (rotaciones y zoom) sobre el conjunto base adicional, incrementando la diversidad de los datos de entrenamiento.
4. **Conjunto base aumentado escala de grises:** implementación de un modelo con escala de grises a partir del conjunto base aumentado.
5. **Conjunto final:** Implementación del modelo final con 6.388 imágenes y 13.894 cargas con hormigas anotadas.

Esta estrategia de experimentación con diferentes configuraciones del conjunto de datos, manteniendo constantes los hiperparámetros de entrenamiento

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

(detallados en la Tabla 2), nos permitió evaluar sistemáticamente la influencia de la preparación de datos en el rendimiento final del modelo.

El entrenamiento se ejecutó utilizando infraestructura de computación acelerada por GPU(Graphic Processing Unit), o placa de vídeo como comúnmente es conocida. Específicamente, empleamos una NVIDIA GeForce RTX 3080 con 12GB de memoria VRAM en entorno local, complementada con sesiones en Google Colab utilizando GPUs NVIDIA Tesla T4 para experimentos paralelos. Esta implementación aprovecha la arquitectura CUDA y las optimizaciones específicas para redes neuronales convolucionales disponibles en estas unidades. El framework PyTorch (versión 2.0) fue seleccionado como entorno de desarrollo por su flexibilidad, eficiencia computacional y soporte optimizado para arquitecturas YOLO [42].

Parámetro	Valor
Épocas	100
Tamaño de batch	8
Tamaño de imagen	640 [px]

Tabla 2 - Parámetros de entrenamiento

Como resultado del entrenamiento, obtenemos diversos gráficos de evolución (Figura 18) que reflejan el progreso y rendimiento del modelo a lo largo de las épocas de entrenamiento. Estas visualizaciones proporcionan información crucial sobre cómo el modelo mejora su capacidad para detectar y segmentar hormigas y sus cargas con el avance del entrenamiento.

Las gráficas están organizadas en dos categorías principales: métricas de detección de cuadros delimitadores (B) en la columna izquierda, y métricas de segmentación de máscaras (M) en la columna derecha. En ambos casos, podemos observar una tendencia ascendente clara que indica el aprendizaje progresivo del modelo a medida que procesa más datos de entrenamiento.

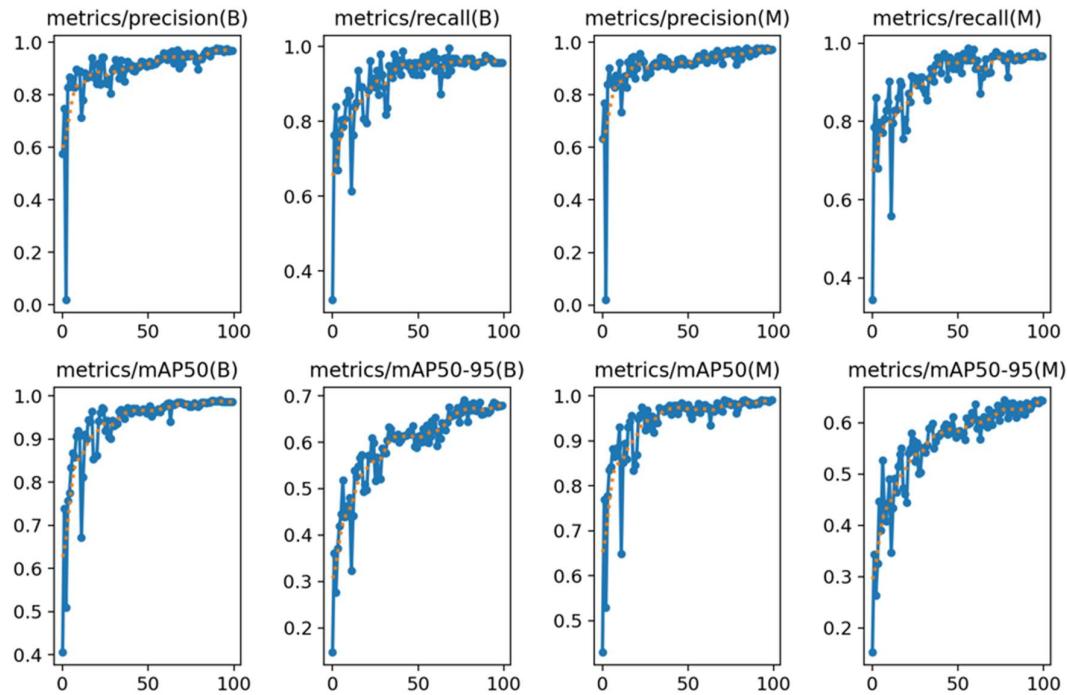


Figura 18 - Métricas del modelo en detección y segmentación entrenado en 100 épocas.

Específicamente, las métricas visualizadas representan:

- Precisión (Precision): La primera fila muestra cómo el modelo mejora su capacidad para realizar predicciones correctas. Esta métrica responde a la pregunta: "De todos los objetos que el modelo identificó como hormigas o cargas, ¿qué porcentaje realmente lo son?". Valores cercanos a 1.0 indican que el modelo genera muy pocos falsos positivos, lo que en nuestro caso significa que raramente confunde otros elementos del entorno con cargas de la hormiga. [43] [44]
- Exhausitividad (Recall): La segunda fila ilustra el progreso en la capacidad del modelo para encontrar todos los objetos relevantes en las imágenes. Responde a la pregunta: "¿qué proporción de hormigas y cargas realmente presentes en las imágenes logra detectar el modelo?" El aumento gradual hasta valores cercanos a 0.95 sugiere que el modelo ha desarrollado una

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

excelente capacidad para detectar casi todas las hormigas presentes, incluso aquellas parcialmente visibles o en posiciones difíciles [43] [44].

- mAP50: La tercera fila representa la precisión media con un criterio de coincidencia relativamente permisivo (50% de solapamiento entre la predicción y la realidad). Es como evaluar al modelo con un estándar básico, donde se considera correcto si la ubicación predicha coincide al menos en la mitad con la ubicación real. El rápido incremento inicial seguido de estabilización cerca de 1.0 indica que el modelo alcanzó excelente rendimiento en este nivel de exigencia [45].
- mAP50-95: La última fila presenta la evaluación más rigurosa, promediando el rendimiento a través de múltiples niveles de exigencia. Equivale a evaluar al modelo tanto con criterios básicos como con estándares muy estrictos donde se requiere casi perfecta coincidencia entre predicción y realidad. Los valores más bajos (aproximadamente 0.65) en comparación con otras métricas son esperables, ya que representan un desafío significativamente mayor [45].

La comparación entre columnas revela un rendimiento ligeramente superior en segmentación (M) respecto a detección de cajas (B), lo que sugiere que el modelo ha desarrollado una comprensión más precisa de la forma exacta de las hormigas y sus cargas que de su mera ubicación aproximada. Asimismo, las fluctuaciones iniciales seguidas de estabilización progresiva son indicativas de un proceso de aprendizaje exitoso, donde el modelo inicialmente explora diferentes configuraciones antes de consolidar su conocimiento.

### 3.2.6 Evaluación del modelo

En este capítulo establecemos el marco metodológico para evaluar el desempeño de nuestros modelos entrenados, basándonos en métricas estándar de la comunidad de visión por computadora [37]. Estas métricas nos proporcionarán una evaluación objetiva y cuantitativa de la capacidad del modelo para detectar y

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

segmentar hormigas y sus cargas, permitiendo comparaciones rigurosas entre diferentes configuraciones de entrenamiento [46] [47].

### 3.2.6.1 Métricas fundamentales de evaluación

#### Exhaustividad (Recall)

La exhaustividad representa la capacidad del modelo para identificar todos los objetos relevantes presentes en las imágenes. Esta métrica responde a la pregunta: "¿qué proporción de hormigas y cargas realmente presentes logra detectar nuestro modelo?" Se calcula mediante la ecuación:

$$R = \frac{TP}{(TP + FN)}$$

Donde:

- **TP (Verdaderos Positivos):** Número de cargas correctamente detectadas por el modelo
- **FN (Falsos Negativos):** Número de cargas presentes en la imagen que el modelo no logró detectar

El denominador ( $TP + FN$ ) representa el total de objetos reales existentes en el conjunto de datos de evaluación. Un valor de recall de 0.95, por ejemplo, indica que el modelo detecta exitosamente el 95% de todas las hormigas y cargas presentes, omitiendo únicamente el 5% restante [47] [48].

#### Precisión (Precision)

La precisión evalúa la calidad de las detecciones realizadas por el modelo, cuantificando la proporción de predicciones correctas entre todas las detecciones reportadas. Se define como:

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

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

Donde:

- **FP (Falsos Positivos):** Número de detecciones incorrectas (elementos del fondo identificados erróneamente como hormigas o cargas, o detecciones con localización imprecisa)

El denominador ( $TP + FP$ ) corresponde al total de detecciones realizadas por el modelo. Una precisión de 0.92 significa que el 92% de las detecciones reportadas son correctas, mientras que el 8% restante son falsos positivos que deben ser filtrados [47] [48].

### Precisión Media (Average Precision - AP)

La Precisión Media integra el rendimiento del modelo a través de todos los niveles de confianza, proporcionando una métrica única y robusta para la evaluación. Geométricamente, representa el área bajo la curva Precisión-Recall, donde el eje horizontal representa recall y el eje vertical representa precisión:

$$AP = \int_0^1 P(R) dR$$

### Variantes de mAP según umbrales de IoU

#### mAP@0.5 (mAP a IoU 0.5)

Esta métrica evalúa el rendimiento de la AP a través de todas las clases de un dataset específico, utilizando un umbral de Intersección sobre Unión (IoU) de 0.5 para clasificar detecciones como verdaderos positivos. El IoU se calcula como:

$$IoU = \frac{\text{Área intersección}}{\text{Área unión}}$$

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

Un umbral de 0.5 significa que una detección se considera correcta si la superposición entre la predicción y el objeto real (Ver figura 19) cubre al menos el 50% del área total ocupada por ambos. Esta métrica es más permisiva y típicamente arroja valores más altos, siendo útil para evaluar la capacidad básica de localización del modelo [45] [47] [48].

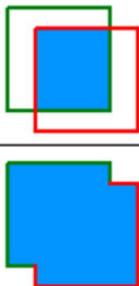
$$IoU = \frac{\text{Área intersección}}{\text{Área unión}} = \frac{\text{Área del cuadro azul}}{\text{Área del cuadro azul} + \text{Área del cuadro rojo} - \text{Área del cuadro azul}}$$


Figura 19 - Intersección sobre unión [46].

La ecuación de mAP estará dada por:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$$

Donde N será la cantidad de clases dentro del dataset

#### mAP@0.5:0.95 (mAP Multi-umbral)

Esta métrica más rigurosa promedia la AP a través de múltiples umbrales de IoU desde 0.5 hasta 0.95 con incrementos de 0.05, siguiendo el protocolo estándar de MS COCO [37]

Esta formulación exige diferentes niveles de precisión espacial: desde detecciones aproximadas ( $IoU = 0.5$ ) hasta localizaciones extremadamente precisas ( $IoU = 0.95$ ). Los valores resultantes son típicamente menores que mAP@0.5, pero proporcionan una evaluación más exhaustiva de la capacidad del modelo para localizar objetos con alta precisión [47] [48].

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

YOLO genera automáticamente matrices de confusión post-entrenamiento que tabulan las predicciones versus las etiquetas reales para cada clase. Esta representación matricial facilita la extracción directa de los valores TP, FP y FN necesarios para el cálculo de las métricas:

Elementos diagonales: Verdaderos positivos para cada clase

Suma de fila menos elemento diagonal: Falsos negativos para esa clase

Suma de columna menos elemento diagonal: Falsos positivos para esa clase.

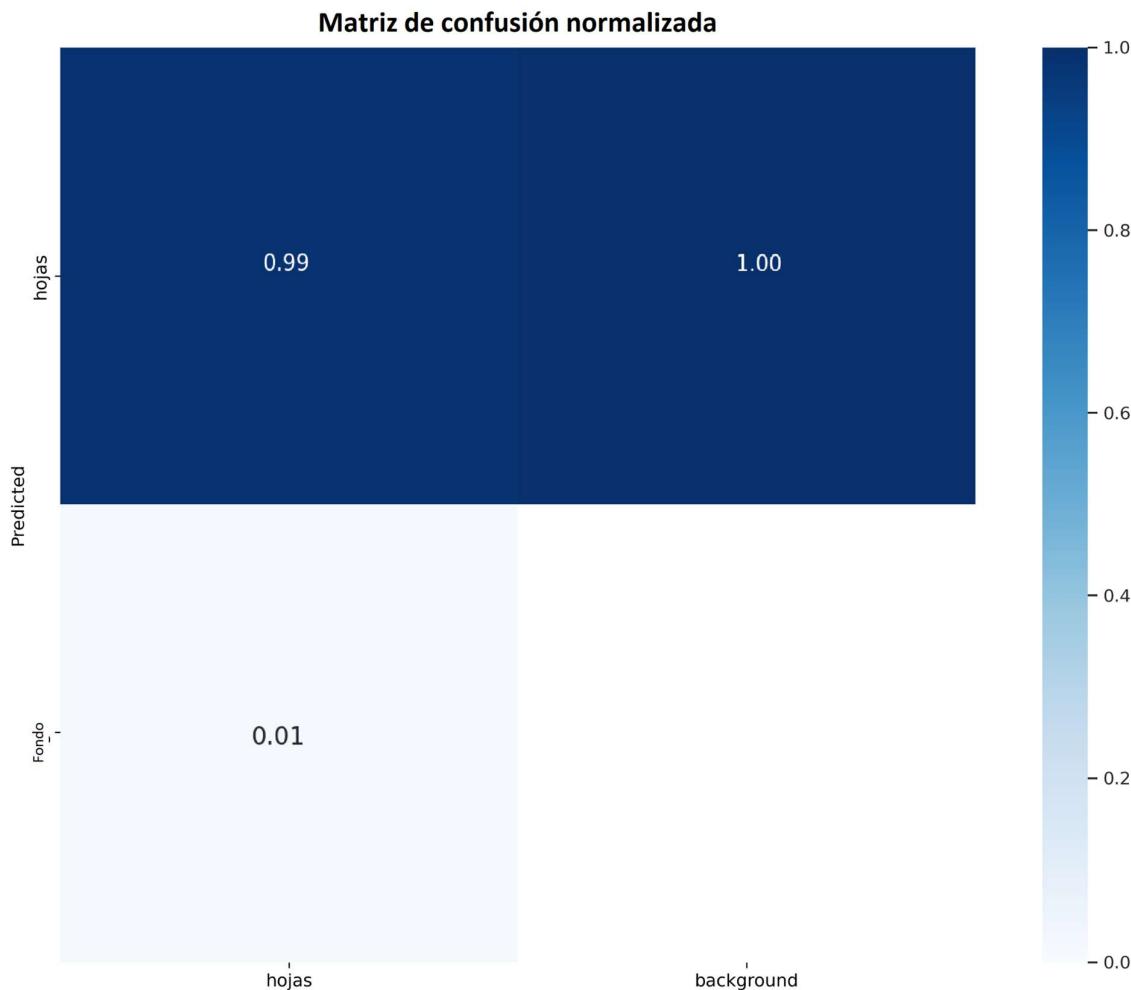


Figura 20 - Matriz de confusión del Conjunto base adicional

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

### 3.2.6.2 Métricas de los modelos

Se desarrollaron cinco modelos YOLOv8 diferentes, dos de los cuales fueron entrenados en PC local, dos en la plataforma Google Colab y el último siendo entrenado en Roboflow. Los modelos locales utilizaron conjuntos de datos de 185 y 372 imágenes respectivamente, mientras que los modelos desarrollados en Roboflow emplearon un dataset de 770 imágenes, siendo uno entrenado con imágenes a color y otro con imágenes en escala de grises, tal como se describió previamente.

La evaluación del rendimiento de estos modelos se realizó mediante un conjunto de métricas estándar ampliamente reconocidas en la literatura de detección de objetos [46] [47] [48]. Estas métricas permiten cuantificar la precisión del modelo al comparar sus predicciones con las anotaciones de referencia (ground truth), empleando imágenes que no fueron utilizadas durante el proceso de entrenamiento [90]. Usamos como imágenes de prueba aquellas que forman parte del conjunto test del dataset *Conjunto base aumentado*. Para el conjunto final aplicamos como test 812 imágenes que fueron generadas por Roboflow. La selección de estas métricas de evaluación se fundamenta en su capacidad para proporcionar una evaluación integral del rendimiento del modelo, considerando tanto la exactitud en la detección como la localización de los objetos.

Además de las métricas ya explicadas anteriormente, a esta evaluación también agregamos **F1-Score**, que representa la media armónica entre precisión y recall, proporcionando una métrica balanceada que considera tanto falsos positivos como falsos negativos [91] [92] [93]. Su fórmula es  $F1 = 2 \times (\text{Precisión} \times \text{Recall}) / (\text{Precisión} + \text{Recall})$ , siendo especialmente útil para datasets desbalanceados [93].

A continuación, se presentan los resultados obtenidos para cada modelo:

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

Modelo	mAP50	mAP50-95	Precision	Recall	F1
Conjunto final	0.83	0.54	0.83	0.77	0.79
Conjunto base	0.35	0.209	0.33	0.57	0.42
Conjunto base adicional	0.55	0.31	0.67	0.47	0.55
Conjunto base aumentado	0.97	0.77	0.92	0.97	0.94
Conjunto base aumentado escala de grises	0.97	0.75	0.97	0.93	0.95

Tabla 3 - Tabla comparativa con métricas de los modelos entrenados.

El análisis de las métricas obtenidas revela que el modelo entrenado con el dataset conjunto base aumentado presentó el mejor rendimiento general, destacándose consistentemente en las métricas evaluadas (Figura 21). Aunque sus métricas son superiores a los demás modelos, tiene una cantidad de imágenes inferior al conjunto final, a partir de esto, utilizamos el modelo basado en el dataset Conjunto final, ya que tiene la mayor cantidad de imágenes y métricas consistentes en la detección.

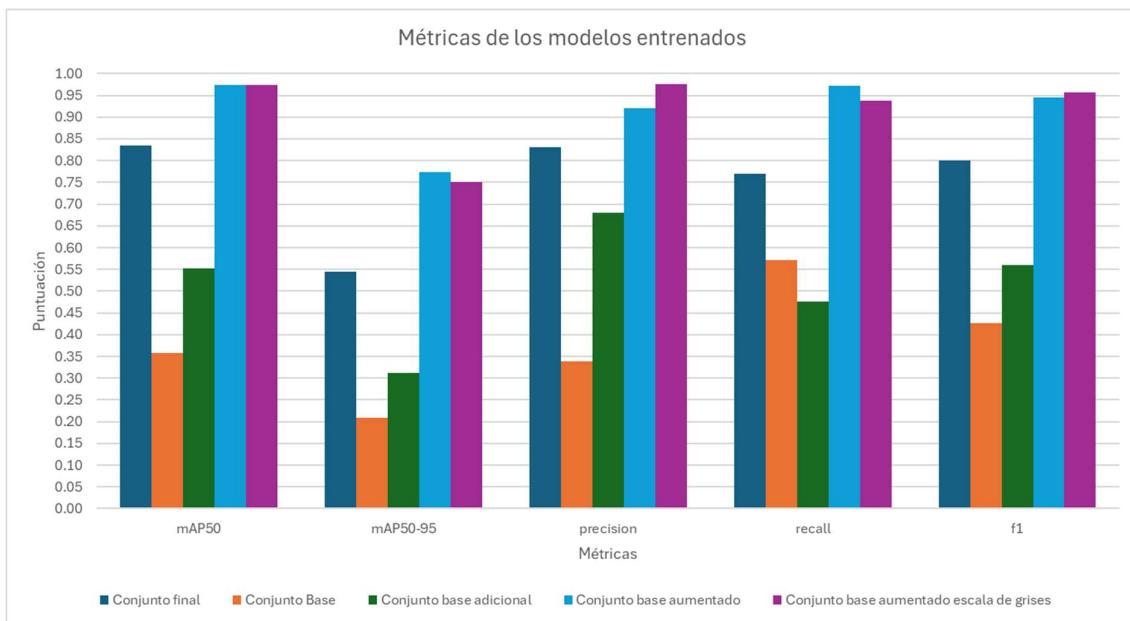


Figura 21 - Gráfica de comparación de métricas de los modelos entrenados.

Es importante destacar que, el modelo seleccionado es el más rápido entre todos, destacándose con un tiempo de inferencia de 1,46 [ms], en la tabla 4 se visualiza la comparación con los otros modelos entrenados.

Modelo	Tiempo de Inferencia [ms]
Conjunto final	1.46
Conjunto Base	9.08
Conjunto base adicional	10.36
Conjunto base aumentado	8.08
Conjunto base aumentado escala de grises	14.8

Tabla 4 - Comparación de tiempos de inferencia.

### 3.2.7 Seguimiento multi-objeto

En la siguiente sección se analizará en detalle el método utilizado para el seguimiento de la carga de hormiga a medida que progresó a través de los frames

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

del video. Aunque el seguimiento multi-objeto (Multi-Object Tracking, MOT) ha sido aplicado exitosamente en diversos campos y existen algoritmos con desarrollo muy avanzado, el objeto a detectar en nuestro caso presenta características particulares que complejizan el proceso de seguimiento: la carga transportada por hormigas no es uniforme y el ángulo de visualización varía constantemente durante el transporte.

### 3.2.7.1 Método de seguimiento

El seguimiento multi-objeto en secuencias de vídeo constituye un área de investigación activa en visión por computadora, siendo fundamental para diversas aplicaciones como vigilancia, conducción autónoma y análisis de comportamiento [50]. Los enfoques actuales se basan predominantemente en el paradigma "tracking-by-detection", donde se realiza la detección de objetos en cada frame seguida de asociación temporal para mantener las identidades de los objetos a través del tiempo.

Entre los algoritmos estado del arte se encuentran SORT (Simple Online and Realtime Tracking) [50], DeepSORT [51], ByteTrack [52] y BoT-SORT [53], cada uno con diferentes enfoques para abordar los desafíos inherentes al seguimiento multi-objeto. SORT utiliza el filtro de Kalman y el algoritmo húngaro para la asociación de datos, logrando un balance entre simplicidad y eficiencia computacional [50]. DeepSORT extiende SORT incorporando información de apariencia mediante descriptores profundos, mejorando significativamente el manejo de occlusiones y reduciendo los cambios de identidad [51].

ByteTrack introduce una estrategia novedosa que utiliza todas las cajas de detección, incluyendo aquellas con puntuaciones bajas, para recuperar objetos verdaderos que podrían perderse con umbrales de confianza estrictos [52]. Por su parte, BoT-SORT combina las ventajas de la información de movimiento y apariencia, junto con compensación de movimiento de cámara y un vector de estado del filtro de Kalman más preciso [53].

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

### 3.2.7.2 Desafíos específicos del seguimiento de carga de hormigas

El primer acercamiento al seguimiento de objetos fue implementado utilizando estos algoritmos mencionados, los cuales no proporcionaron resultados óptimos debido a las características específicas de nuestro objeto de interés. La principal limitación identificada fue la falta de uniformidad de la carga transportada por las hormigas, lo que provocaba pérdidas de detección intermitentes. Cuando el objeto era detectado nuevamente después de varios frames sin detección, los algoritmos estándar le asignaban un nuevo identificador, perdiendo así la continuidad del seguimiento del objeto original.

Este comportamiento genera inconsistencias significativas en las trayectorias, fenómeno bien documentado en la literatura como uno de los principales desafíos en MOT. Las occlusiones, cambios en la apariencia del objeto y movimientos rápidos son factores que contribuyen a la fragmentación de trayectorias y al incremento en los cambios de identidad [54] [55].

Debido a las limitaciones observadas con los algoritmos estándar, se optó por implementar un algoritmo personalizado que tiene la capacidad de relacionar las detecciones entre frames de manera más robusta para nuestro caso específico. Esta decisión se fundamenta en la necesidad de adaptar el proceso de asociación de datos a las características particulares de la carga de hormigas, proporcionando al usuario la posibilidad de modificar parámetros de configuración para cada caso de aplicación.

El enfoque desarrollado incorpora estrategias de asociación más flexibles que consideran no solo la proximidad espacial y temporal, sino también características específicas del dominio que permiten mantener la continuidad del seguimiento incluso cuando las detecciones son intermitentes. Esta aproximación personalizada permite un control más granular sobre los parámetros de seguimiento, aspecto crucial para optimizar el rendimiento en aplicaciones específicas como el análisis de comportamiento de hormigas.

Debido a las limitaciones observadas con los algoritmos estándar, se optó por implementar un algoritmo personalizado que tiene la capacidad de relacionar las

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

detecciones entre frames para nuestro caso específico. Esta decisión se fundamenta en la necesidad de adaptar el proceso de asociación de datos a las características particulares de la carga de hormigas, proporcionando al usuario la posibilidad de modificar parámetros de configuración en la interfaz para cada caso de aplicación.

El enfoque desarrollado incorpora estrategias de asociación más flexibles que consideran no solo la proximidad espacial y temporal, sino también características específicas del dominio que permiten mantener la continuidad del seguimiento incluso cuando las detecciones son intermitentes. Esta aproximación personalizada permite un control más granular sobre los parámetros de seguimiento, aspecto crucial para optimizar el rendimiento en aplicaciones específicas como el análisis de comportamiento de hormigas.

### 3.2.7.3 Algoritmo de seguimiento de objetos

El algoritmo de seguimiento desarrollado implementa un enfoque basado en distancia euclíadiana para mantener la continuidad de las trayectorias de objetos a través de secuencias de vídeo. El sistema opera bajo el paradigma "tracking-by-detection", donde cada frame es procesado independientemente para detectar objetos, seguido de un proceso de asociación que vincula las nuevas detecciones con trayectorias existentes utilizando criterios de proximidad espacial [56] [57].

El proceso inicia con la extracción de características geométricas básicas de cada detección en el frame actual. Para cada bounding box detectado, se calculan las coordenadas del centroide ( $x_{\text{med}}$ ,  $y_{\text{med}}$ ) y las dimensiones ( $\Delta x$ ,  $\Delta y$ ), información que constituye la base para el posterior proceso de asociación [56]. A partir de estas dimensiones, se computa la distancia hipotenusa  $hp = \sqrt{(\Delta x^2 + \Delta y^2)}$ , que actúa como una medida representativa del tamaño del objeto y sirve para definir regiones de búsqueda adaptativas alrededor de cada detección. Esta métrica dimensional permite que el algoritmo adapte automáticamente las tolerancias de asociación según el tamaño de cada objeto individual, proporcionando mayor flexibilidad para manejar objetos de diferentes escalas dentro de la misma secuencia.

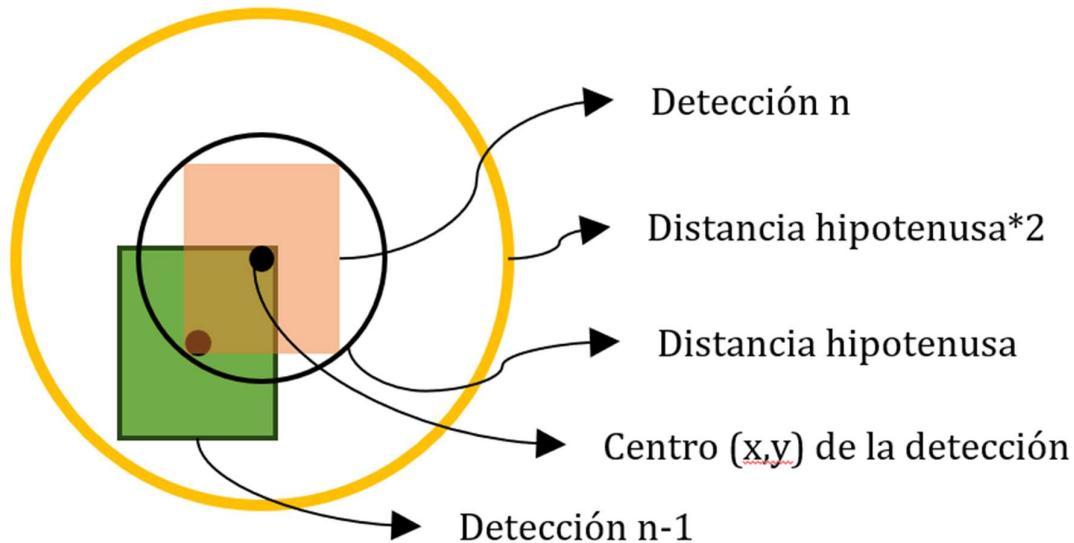


Figura 22 - Visualización de una detección y comparación positiva.

La estrategia de asociación implementada utiliza un enfoque de múltiples niveles de tolerancia espacial que permite manejar diferentes grados de variabilidad en el movimiento de los objetos. En el primer nivel de asociación, se define un área de búsqueda circular con dimensiones  $\pm k \times hp$  alrededor de la posición del centroide de cada objeto previamente rastreado, donde  $k$  representa un factor de escalamiento configurable. Esta región primaria está diseñada para capturar movimientos normales del objeto bajo condiciones ideales de seguimiento. La condición de asociación en este nivel se expresa matemáticamente como  $(-k \times hp) < (x_c - x_{med}) < (k \times hp) \wedge (-k \times hp) < (y_c - y_{med}) < (k \times hp)$ , donde  $(x_c, y_c)$  representa el centroide del objeto rastreado en la aparición anterior y  $(x_{med}, y_{med})$  son las coordenadas de la detección actual.

La validación del algoritmo fue realizada mediante métricas tales como TSE y TCT, explicados profundamente en la sección 3.2.10, que son aquellas que determinaron la puntuación de nuestro seguimiento en cada intervalo evaluado.

Adicionalmente y a fin fortalecer la validación del algoritmo, planteamos realizar pruebas con un vídeo corto de 2 minutos, en el cual pasan 35 cargas hacia la boca del hormiguero, cambiando los parámetros de configuración, explicados detalladamente en la sección 3.2.9.1, obteniéndose los siguientes resultados.

	UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica		3 - Materiales y Métodos

Parámetro	#1	#2	#3	#4	#5	#6
FPS	30	30	30	30	30	30
Distancia	2	1	2	1	1	4
Frames						
Frames	15	3	5	5	2	20
aparición						
Confianza	0,6	0,5	0,4	0,4	0,4	0,6
Cantidad de	10	2	2	2	1	18
apariciones						
Tiempo de	1	1	1	1	1	1
guardado						
#Hojas contadas	9	25	29	27	42	1

Tabla 5 - Comprobación de comportamiento ante el cambio de parámetros.

En la tabla 5 se observa el comportamiento del conteo ante la variación de parámetros de configuración. Cuanto más permisivos son estos parámetros, mayor es la cantidad de detecciones. A la inversa, cuanto más estricto son, menor es la cantidad de detecciones.

En situaciones donde no se encuentra una asociación válida en el área primaria, el algoritmo expande la búsqueda a una región secundaria con tolerancia  $2k \times hp$ . Esta expansión permite recuperar objetos que han experimentado movimientos más pronunciados o que han sufrido interrupciones temporales en su detección debido a occlusiones parciales o cambios en las condiciones de imagen.

El algoritmo incorpora varios mecanismos para manejar casos especiales que frecuentemente aparecen en aplicaciones de seguimiento multi-objeto. La inicialización de nuevos objetos ocurre cuando una detección no puede ser asociada con ninguna trayectoria existente dentro de las regiones de búsqueda definidas y cuando la posición de la detección se encuentra dentro de regiones espaciales de

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

interés predefinidas. En estos casos, se crea una nueva instancia de objeto con un identificador único, estableciendo el inicio de una nueva trayectoria de seguimiento. Para prevenir la creación de múltiples tracks para el mismo objeto, se implementa un sistema de verificación de duplicados que evalúa la proximidad espacial entre la nueva detección y las posiciones de objetos existentes utilizando una distancia mínima configurable, detallado en la sección 3.2.7.4.

La gestión temporal del seguimiento incluye consideraciones sobre la continuidad de las trayectorias a través de gaps temporales. El algoritmo evalúa la diferencia entre el frame actual y el último frame donde cada objeto fue detectado, permitiendo asociaciones incluso cuando han transcurrido varios frames sin detección, siempre que la diferencia temporal sea positiva. Esta funcionalidad es particularmente importante en escenarios donde los objetos pueden experimentar occlusiones temporales o donde la calidad de detección puede variar entre frames consecutivos debido a cambios en la iluminación, movimientos de cámara o características inherentes del objeto.

Un aspecto distintivo del algoritmo es su capacidad de adaptación a las características específicas del dominio de aplicación. Los parámetros de tolerancia espacial, umbrales de distancia y criterios de inicialización pueden ser ajustados dinámicamente según las características observadas del movimiento de los objetos y las condiciones específicas del experimento. Esta flexibilidad paramétrica permite optimizar el rendimiento del seguimiento para diferentes tipos de objetos y patrones de movimiento. La simplicidad del enfoque basado en distancia euclíadiana también facilita la interpretación y depuración del comportamiento del algoritmo, aspectos cruciales durante el desarrollo y validación de sistemas de seguimiento personalizados.

El sistema mantiene un registro histórico de las apariciones de cada objeto a través de una estructura de datos que almacena información sobre posición, área, frame de detección y características geométricas relevantes.

La visualización del proceso de seguimiento se facilita mediante la superposición de indicadores gráficos en el frame anotado, incluyendo círculos que

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

representan las regiones de búsqueda y marcadores que identifican los centroides de los objetos detectados, proporcionando retroalimentación visual inmediata sobre el comportamiento del algoritmo durante la ejecución

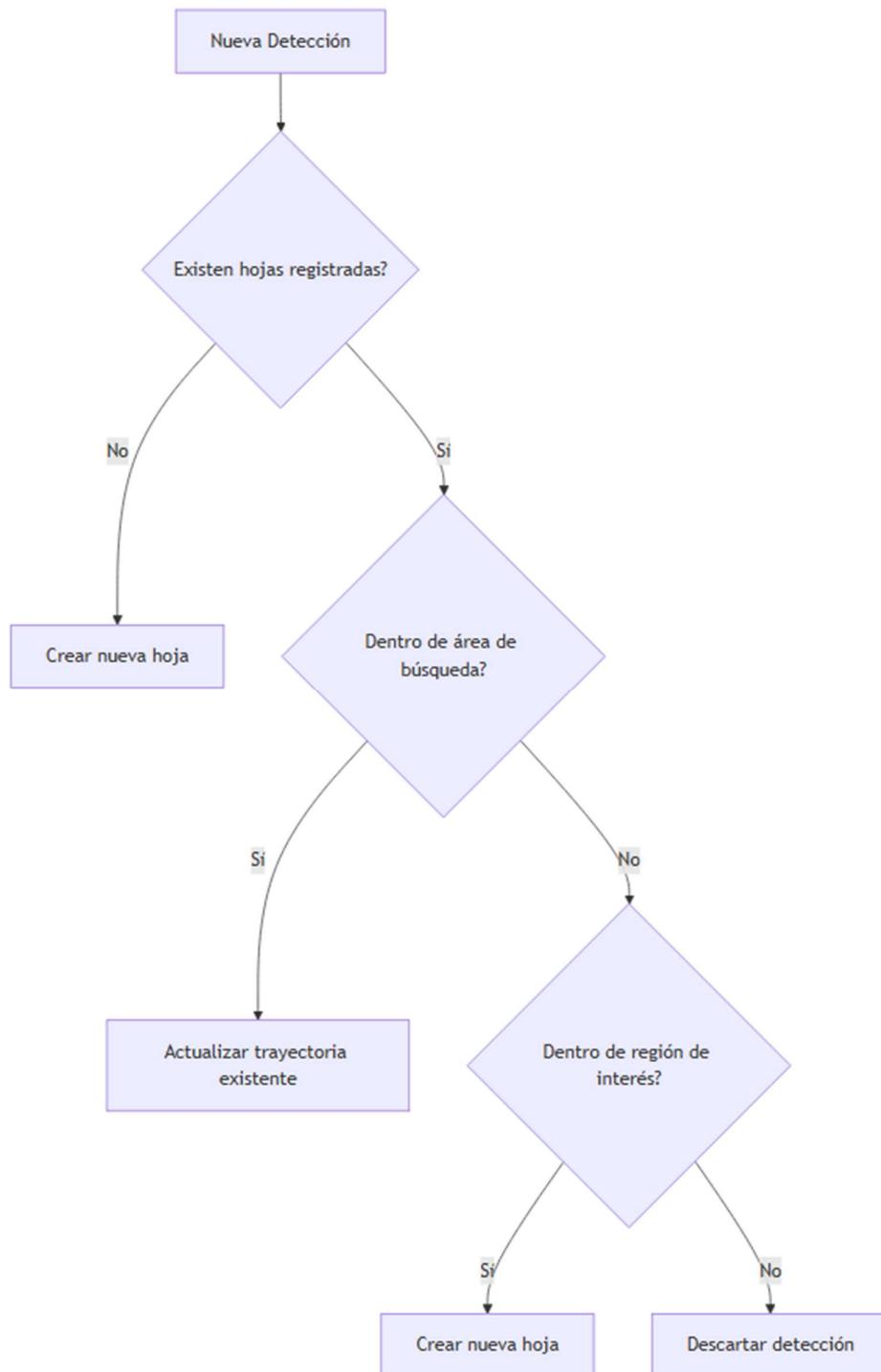


Figura 23 - Diagrama de flujo del algoritmo de seguimiento.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

### 3.2.7.4 Eliminación de Duplicados y falsas detecciones

La eliminación de duplicados y falsas detecciones constituye una etapa de post-procesamiento en sistemas de seguimiento multi-objeto, especialmente cuando los algoritmos enfrentan situaciones donde objetos se encuentran en proximidad espacial o experimentan cruzamientos entre sus trayectorias. Este problema es particularmente prevalente en escenarios de seguimiento denso, donde la asociación de datos basada únicamente en distancia euclíadiana puede generar ambigüedades significativas y resultar en asignaciones incorrectas de identificadores [58] [59].

El desarrollo de métodos para la detección y eliminación de duplicados representa un desafío fundamental en visión por computadora, ya que los falsos positivos y las detecciones redundantes pueden comprometer significativamente la calidad del análisis posterior. En sistemas de seguimiento que operan en tiempo real, la acumulación de errores de asociación puede propagarse a través de múltiples frames, aumentando el problema y resultando en fragmentación de trayectorias o fusión incorrecta de tracks independientes.

El algoritmo implementado aborda este problema mediante un enfoque multicriterio que evalúa tanto la proximidad espacial como la similitud temporal y geométrica entre objetos rastreados. El proceso inicia con un análisis de todas las trayectorias almacenadas en la estructura de datos global, comparando sistemáticamente cada par de objetos para identificar potenciales duplicados. La evaluación de proximidad espacial se realiza calculando la distancia euclíadiana entre los centroides de las apariciones más recientes de cada objeto, estableciendo un umbral de 5 píxeles como criterio de proximidad crítica. Este valor umbral ha sido seleccionado empíricamente para balancear la sensibilidad de detección con la especificidad, minimizando tanto falsos positivos como falsos negativos en la identificación de duplicados.

El componente temporal del algoritmo introduce una dimensión adicional al evaluar la sincronización entre apariciones de diferentes objetos. El sistema verifica si las detecciones ocurren en frames consecutivos o dentro de una ventana temporal

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

limitada, típicamente 5 frames, lo que permite identificar situaciones donde el mismo objeto físico ha sido erróneamente asociado con múltiples identificadores debido a variaciones en la calidad de detección o movimientos abruptos. Esta evaluación temporal es fundamental para distinguir entre objetos genuinamente separados que temporalmente ocupan posiciones cercanas y duplicados reales que representan la misma entidad física.

El algoritmo incorpora criterios geométricos adicionales mediante la comparación de áreas de los bounding boxes, evaluando si la diferencia relativa entre las áreas de dos objetos candidatos excede un umbral del 30%. Esta verificación geométrica proporciona una capa adicional de validación, ya que duplicados genuinos típicamente exhiben características dimensionales similares, mientras que objetos distintos pueden presentar variaciones significativas en su tamaño aparente [61]. La inclusión de múltiples criterios de evaluación reduce sustancialmente la probabilidad de clasificaciones erróneas, especialmente en escenarios donde la proximidad espacial por sí sola podría resultar en decisiones ambiguas [62].

El proceso de toma de decisiones para la eliminación de duplicados emplea una estrategia conservadora que prioriza la preservación de información sobre la eliminación agresiva. Cuando se identifica un par de objetos como duplicados potenciales, el algoritmo compara el historial de apariciones de ambos, manteniendo el objeto con mayor cantidad de detecciones registradas o, en caso de empate, el que apareció primero cronológicamente. Esta estrategia asegura que se preserve la trayectoria más completa y establecida, minimizando la pérdida de información valiosa sobre el comportamiento del objeto.

La gestión de objetos perdidos o falsas detecciones representa otro aspecto crucial del sistema de filtrado. El algoritmo evalúa continuamente la diferencia temporal entre la última aparición registrada de cada objeto y el frame actual de procesamiento. Cuando esta diferencia excede un umbral predefinido de 15 frames por defecto, se activa un proceso de evaluación que considera múltiples factores antes de decidir la eliminación del objeto. Estos factores incluyen la cantidad total de apariciones registradas, la consistencia de la trayectoria, y la dirección del movimiento

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

del objeto. Esta evaluación previene la eliminación prematura de objetos que pueden estar experimentando oclusiones temporales o que han salido momentáneamente del campo de visión debido a movimientos de cámara o cambios en las condiciones de iluminación.

La implementación del algoritmo utiliza estructuras de datos que permiten comparaciones rápidas entre múltiples objetos sin comprometer significativamente el rendimiento computacional manteniendo índices ordenados de objetos por identificador, facilitando búsquedas sistemáticas y evitando comparaciones redundantes.

La efectividad del sistema de filtrado ha demostrado ser particularmente valiosa en situaciones problemáticas como cruzamientos de trayectorias, donde objetos independientes pueden aproximarse lo suficiente como para generar confusión en algoritmos de asociación más simples. Al incorporar múltiples criterios de evaluación y estrategias conservadoras de eliminación, el algoritmo mantiene la integridad de las trayectorias individuales mientras reduce significativamente la incidencia de duplicados y falsas detecciones, resultando en datasets más limpios y análisis posteriores más precisos.

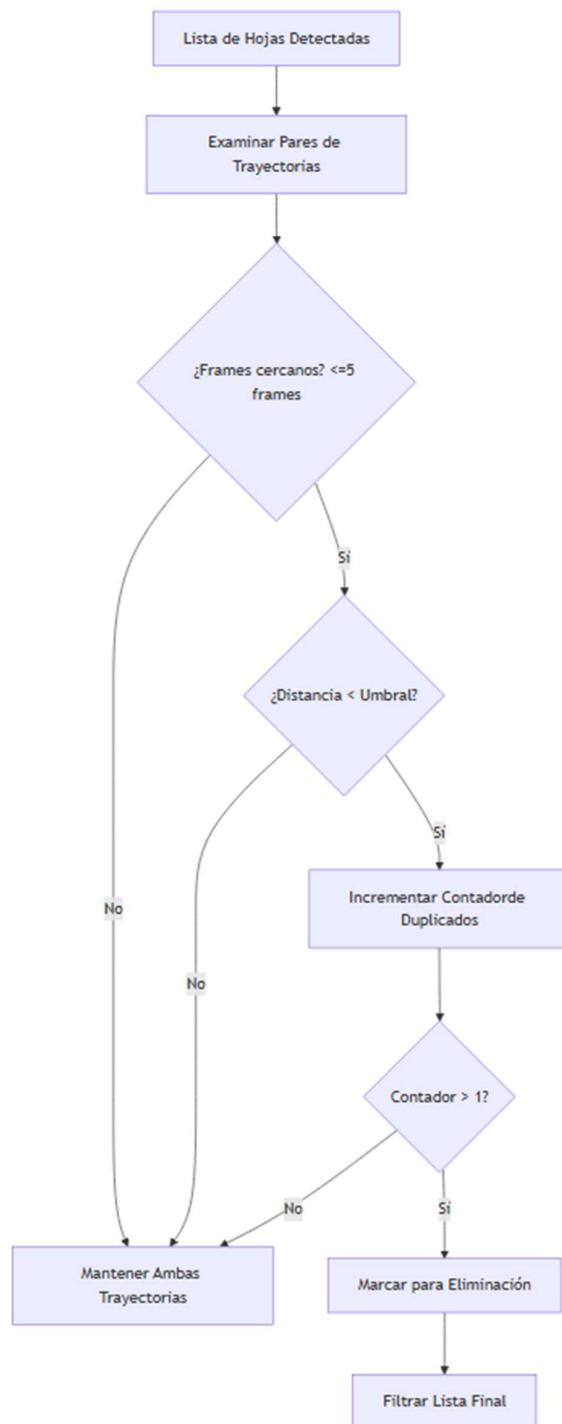


Figura 24 - Diagrama de flujo del algoritmo de filtrado.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

### 3.2.8 Cálculo del área

#### 3.2.8.1 Determinación del Contorno del Objeto

El proceso de determinación morfométrica se fundamenta en las máscaras de segmentación generadas por el modelo entrenado, que proporcionan una delimitación precisa a nivel de píxel de cada objeto detectado . A diferencia de los métodos tradicionales basados únicamente en bounding boxes rectangulares, la segmentación de instancias permite capturar la geometría irregular característica de fragmentos orgánicos transportados por hormigas [63].

Para cada detección  $i$  en el frame  $t$ , YOLOv8 genera:

$$D_i(t) = \{BB_i(t), M_i(t), conf_i(t)\}$$

Donde:

**BB<sub>i</sub>(t)** = [x\_sup, y\_sup, x\_inf, y\_inf]: Coordenadas del bounding box rectangular.

**M<sub>i</sub>(t)**: Máscara de segmentación binaria de dimensiones correspondientes al bounding box.

**conf<sub>i</sub>(t)**: Puntuación de confianza de la detección.

La máscara de segmentación  $M_i(t)$  se procesa para generar una imagen binaria donde cada píxel se clasifica como perteneciente o no al objeto de interés [8].

Utilizando la función `cv2.findContours()` de OpenCV, se extraen los contornos del objeto a partir de la imagen binaria [64]. Esta función implementa el algoritmo de Suzuki-Abe para detección de contornos, generando una secuencia ordenada de puntos que define el perímetro exterior del objeto [65].

#### 3.2.8.2 Cálculo del Área Basado en el Contorno

Una vez extraído el contorno del objeto, procedemos al cálculo de su área utilizando métodos computacionales establecidos en visión por computadora.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

OpenCV proporciona la función *contourArea()*, que implementa algoritmos geométricos para determinar el área encerrada por cualquier contorno cerrado. Esta función procesa el conjunto ordenado de puntos que definen el perímetro del objeto y calcula numéricamente el área total en unidades de píxeles cuadrados [66].

La implementación de *contourArea()* se basa en el teorema de Green, un principio del cálculo vectorial que permite convertir integrales de área en integrales de línea más simples de calcular [67]. Este enfoque matemático garantiza precisión incluso para formas irregulares y complejas, características inherentes a las cargas transportadas por hormigas cortadoras. La función maneja automáticamente casos especiales como contornos cóncavos o con múltiples componentes conexos, proporcionando resultados confiables independientemente de la complejidad morfológica del objeto [66].

La precisión del cálculo de área depende críticamente de dos factores técnicos principales. Primero, la resolución espacial del video determina el nivel de detalle con que se pueden capturar los bordes irregulares del objeto. Resoluciones más altas permiten una delimitación más precisa de características morfológicas finas, mientras que resoluciones bajas pueden introducir errores de discretización en objetos pequeños. Segundo, la calidad de la segmentación generada por el modelo entrenado influye directamente en la fidelidad del contorno extraído. Errores en la segmentación, como bordes suavizados o fragmentación de regiones, se propagan al cálculo de área, enfatizando la importancia de un modelo de detección bien entrenado.

Un aspecto fundamental a considerar en el análisis de cargas de hormigas es la variabilidad del área calculada. Las cargas orgánicas, por su naturaleza deformable, experimentan cambios de forma durante el transporte debido a factores como reorientación espacial, compresión por el agarre de la hormiga, y movimientos oscilatorios durante la locomoción [68]. Esta variabilidad natural introduce fluctuaciones en las mediciones de área entre frames consecutivos, incluso cuando se trata del mismo objeto físico.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

### 3.2.8.3 Acumulación de Áreas

El análisis longitudinal de objetos transportados requiere no solo mediciones instantáneas, sino también la acumulación sistemática de datos a lo largo de trayectorias completas. Este proceso permite caracterizar patrones de comportamiento, identificar tendencias temporales y generar estadísticas sobre las características de carga en colonias de hormigas [68]. La acumulación de áreas se fundamenta en el mantenimiento de estructuras de datos que almacenan el historial completo de cada objeto rastreado, incluyendo coordenadas espaciales y mediciones de área.

Cuando un objeto es detectado en un frame posterior, el sistema de seguimiento ejecuta algoritmos de asociación de datos para determinar si corresponde a una entidad previamente registrada. Esta verificación se basa en criterios de proximidad espacial y coherencia temporal que evalúan la plausibilidad de que dos detecciones separadas en el tiempo representen el mismo objeto físico, tal como fue explicado en la sección 3.2.7.4. Una vez confirmada la correspondencia, la nueva medición de área se incorpora al conjunto acumulado asociado con ese objeto específico, preservando la integridad del historial de seguimiento.

La acumulación de datos de área facilita análisis estadísticos que proporcionan información importante sobre el comportamiento de forrajeo de las hormigas. Se pueden calcular medidas de tendencia central como la media y mediana del área transportada, así como indicadores de variabilidad como percentiles. Estos estadísticos descriptivos permiten caracterizar cuantitativamente la consistencia del comportamiento de selección de carga. Además, el análisis temporal de estas métricas puede revelar adaptaciones comportamentales a condiciones ambientales cambiantes o estados fisiológicos de la colonia [69].

Un componente esencial del sistema de análisis morfométrico es la conversión de mediciones en píxeles a unidades físicas reales, proceso fundamental para garantizar la comparabilidad de resultados entre diferentes configuraciones experimentales y estudios. Esta calibración se realiza mediante marcadores de

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

referencia de dimensiones conocidas integrados en el escenario experimental. En nuestro protocolo, utilizamos un cuadrado de 10 mm × 10 mm ubicado en la superficie de tránsito como estándar de referencia.

El proceso de calibración implica la selección por el usuario del marcador de referencia, el cálculo de su área en píxeles mediante los mismos algoritmos utilizados para las cargas de hormigas, y la determinación del factor de conversión que relaciona píxeles cuadrados con milímetros cuadrados. Este factor se aplica posteriormente a todas las mediciones de área, transformando los datos de unidades de imagen a unidades métricas estándar. La implementación de este sistema de calibración asegura que las mediciones sean reproducibles, precisas y comparables con estudios realizados en diferentes entornos o configuraciones de cámara.

### 3.2.9 Interfaz gráfica

La interfaz implementada en el marco de la biblioteca CustomTkinter [36], representa una pestaña dentro de una interfaz gráfica diseñada para gestionar y visualizar procesos de detección y seguimiento multi-objeto. Esta pestaña está estructurada en tres secciones principales: 1)- un marco dedicado a la visualización de video, 2) - un marco para botones de control y 3)- un marco para la presentación de información textual. Cada uno de estos componentes cumple un rol específico en la interacción del usuario con el sistema. Además se incorporó un sistema para visualizar las detecciones realizadas con el modelo de YOLO que sea seleccionado por el usuario y un menú desplegable en el que se podrá elegir un filtrado de cruces y duplicados estricto.

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	<b>3 - Materiales y Métodos</b>

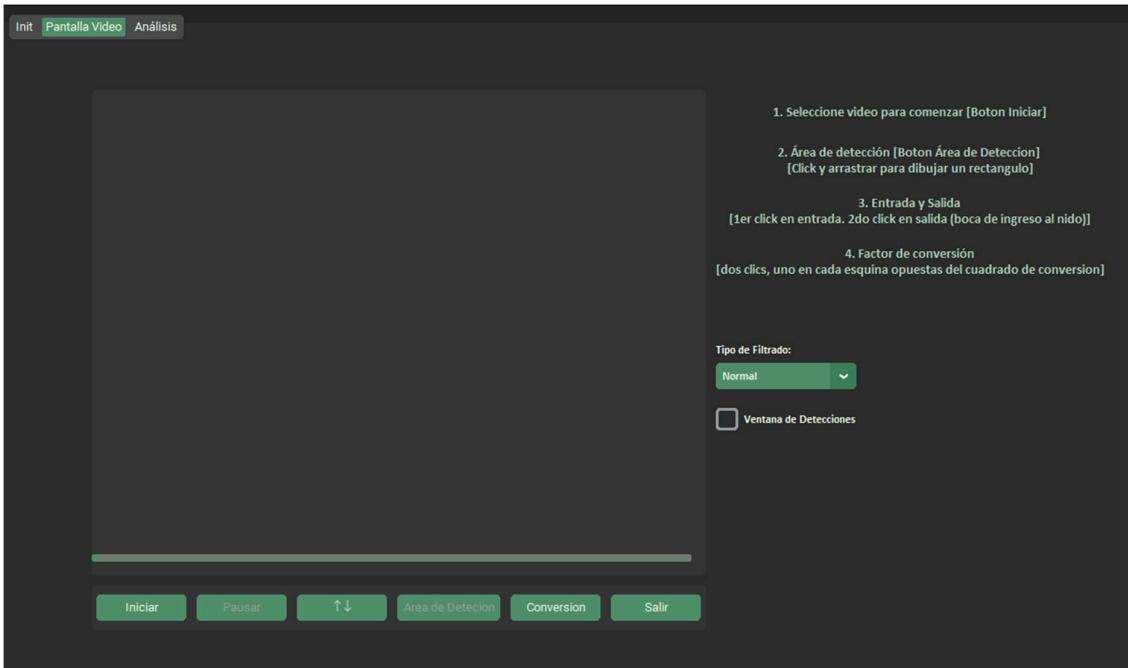


Figura 25 - Interfaz gráfica realizada con CustomTKinter (Sin reproducción de vídeo activa)

El marco de vídeo es la parte de la interfaz donde se muestra el video en tiempo real. Configurado con dimensiones que se ajustan a las dimensiones del vídeo, actúa como un lienzo dinámico que permite no solo la reproducción del video, sino también la interacción directa a través de eventos de clic y movimiento del mouse. Esta funcionalidad permite al usuario definir regiones de interés, como el área de detección o las coordenadas de entrada hacia la boca del hormiguero y salida, mediante la captura de eventos de clic y la actualización visual de rectángulos en tiempo real. Además, se incluye una barra de progreso que proporciona una representación visual del avance del procesamiento del video, mejorando la experiencia del usuario al ofrecer retroalimentación sobre el estado del sistema.

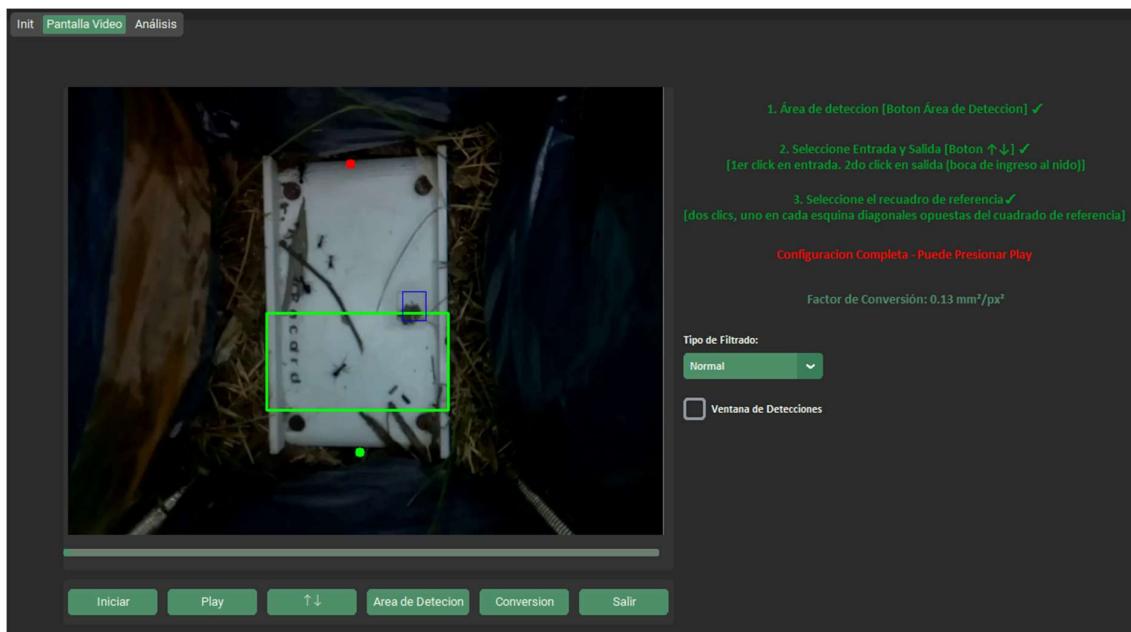


Figura 26 - Interfaz gráfica realizada con CustomTKinter (Con reproducción de vídeo activa)

El marco para botones contiene una serie de ellos que permiten al usuario controlar el flujo del sistema. El botón "Iniciar" activa el procesamiento del video, mientras que los botones "Pausar", "Área de Detec.", " $\uparrow \downarrow$ " y "Conversion" habilitan funciones específicas, como la pausa del video, la selección del área de detección, la elección de entrada y salida del hormiguero, y la conversión de unidades de medición. Estos botones están inicialmente deshabilitados y se activan secuencialmente según el estado del sistema, garantizando que el usuario siga una secuencia lógica y ordenada. El botón "Salir" permite finalizar la aplicación de manera segura. La implementación de estados dinámicos en los botones refuerza el funcionamiento de la interfaz, evitando interacciones inapropiadas en momentos inadecuados.

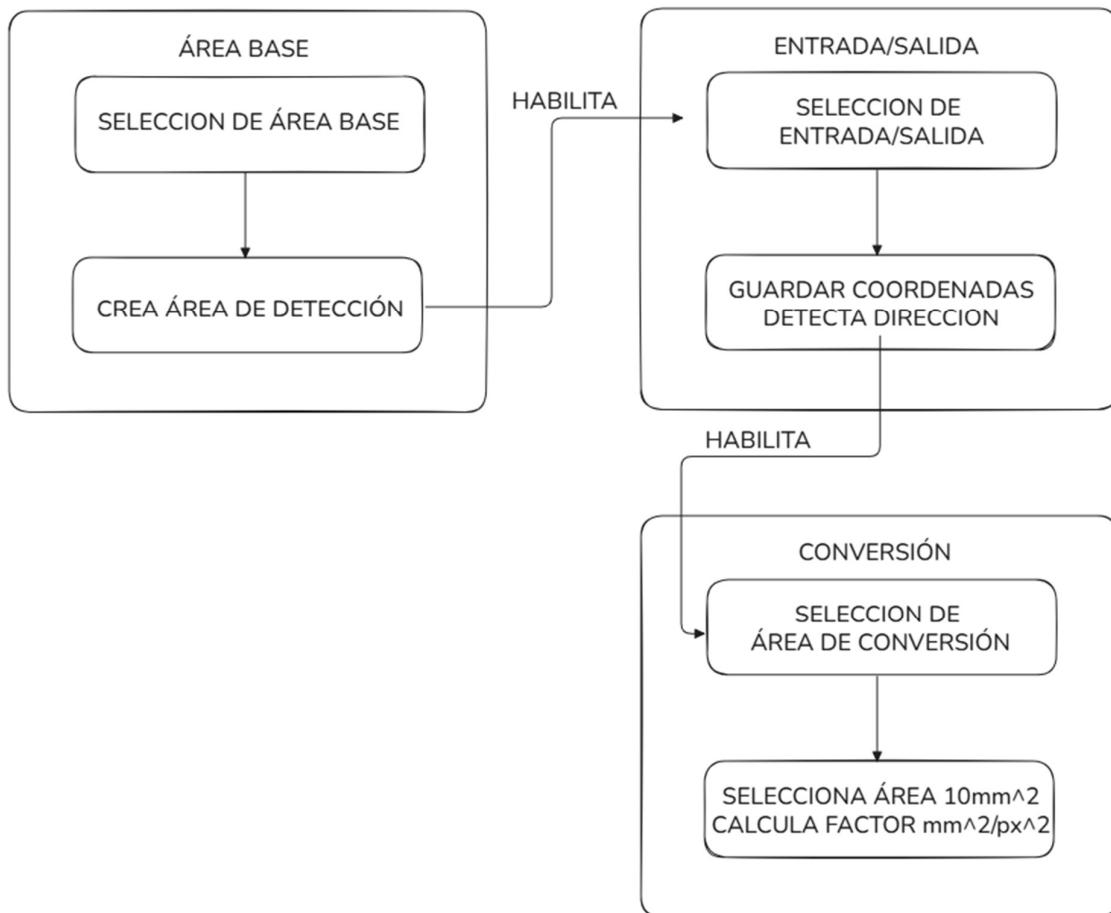


Figura 27 - Secuencia de botones.

El marco de textos actúa como un panel informativo, donde se presentan mensajes de texto que guían al usuario a través de las diferentes etapas del proceso. Estos muestran instrucciones claras y actualizaciones en tiempo real, como la confirmación de la selección de áreas o la visualización de valores calculados, así como la constante de conversión de unidades.

### 3.2.9.1 Parámetros de configuración

La interfaz del software implementado proporciona un conjunto integral de parámetros configurables que permiten la adaptación del sistema a diversos

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

escenarios operativos. Si bien estos parámetros poseen valores predeterminados que garantizan la funcionalidad básica del programa, su ajuste es altamente recomendable para optimizar el rendimiento según los requisitos específicos de cada aplicación.



Figura 28 - Pestaña de configuración de parámetros.

Los parámetros configurables se describen a continuación:

- **Fecha del Video (DD-MM-YYYY):** Este campo permite establecer la referencia temporal del video analizado. La correcta especificación de metadatos temporales es fundamental para la trazabilidad y análisis posteriores de los datos.
- **Hora de Inicio del Video (HH:MM):** La sincronización temporal precisa facilita la correlación de eventos detectados con registros externos. Este parámetro es esencial cuando se requiere integración con otros sistemas de monitoreo o cuando se realizan análisis multi-fuente.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

- **Fotogramas por segundo (FPS):** Este parámetro especifica la velocidad de captura del video, es decir, la cantidad de imágenes (frames) que se reproducen por cada segundo de video. Por ejemplo, un video a 30 FPS contiene 30 imágenes por cada segundo de duración. Es fundamental configurar este valor correctamente para que coincida con el FPS real del video de entrada, ya que afecta directamente el cálculo de tiempos, velocidades y la sincronización temporal de las detecciones.
- **Distancia entre Frames de Detección:** Este parámetro implementa el concepto de "frame skipping", una técnica ampliamente utilizada para reducir la carga computacional en aplicaciones de tiempo real. El balance entre eficiencia y precisión debe considerarse cuidadosamente, ya que intervalos mayores pueden comprometer la continuidad del seguimiento [72].
- **Frames de Aparición:** Define el número máximo de frames consecutivos que un objeto puede estar ausente antes de que su seguimiento se descarte definitivamente. Por ejemplo, si se configura en 10 frames, el sistema mantendrá el seguimiento de un objeto aunque desaparezca temporalmente (por occlusiones o fallas de detección) hasta por 10 frames consecutivos. Si el objeto reaparece dentro de este intervalo, el seguimiento continúa; si no, se considera que el objeto ha salido de escena y su seguimiento se elimina.
- **Umbral de Confianza (0-1):** El ajuste del umbral de confianza es crítico para el balance entre sensibilidad y especificidad del detector. Valores típicos oscilan entre 0.5 y 0.8, dependiendo de la aplicación y la calidad del video de entrada.
- **Cantidad de Apariciones Mínimas:** Este parámetro complementa el filtrado temporal establecido por los "Frames de Aparición", proporcionando una capa adicional de validación para garantizar la consistencia de las detecciones a lo largo del tiempo.
- **Intervalo de Tiempo de Guardado (minutos):** Será el intervalo en minutos en los que se guardaron los datos de salida.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

- **k (Factor de Sensibilidad):** Ajusta la tolerancia del sistema de seguimiento basado en distancia euclíadiana. Un valor bajo de k hace el seguimiento más estricto, mientras que valores altos permiten mayor flexibilidad en la asociación de trayectorias
- **Procesar con:** Define, a través de un menú desplegable, si el procesamiento se realiza en CPU o GPU. La selección de GPU acelera significativamente la inferencia y el análisis del sistema. El software reconocerá la GPU de inmediato en caso de ser compatible.
- **Carpeta de Guardado:** Permite especificar el directorio donde se almacenarán los archivos de resultados generados por el sistema.
- **Modelo YOLO:** Permite seleccionar un modelo de detección YOLO propio del usuario. La elección del modelo influye en la precisión y velocidad de la detección de objetos en cada video procesado. La posibilidad de elegir el modelo específico facilita la adaptación del sistema a distintos escenarios.

Una vez configurados todos los parámetros y seleccionada la carpeta de destino, el botón "Confirmar" se habilitará, permitiendo iniciar el procesamiento del video.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

### 3.2.10 Etapa de validación.

#### 3.2.10.1 Fundamentación de las métricas utilizadas

En el desarrollo de sistemas de detección y seguimiento multi-objeto (MOT, por sus siglas en inglés), es fundamental contar con métricas que permitan evaluar de manera objetiva el rendimiento del sistema. La evaluación del rendimiento en el seguimiento de múltiples objetos representa un aspecto crítico para el diseño, optimización de parámetros y comparación de sistemas de seguimiento. Sin embargo, la elección de estas métricas depende en gran medida de las características específicas del problema abordado, así como de las limitaciones técnicas y metodológicas inherentes al sistema desarrollado [78] [79].

Además de métricas estándar como Precision, Recall, mAP@50-95 y F-Score, ampliamente utilizadas en tareas de detección de objetos [46] [37], hemos incorporado métricas específicas para evaluar la capacidad del sistema en términos de seguimiento continuo y robustez. Estas métricas incluyen la Tasa de Seguimiento Exitoso y la Tasa de Completitud de Trayectoria, las cuales abordan aspectos fundamentales del seguimiento como la continuidad temporal y la precisión de asignación de trayectorias.

#### 3.2.10.2 Limitaciones para la Implementación de MOTA y MOTP

El algoritmo desarrollado para el seguimiento de la carga de las hormigas se basa en un enfoque que combina técnicas de procesamiento de imágenes y análisis de trayectorias. Sin embargo, no fue posible calcular métricas como MOTA (Multiple Object Tracking Accuracy) y MOTP (Multiple Object Tracking Precision). Estas métricas, definidas originalmente en el marco CLEAR MOT [80], requieren una correspondencia precisa entre las detecciones predichas y las anotaciones de campo, lo cual implica la necesidad de un etiquetado manual exhaustivo y una implementación específica para manejar eventos como falsos positivos, falsos negativos e intercambios de identidad.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

MOTA se define como una métrica que considera todos los errores de configuración de objetos realizados por el rastreador (falsos positivos, pérdidas y desajustes) en todos los frames, mientras que MOTP evalúa la capacidad del rastreador para estimar posiciones precisas de objetos, independientemente de su habilidad para reconocer configuraciones de objetos [81]. Dado que nuestro sistema no cuenta con una integración directa con herramientas que faciliten este tipo de evaluación, se optó por métricas alternativas que pudieran ser calculadas de manera directa y eficiente.

El seguimiento de la carga de las hormigas presenta desafíos únicos documentados en la literatura de seguimiento de insectos [82], como la variabilidad en la forma y tamaño de la carga, la superposición de objetos en el campo visual y el movimiento errático de las hormigas. Estos factores dificultan la aplicación de métricas como MOTA y MOTP, que asumen cierto grado de estabilidad en las trayectorias y una correspondencia clara entre los objetos detectados y los ground truth.

### 3.2.10.3 Tasa de Seguimiento Exitoso

La Tasa de Seguimiento Exitoso, denominada como Correctness Measures [83] o Tracker Detection Rate [86] se define como la proporción de detecciones que fueron correctamente asignadas a una trayectoria válida en relación con el total de detecciones realizadas [83]. Esta métrica se fundamenta en los principios de asociación de datos en seguimiento multi-objeto, donde la calidad de la asignación detección-trayectoria es crucial para el rendimiento del sistema. Matemáticamente, se expresa como:

$$TSE = \frac{\text{Número de detecciones exitosas}}{\text{Número total de detecciones}}$$

Esta métrica permite evaluar la capacidad del sistema para mantener un seguimiento consistente de los objetos detectados.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

A diferencia de MOTA, que penaliza errores como falsos positivos y falsos negativos de manera acumulativa [80], la Tasa de Seguimiento Exitoso se enfoca exclusivamente en la calidad de las asignaciones realizadas. Esta característica es crucial en escenarios donde las detecciones erróneas pueden ser frecuentes debido a la complejidad del entorno, como es el caso del seguimiento de hormigas en condiciones naturales.

Al centrarse en la proporción de detecciones exitosas, esta métrica proporciona una visión clara de la eficiencia del sistema en términos de asignación de trayectorias.

### 3.2.10.4 Tasa de Completitud de Trayectoria

La Tasa de Completitud de Trayectoria, también denominado como Track Completeness [86] mide la proporción de tiempo durante el cual la carga de una hormiga fue detectada y seguida correctamente en relación con la duración total de su trayectoria. Se define matemáticamente como:

$$TCT = \frac{\text{Tiempo detectado}}{\text{Duración total de la Trayectoria}}$$

Esta métrica es especialmente útil para evaluar la capacidad del sistema para mantener un seguimiento continuo de los objetos a lo largo del tiempo. En el contexto específico del seguimiento de la carga de las hormigas, donde las trayectorias pueden ser interrumpidas debido a occlusiones, cambios abruptos en la dirección del movimiento, o comportamientos erráticos característicos de estos insectos [85], la TCT proporciona una medida directa de la estabilidad del sistema.

A diferencia de MOTP, que evalúa la precisión espacial de las detecciones, la Tasa de Completitud de Trayectoria se centra en la continuidad temporal del seguimiento. Esto es particularmente relevante en un escenario dinámico, donde la pérdida temporal de un objeto puede tener un impacto significativo en la calidad general del seguimiento. Además, esta métrica permite identificar patrones

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	3 - Materiales y Métodos

específicos de fallos, como ocurrencias frecuentes de pérdidas de seguimiento en ciertas regiones del campo visual.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	4 - Límites de implementación

## Capítulo 4: Límites de implementación

Todo sistema de visión por computadora aplicado a entornos biológicos presenta limitaciones inherentes que deben ser reconocidas y caracterizadas para garantizar la interpretación apropiada de los resultados. Nuestro desarrollo presenta restricciones específicas derivadas tanto de las decisiones metodológicas adoptadas durante el entrenamiento como de las complejidades al seguimiento de objetos deformables en escenarios con múltiples entidades [71] [72].

### 4.1 Limitaciones del dominio de aplicación

La principal restricción de nuestro sistema radica en su especialización para entornos controlados con fondos homogéneos de color blanco. Esta característica del conjunto de entrenamiento, aunque beneficiosa para maximizar el contraste entre objetos de interés y fondo, limita significativamente la generalización del modelo a escenarios naturales. Los algoritmos de aprendizaje profundo tienden a sobreajustarse a las características visuales predominantes en los datos de entrenamiento, fenómeno conocido como "dataset bias" que puede comprometer el rendimiento en condiciones ambientales diferentes [73].

La dependencia del fondo blanco homogéneo se fundamenta en principios de percepción visual y procesamiento de imágenes, donde el contraste elevado facilita la segmentación de objetos y reduce la ambigüedad en la delimitación de contornos. Sin embargo, esta simplificación experimental introduce vulnerabilidades específicas: elementos dispersos en el área de observación, como fragmentos de hojas no transportadas, pueden generar falsos positivos debido a su similitud visual con las cargas genuinas, este problema de falsos positivos lo abordamos implementando un controlador verificando que la posición inicial y final del objeto sean diferentes, si estas son iguales significa que el elemento es estático y se descarta del análisis. Es por esto que es importante seleccionar un área de detección con la menor cantidad de elementos dispersos.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	4 - Límites de implementación

#### 4.2 Desafíos en la consistencia de identificadores

Una limitación crítica del sistema de seguimiento implementado surge durante eventos de intersección entre objetos múltiples, particularmente cuando hormigas portadoras de carga se cruzan o superponen temporalmente. Los algoritmos de asociación de datos, aunque robustos para trayectorias aisladas, enfrentan ambigüedades cuando múltiples objetos ocupan regiones espaciales similares simultáneamente. Durante estos eventos de oclusión mutua, el sistema puede experimentar intercambio de identidades (ID swapping), donde los identificadores asignados a objetos específicos se transponen erróneamente [76] [77].

Este fenómeno es problemático cuando dos hormigas con cargas se aproximan dentro del umbral de asociación espacial del algoritmo de seguimiento. La proximidad espacial puede resultar en asociaciones erróneas donde la carga de la hormiga A se asocia incorrectamente con la trayectoria de la hormiga B, y viceversa. Estos intercambios de identidad introducen discontinuidades artificiales en las trayectorias y comprometen la integridad de los análisis.

El problema se agrava en escenarios de alta densidad poblacional, donde la probabilidad de intercepciones aumenta con el número de individuos presentes [76]. A diferencia de objetos rígidos con características visuales distintivas y estables, las cargas orgánicas presentan variabilidad morfológica que dificulta la implementación de técnicas de re-identificación basadas en descriptores visuales.

#### 4.3 Criterios de validación para análisis estadístico

Para mitigar los efectos de las limitaciones mencionadas y garantizar los resultados, implementamos criterios estrictos de validación que determinan qué trayectorias se incluyen en el análisis estadístico final. El criterio fundamental requiere que cada objeto complete una trayectoria desde su primera aparición hasta su salida definitiva del área de observación, siguiendo los parámetros de configuración del sistema. Esta condición de "trayectoria completa" asegura que solo objetos con seguimiento confiable y consistente contribuyan a las métricas [74].

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	4 - Límites de implementación

Trayectorias que experimentan interrupciones significativas debido a pérdidas de seguimiento, intercambios de identidad o salidas prematuras del campo visual se excluyen automáticamente del análisis con la ayuda de nuestra función de filtrado. Aunque esta estrategia conservadora puede reducir el tamaño muestral efectivo, el usuario tendrá la posibilidad de modificar los parámetros de configuración logrando así una mayor o menor flexibilidad en el conteo [74].

Adicionalmente, se implementan umbrales mínimos de duración de trayectoria para excluir detecciones espurias u objetos que atraviesan el campo visual demasiado brevemente para constituir observaciones comportamentales significativas. Estos criterios de filtrado múltiple, aunque introducen cierto conservadurismo en el análisis, proporcionan confianza estadística en que los resultados derivados representan fidedignamente los patrones de comportamiento de las hormigas cortadoras bajo las condiciones específicas [75].

Estas limitaciones definen claramente el dominio de aplicabilidad del sistema desarrollado y establecen las bases para futuras mejoras metodológicas dirigidas a expandir el enfoque propuesto.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	5 - Resultados

## Capítulo 5: Resultados

Para evaluar el rendimiento del sistema desarrollado, se utilizaron videos pregrabados de 30 minutos de duración, cada uno capturado bajo condiciones experimentales específicas y controladas. Estas pruebas se complementaron con evaluaciones de 24 horas, ejecutando ambos tipos de pruebas bajo tres regímenes de parámetros: permisivos, semi-permisivos y estrictos. La selección de estos videos respondió a la necesidad de someter el sistema a escenarios diversos y representativos de las condiciones reales de campo, incluyendo variaciones en la iluminación natural, episodios de superposición entre individuos, y patrones de movimiento característicos de las hormigas portadoras de carga.

El modelo seleccionado para la evaluación demostró el mejor balance entre precisión de detección y eficiencia computacional durante las pruebas preliminares, características esenciales para el funcionamiento del sistema en condiciones de campo. La evaluación se centró en dos aspectos fundamentales: la capacidad del sistema para asignar correctamente las detecciones a trayectorias individuales y su habilidad para mantener la continuidad del seguimiento a lo largo de secuencias temporales extendidas.

Los resultados se organizaron según las métricas específicamente desarrolladas para este estudio: la Tasa de Seguimiento Exitoso (TSE) y la Tasa de Completitud de Trayectoria (TCT), las cuales fueron desarrolladas para evaluar cada intervalo de guardado. En el archivo de salida se visualizará la puntuación tanto de TSE como de TCT de cada intervalo de tiempo. Esta aproximación metodológica permitió evaluar tanto la precisión instantánea del sistema como su robustez temporal, proporcionando una visión integral del desempeño en diferentes contextos operativos.

Cada video fue analizado individualmente, documentando las condiciones específicas que influyeron en el rendimiento del sistema. Los análisis incluyen la caracterización de los desafíos particulares presentados por cada secuencia, como la

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	5 - Resultados

densidad de individuos en el campo visual, las condiciones de iluminación, y la presencia de elementos que pudieran interferir con la detección o el seguimiento.

Los resultados detallados se presentan organizados por video de prueba, incluyendo para cada caso: las métricas de rendimiento obtenidas en las tres configuraciones paramétricas y ambas duraciones de prueba, la identificación de los factores que contribuyeron al éxito o las limitaciones observadas, y un análisis de cómo las características específicas del entorno experimental afectaron el desempeño general del sistema. Esta estructura comparativa permitió evaluar simultáneamente la sensibilidad a parámetros, la consistencia del sistema a través de diferentes condiciones operativas e identificar las circunstancias óptimas para su implementación en estudios de campo.

Para cada evaluación se cambiaron tres parámetros claves del sistema para que sea permisivo o estricto, estos fueron cantidad de apariciones, confianza y frames de aparición, visualizados en la siguiente tabla.

Parámetro	Permisivo	Semi-Permisivo	Estricto
Cantidad de apariciones	5	10	15
Confianza	0.4	0.6	0.75
Frames de aparición	10	15	15

Tabla 6 - Definición de parámetros de prueba.

## 5.1 Entorno de menor actividad

El análisis de los resultados se realizó sobre tres intervalos de tiempo, con un entorno como se visualiza en la figura 29.

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	5 - Resultados



*Figura 29 - Entorno del video utilizado*

Los resultados revelan diferencias significativas en el rendimiento del sistema según el régimen de parámetros aplicado, visualizados en las tablas 7, 8 y 9. En la configuración permisiva, se detectaron un total de 44 hojas distribuidas en 18, 11 y 15 hojas por intervalo respectivamente, con medianas de área entre 358-430 mm<sup>2</sup>. La TSE varió entre 0.79-0.86, mientras que la TCT se mantuvo estable entre 0.94-0.98. La configuración semi-permisiva detectó 34 hojas totales (15, 7 y 12 por intervalo), con medianas de área progresivamente mayores (386-480 mm<sup>2</sup>). Notablemente, esta configuración alcanzó una TSE perfecta (1.00) en el tercer intervalo, con valores de TCT consistentemente altos (0.94-0.96). La configuración estricta registró el menor número de detecciones con 21 hojas totales (7, 5 y 9 por intervalo), pero exhibió las mayores medianas de área (514-560 mm<sup>2</sup>). En la figura 30 podemos visualizar directamente estas diferencias.

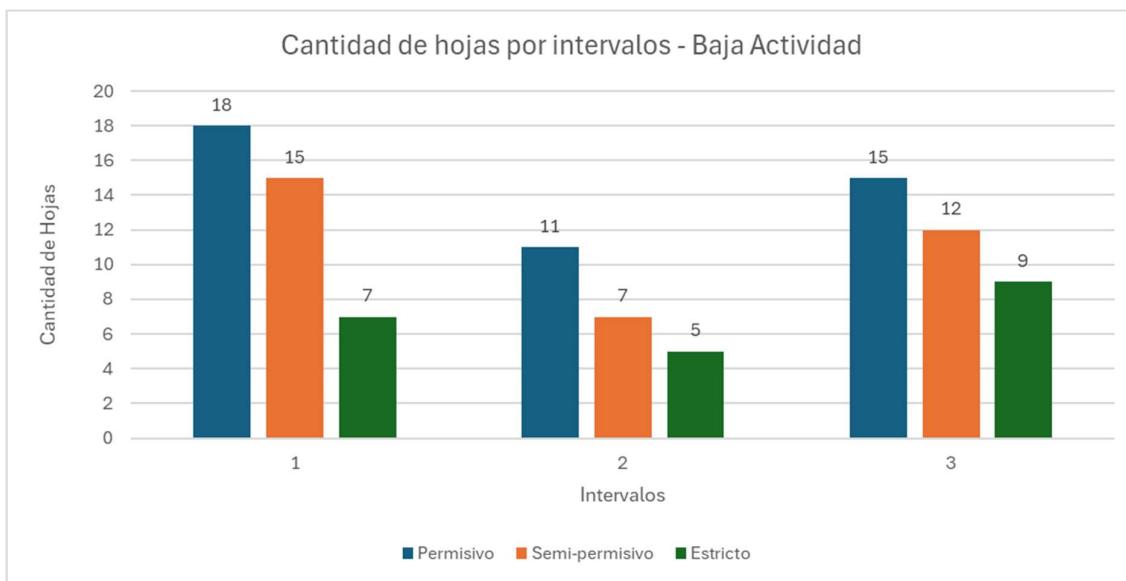


Figura 30 - Comparación de cantidad de hojas por intervalo

La relación inversa entre restrictividad paramétrica y cantidad de detecciones es evidente: los parámetros más permisivos detectaron el doble de objetos que los estrictos (44 vs 21), sugiriendo que configuraciones más permisivas capturan eventos de carga menores que podrían ser filtrados en regímenes restrictivos. Sin embargo, la calidad del seguimiento mostró patrones complejos: la TSE alcanzó valores máximos (1.00) tanto en configuraciones semi-permisivas como estrictas durante el tercer intervalo, mientras que los parámetros permisivos mantuvieron rendimientos más consistentes pero nunca perfectos. La TCT demostró mayor estabilidad en todas las configuraciones, variando solo entre 0.94-0.98.

Un hallazgo emerge en la correlación entre tamaño de carga y régimen paramétrico: las configuraciones estrictas favorecieron la detección de cargas significativamente mayores (mediana de aproximadamente 540 mm<sup>2</sup>) comparado con los permisivos (aproximadamente 360 mm<sup>2</sup>), confirmando que la restrictividad paramétrica funciona como filtro selectivo basado en características físicas del objeto. Esta selectividad explica simultáneamente la menor cantidad de detecciones y la tendencia hacia mejor calidad de seguimiento en objetos de mayor tamaño, validando la relación establecida entre dimensiones del objeto y facilidad de tracking.

Resultado parámetros permisivos.

Intervalo	Cantidad de Hojas	Mediana mm <sup>2</sup>	Percentil 25 mm <sup>2</sup>	Percentil 75 mm <sup>2</sup>	TSE	TCT
1	18	358.09	264.76	422.03	0.86	0.94
2	11	358.07	270.92	394.14	0.79	0.94
3	15	430.46	317.38	499.89	0.83	0.98

Tabla 7 - Resultados parámetros permisivos.

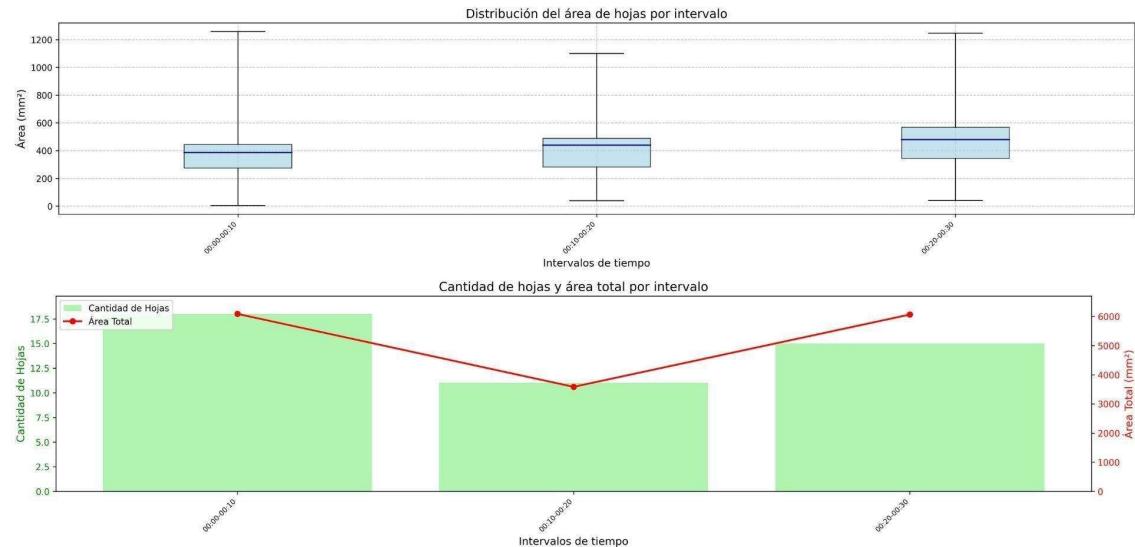


Figura 31 - Gráfico de la distribución por intervalo en parámetros permisivos.

Resultados parámetros semi-permisivos

Intervalo	Cantidad de Hojas	Mediana mm <sup>2</sup>	Percentil 25 mm <sup>2</sup>	Percentil 75 mm <sup>2</sup>	TSE	TCT
1	15	386.7	273.62	444.55	0.88	0.94
2	7	439.3	281.38	489.26	0.78	0.96
3	12	480.5	343.67	567.64	1	0.95

Tabla 8 - Resultados parámetros semi-permisivos.

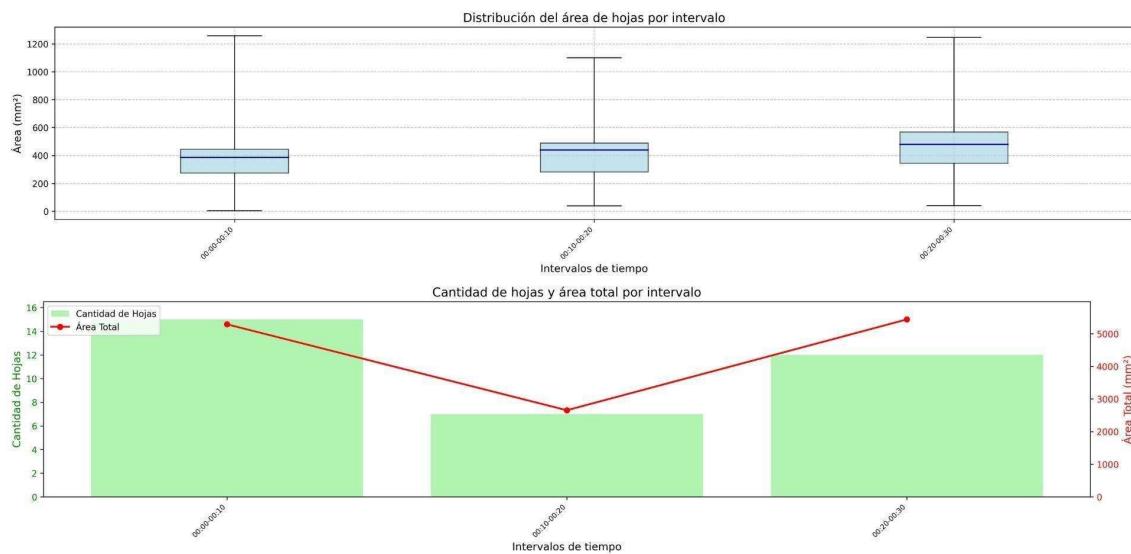


Figura 32 - Gráfico de la distribución por intervalo en parámetros semi-permisivos

Resultados parámetros estrictos

Intervalo	Cantidad de Hojas	Mediana mm <sup>2</sup>	Percentil 25 mm <sup>2</sup>	Percentil 75 mm <sup>2</sup>	TSE	TCT
1	7	513.96	333.66	593.88	0.88	0.97
2	5	534.78	380.64	619.78	0.71	0.94
3	9	560.19	418.09	656.45	1	0.96

Tabla 9 - Resultados parámetros estrictos

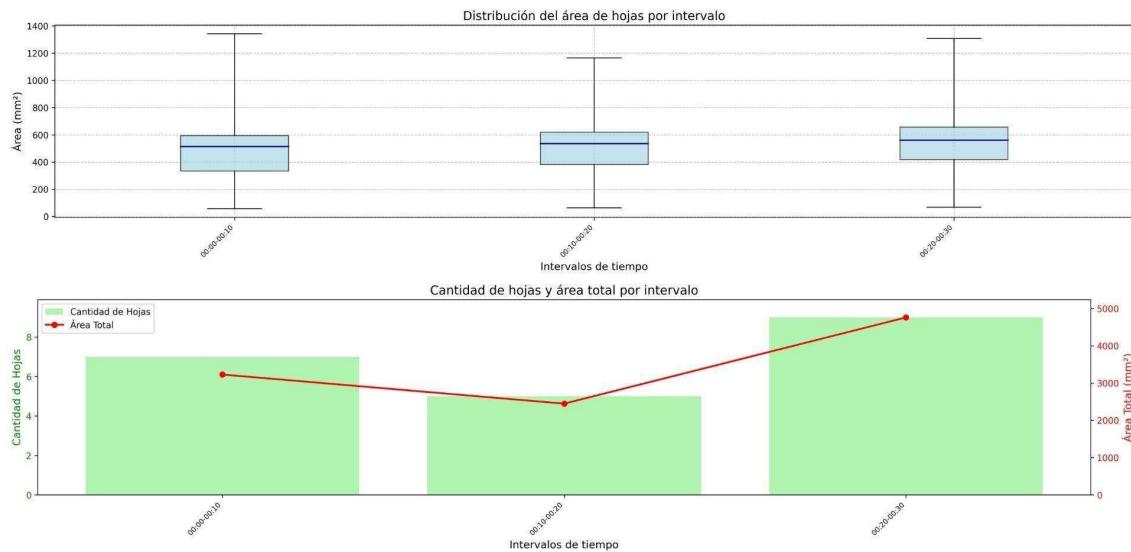


Figura 33 - Gráfico de la distribución por intervalo en parámetros estrictos.

Los resultados demuestran que no existe una configuración paramétrica universalmente superior, sino intercambios específicos entre sensibilidad de detección y precisión de seguimiento, donde la selección óptima depende de los objetivos: maximizar cobertura de eventos (permisivo) versus optimizar calidad de tracking en cargas prominentes (estricto).

## 5.2 Entorno de mayor actividad

El análisis del segundo entorno se realizó bajo condiciones ambientales distintas, incluyendo variaciones en la iluminación y rotación de la cámara, como se observa en la Figura 34. Este entorno se caracterizó por una actividad significativamente mayor del hormiguero, proporcionando un escenario desafiante para evaluar la robustez del sistema de seguimiento en condiciones de alta densidad de objetos.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	5 - Resultados



Figura 34 - Entorno del segundo video utilizado

En el entorno de mayor actividad, la comparación entre los tres regímenes de parámetros (permisivo, semi-permisivo y estricto) revela diferencias notables en la sensibilidad y precisión del sistema de seguimiento ante altas densidades de objetos. Bajo parámetros permisivos, se detectó el mayor número de hojas en todos los intervalos (37, 42 y 48), aunque la Tasa de Seguimiento Exitoso (TSE) se mantuvo moderada (0.53–0.59) y la Tasa de Completitud de Trayectoria (TCT) relativamente alta (0.89–0.93). Esta configuración maximizó la cobertura de eventos, pero con menor precisión en la asignación de trayectorias individuales.

La configuración semi-permisiva presentó un equilibrio entre sensibilidad y precisión, con cantidades intermedias de hojas detectadas (27, 32 y 32) y mejoras en la TSE en el segundo intervalo (0.62). Sin embargo, la TSE fue menor en el primer y tercer intervalo (0.44 y 0.51), reflejando la dificultad de mantener la precisión bajo condiciones de alta actividad. La TCT se mantuvo estable y elevada (0.88 -- 0.95), lo que indica un seguimiento continuo una vez establecida la trayectoria.



Por último, la configuración estricta resultó en la menor cantidad de detecciones (22, 28 y 29 hojas), pero con valores de TSE y TCT similares a los de la configuración semi-permisiva (TSE: 0.45–0.77, TCT: 0.90–0.95). Este régimen favoreció la detección de cargas de mayor tamaño, como lo evidencian las medianas de área superiores, y filtró eventos ambiguos, priorizando la calidad sobre la cantidad de trayectorias. En el siguiente gráfico podemos observar la cantidad de hojas por intervalo comparando cada configuración de parámetros.

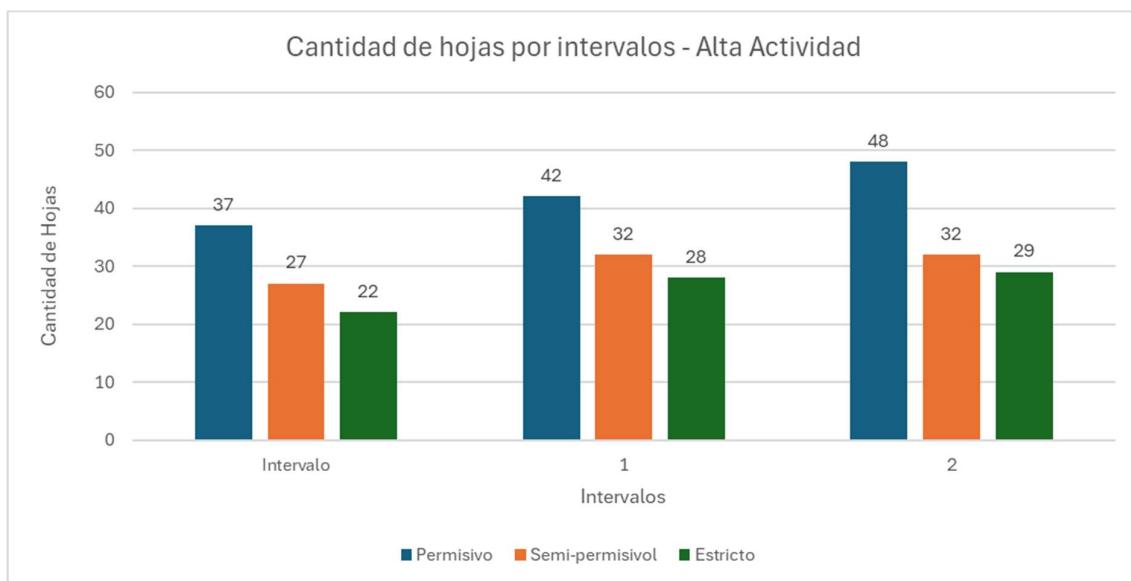


Figura 35 - Comparación de cantidad de hojas por intervalo

Estos resultados confirman que, en escenarios de alta actividad, existe un compromiso entre la sensibilidad del sistema y la precisión del seguimiento: parámetros permisivos favorecen la detección masiva a expensas de la precisión, mientras que parámetros estrictos priorizan la calidad de las trayectorias, filtrando detecciones menos confiables. La configuración semi-permisiva ofrece un balance intermedio, aunque la dificultad inherente al entorno de alta densidad limita la TSE en todos los casos. La TCT se mantuvo elevada en todos los regímenes, destacando la robustez del sistema para mantener el seguimiento una vez establecida la asociación inicial, incluso bajo condiciones desafiantes.

Resultado parámetros permisivos.



Intervalo	Cantidad de Hojas	Mediana mm <sup>2</sup>	Percentil 25 mm <sup>2</sup>	Percentil 75 mm <sup>2</sup>	TSE	TCT
1	37	602,68	450,79	682,2	0,53	0,93
2	42	679,84	487,52	781,61	0,59	0,92
3	48	562,7	437,55	641,03	0,48	0,89

Tabla 10 - Resultados parámetros permisivos.

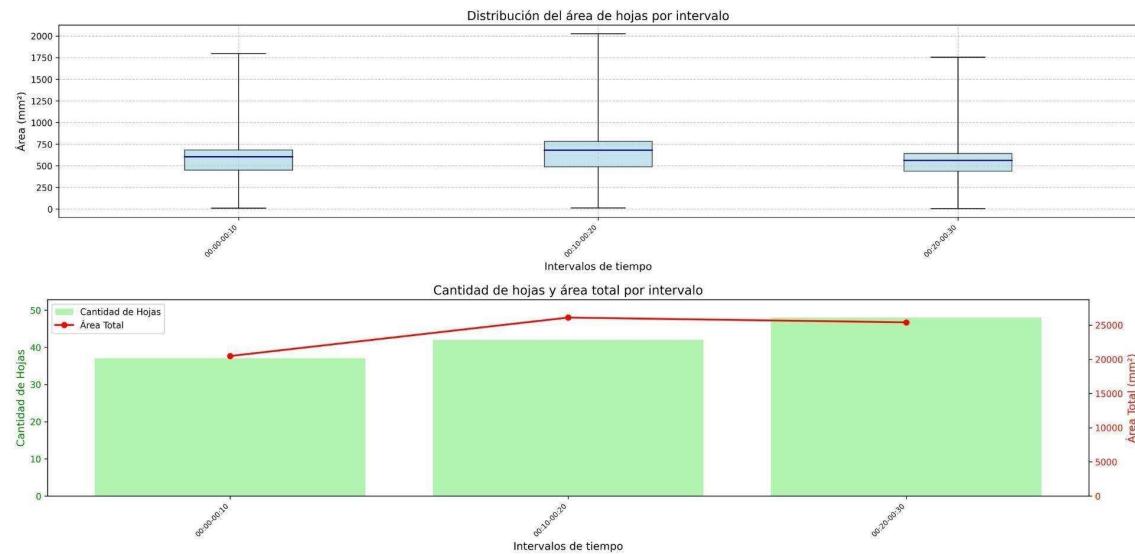


Figura 36 - Gráfico de la distribución por intervalo en parámetros permisivos.

Resultado parámetros semi-permisivos

Intervalo	Cantidad de Hojas	Mediana mm <sup>2</sup>	Percentil 25 mm <sup>2</sup>	Percentil 75 mm <sup>2</sup>	TSE	TCT
1	27	860,8	663,26	966,48	0,44	0,95
2	32	906,15	621,5	1044,72	0,62	0,91
3	32	684,94	535,37	779,03	0,51	0,88

Tabla 11 - Resultados parámetros semi-permisivos.

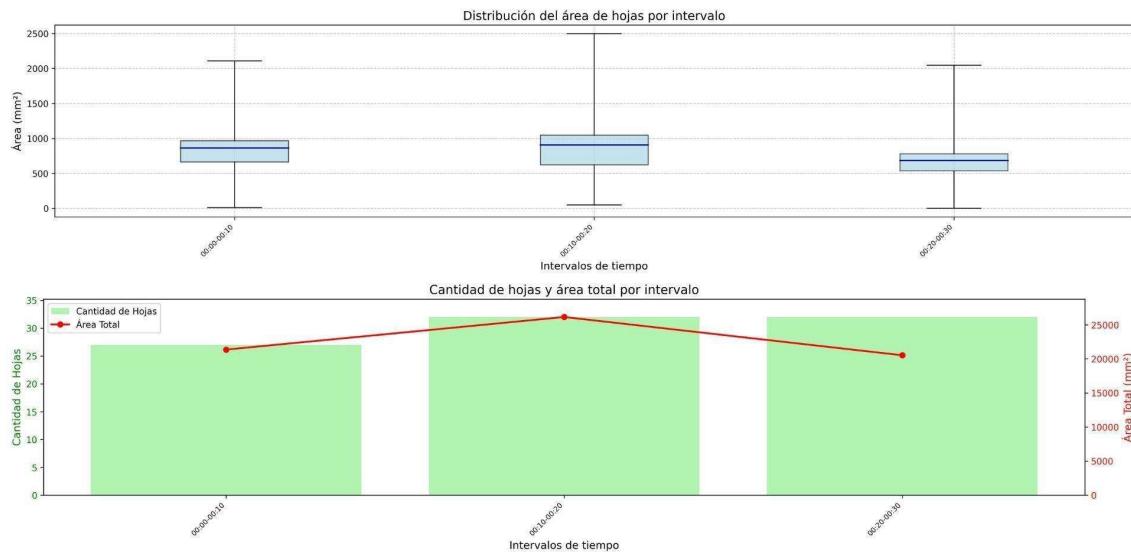


Figura 37 - Gráfico de la distribución por intervalo en parámetros semi-permisivos

Resultado parámetros estrictos.

Intervalo	Cantidad de Hojas	Mediana mm <sup>2</sup>	Percentil 25 mm <sup>2</sup>	Percentil 75 mm <sup>2</sup>	TSE	TCT
1	22	711,32	565,02	784,85	0,48	0,95
2	28	782,39	565,18	891,3	0,7	0,93
3	29	569,71	451,55	639,22	0,45	0,90

Tabla 12 - Resultados parámetros estrictos.

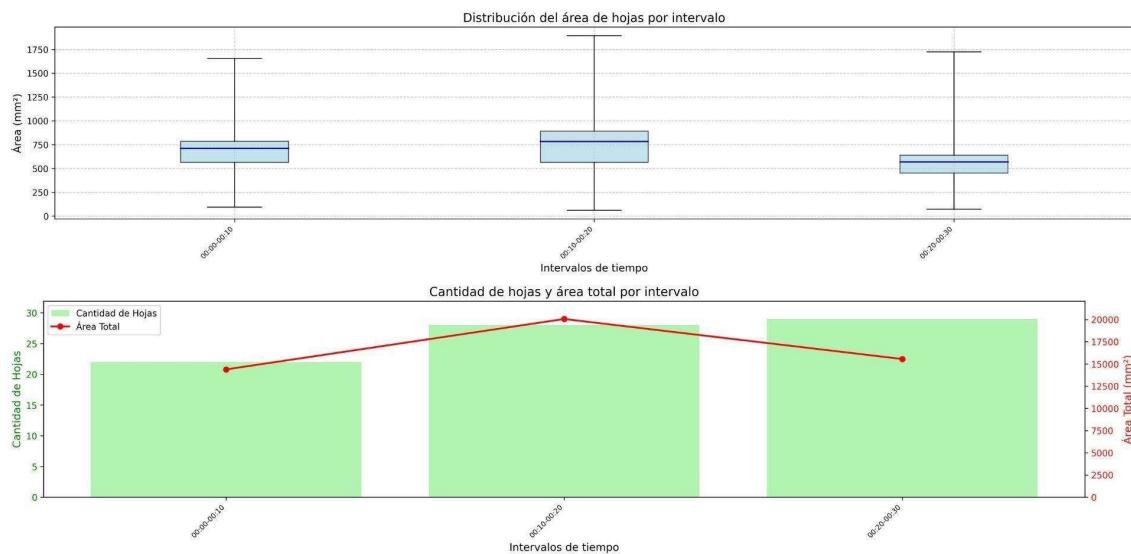


Figura 38 - Gráfico de la distribución por intervalo en parámetros estrictos.

### 5.3 Análisis comparativo entre entornos

La comparación entre ambos entornos revela diferencias fundamentales en el comportamiento del sistema bajo distintos niveles de actividad y densidad de objetos. En el entorno de baja actividad, la cantidad promedio de hojas detectadas por intervalo fue considerablemente menor en todos los regímenes de parámetros, mientras que en el entorno de alta actividad los promedios aumentaron significativamente. Esta diferencia de densidad se tradujo en una disminución consistente de la TSE al pasar de un entorno controlado a uno de mayor congestión. Esta relación inversa entre densidad de objetos y rendimiento de asignación confirma las predicciones teóricas sobre las limitaciones de sistemas de seguimiento multi-objeto en escenarios de alta congestión [87] [88].

Los resultados demuestran que el sistema desarrollado mantiene funcionalidad efectiva a través de entornos con características marcadamente diferentes, desde condiciones controladas de baja densidad hasta escenarios de alta actividad con múltiples interacciones simultáneas. Esta adaptabilidad representa una característica



crucial para aplicaciones de seguimiento de comportamiento animal en condiciones naturales, donde la variabilidad ambiental es inherente [89].

#### 5.4 Resultados en vídeos de larga duración.

Además de las pruebas antes señaladas con videos de 30 minutos, también efectuamos las mismas sobre videos de 24 horas de duración, comprobando así el funcionamiento exitoso del sistema. Los resultados fueron plasmados en gráficos, debido a que, si mostrábamos las tablas de salida hubieran sido 3 tablas de 144 filas, es por esto que decidimos anteriormente demostrar el funcionamiento con videos de 30 minutos. En las siguientes imágenes pudimos comparar la cantidad de hojas y el área acumulada por intervalo.

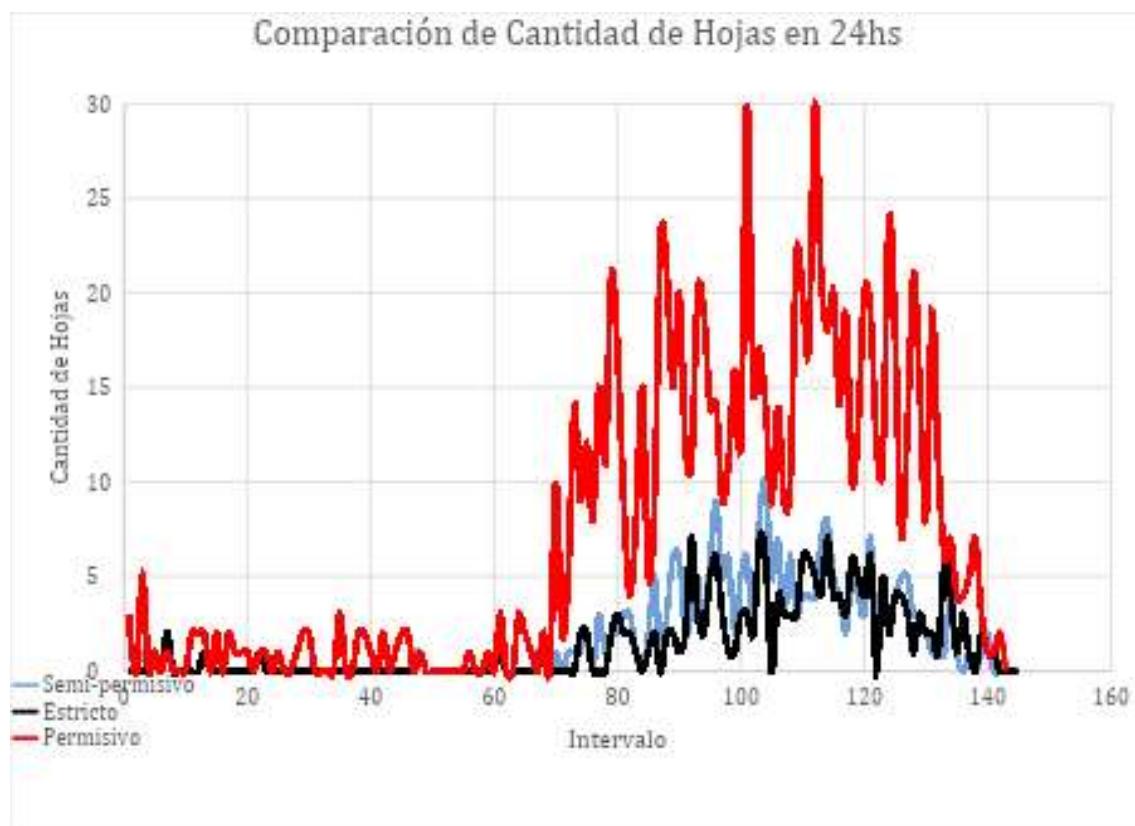


Figura 39 - Cantidad de hojas por intervalo.

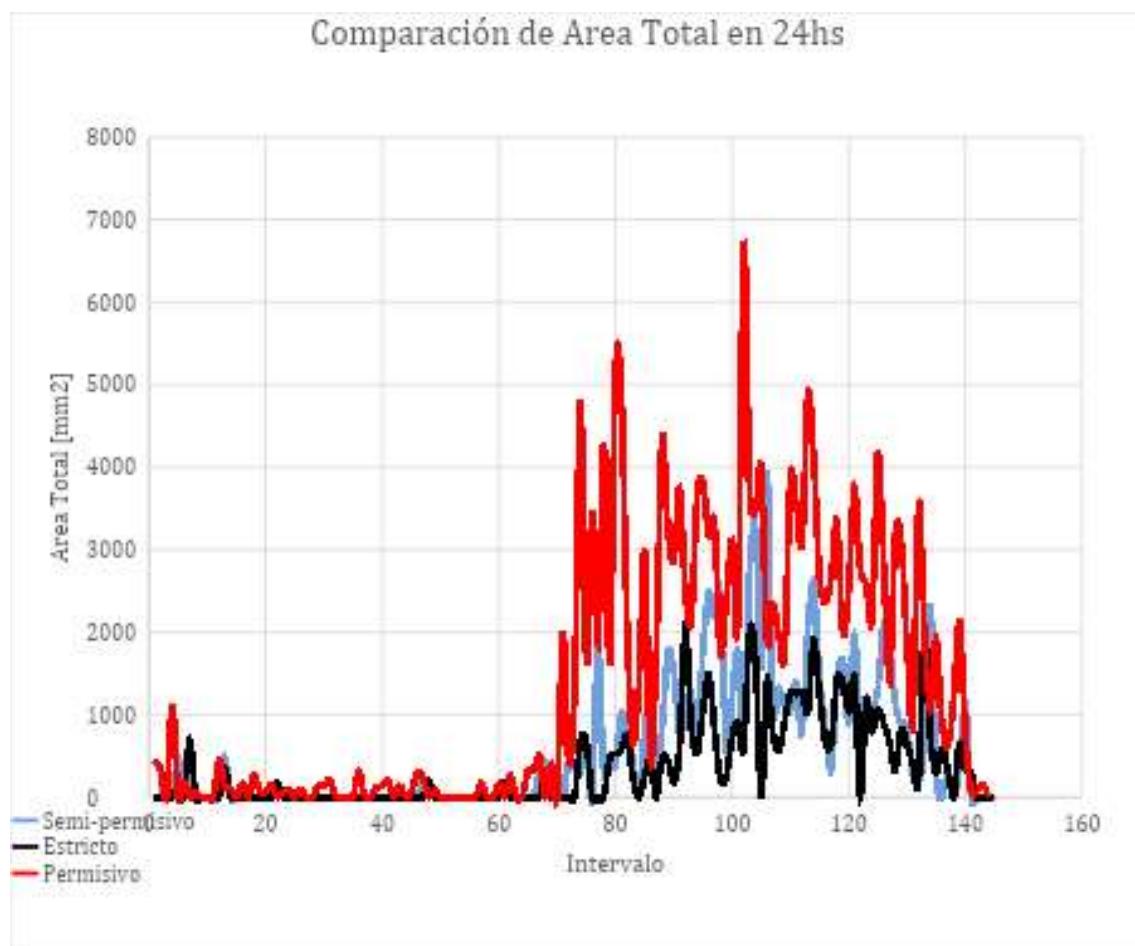


Figura 40 - Área total por intervalo.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	6- Mejoras Futuras y Líneas de Investigación

## Capítulo 6: Mejoras Futuras y Líneas de Investigación

Una vez alcanzados los objetivos planteados y validado el funcionamiento del sistema desarrollado, se identificaron diversas oportunidades de mejora que pueden expandir las capacidades y aplicabilidad del sistema. Estas mejoras se enfocan en tres áreas principales: optimización temporal, robustez del seguimiento, y adaptabilidad a entornos naturales.

### 6.1 Implementación en Tiempo Real

El sistema desarrollado presenta un potencial considerable para la implementación en tiempo real, aprovechando las características de velocidad inherentes de la arquitectura YOLO. La transición hacia procesamiento en tiempo real requeriría optimizaciones específicas en la gestión de memoria y paralelización de procesos, así como la implementación de técnicas de predicción anticipada para reducir la latencia computacional [105] [106]. Esta mejora permitiría la aplicación del sistema en estudios de comportamiento animal que requieren respuesta inmediata, como experimentos de condicionamiento o análisis de interacciones sociales dinámicas [107].

La implementación en tiempo real también facilitaría la integración con sistemas de retroalimentación experimental, donde las decisiones basadas en el comportamiento observado puedan influir en las condiciones experimentales de manera instantánea. Esto abriría nuevas posibilidades en el estudio del comportamiento adaptativo y la toma de decisiones en colonias de insectos [108].

### 6.2 Robustez ante Oclusiones Complejas

Una limitación identificada en el sistema actual es la pérdida de seguimiento durante eventos de oclusión prolongada, particularmente cuando múltiples individuos se superponen o cruzan simultáneamente. El desarrollo de algoritmos avanzados de manejo de oclusiones, basados en técnicas de predicción de trayectorias y modelos

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	6- Mejoras Futuras y Líneas de Investigación

de movimiento probabilísticos, podría mantener la continuidad del seguimiento incluso durante eventos de oclusión completa [109].

La implementación de métodos de re-identificación basados en características morfológicas específicas de la carga transportada, combinados con modelos de predicción de movimiento a largo plazo, podría resolver los desafíos asociados con la pérdida temporal de visibilidad [110]. Estas mejoras son particularmente relevantes para el estudio de comportamientos de forrajeo en alta densidad, donde las interacciones físicas entre individuos son frecuentes y complejas [111].

### 6.3 Adaptación a Entornos Naturales

El entrenamiento del modelo actual se realizó utilizando un fondo controlado de color blanco, lo que limita su aplicabilidad directa en entornos naturales. Una extensión fundamental del trabajo involucraría el desarrollo de un dataset más diverso que incluya variaciones en substrate, iluminación natural, y elementos de vegetación típicos del hábitat natural de las hormigas cortadoras.

La adaptación a entornos naturales requeriría la implementación de técnicas de segmentación semántica avanzadas para distinguir entre la carga transportada y elementos del fondo con características visuales similares. Además, sería necesario incorporar algoritmos de corrección automática de iluminación y técnicas de realce de contraste adaptativo para mantener la precisión de detección bajo condiciones lumínicas variables típicas del campo.

Esta mejora representaría un avance significativo hacia la aplicación del sistema en estudios ecológicos de campo, permitiendo el análisis no invasivo del comportamiento de forrajeo en condiciones completamente naturales y contribuyendo al entendimiento de los patrones de transporte de material en ecosistemas reales [112].

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	7 - Discusión

## Capítulo 7: Discusión

Los resultados obtenidos en este trabajo demuestran un rendimiento competitivo cuando se comparan con sistemas similares en la literatura. Alejandro Requena García [101] desarrolló un sistema de visión computacional para el seguimiento de hormigas cortadoras, reportando una precisión de detección de 90% en su modelo, con un recall superior a 85%. Nuestro sistema alcanzó valores de precisión y recall similares en condiciones controladas, con diferencia que en la tarea de seguimiento implementó algoritmos como Byte-Track y BOT-Sort, ejecutando cada frame en 30-40 ms, mientras que nuestro seguimiento basado en distancia euclidiana se realiza en tiempos entre 10-20 ms por frame, llegando a valores aún menores con el hardware adecuado. También afirma que la detección de carga de la hormiga fue la tarea más desafiante de su proyecto, mientras que el eje de nuestro proyecto fue la detección de carga, ciertamente es un objetivo desafiante, y por sobre todo difícil de parametrizar, como afirma en su proyecto Francisco Sturniolo [7].

El estudio realizado por Apeinans et al. [102] destaca que los entrenamientos con datasets basados en laboratorios presentan una precisión de 98% mientras que el mismo modelo se pone a prueba en entornos naturales decae a una precisión del 6%, en nuestro proyecto sucede un fenómeno similar donde, si buscamos evaluar la carga de las HCH fuera de un entorno controlado, la precisión caerá significativamente, es por eso que la base blanca sobre la cual se experimentó es crucial al momento de la evaluación del contenido videográfico.

Wu et al. [103] propone un sistema basado en Resnet50 para el estudio de hormigas en ambientes tanto internos como externos, donde se encontró con problemas similares a los que nos encontramos nosotros, afirmando que debido a sus interacciones sumamente cercanas, es muy probable que se produzcan detecciones incorrectas.

Wang et al. [104] realizó un estudio de detección de insectos basado en YOLO, teniendo como resultado un modelo de detección con una precisión de hasta 92,8%

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	7 - Discusión

y valores de recall de hasta 95,3 %, valores similares a nuestro modelo, con la diferencia que nuestro modelo fue basado en la segmentación.

El estudio de Imirzian et al. [85] sobre hormigas carpinteras (*C. rufipes*) en bosques tropicales demostró que el seguimiento manual de trayectorias requiere ~2 horas de trabajo humano por cada 10 minutos de video analizado, mientras que la utilización de nuestro software con el hardware (GPU) adecuado puede procesar videos de 24 horas, en 24 horas o menos.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	8 - Conclusión

## Capítulo 8: Conclusión

Nuestro trabajo ha logrado desarrollar exitosamente un sistema automatizado para el seguimiento y análisis del comportamiento forrajero de hormigas cortadoras, cumpliendo con los objetivos planteados y superando en varios aspectos a sistemas similares reportados en la literatura.

Las contribuciones principales incluyen una integración tecnológica novedosa que, a diferencia de trabajos previos que utilizan métodos tradicionales de visión computacional o redes neuronales más antiguas, presenta un sistema pionero en la aplicación de YOLOv8 para el seguimiento de carga transportada por hormigas, logrando un balance óptimo entre precisión y eficiencia computacional. La introducción de TSE y TCT como métricas de evaluación representa una contribución metodológica que puede ser adoptada por futuros estudios en el campo, proporcionando una evaluación más relevante que las métricas genéricas de MOT. El desarrollo de una interfaz gráfica intuitiva y la capacidad de procesamiento en lotes hacen que nuestro sistema sea más accesible para investigadores sin experiencia en programación, una ventaja significativa sobre sistemas que requieren conocimientos técnicos especializados.

En conclusión, este proyecto no solo cumple con los objetivos académicos planteados, sino que también proporciona una herramienta práctica y funcional que puede acelerar significativamente las investigaciones en comportamiento de insectos sociales. La combinación de técnicas avanzadas de deep learning con un diseño centrado en el usuario posiciona a este sistema como una contribución valiosa al campo de la mirmecología computacional.

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	9 - Bibliografía

## Capítulo 9: Bibliografía

- [1] J. M. Cherrett, "The biology, pest status and control of leaf-cutting ants," in Economic impact and control of social insects, S. B. Vinson, Ed. New York, NY, USA: Praeger Publishers, 1986.
- [2] B. Hölldobler and E. O. Wilson, *The Ants*. Cambridge, MA, USA: Harvard University Press, 1990.
- [3] H. G. Fowler, L. C. Forti, C. R. F. Brandão, J. H. C. Delabie, and H. L. Vasconcelos, "Ecología nutricional de formigas cortadeiras," in Ecología nutricional de insetos e suas implicações no manejo de pragas, A. R. Panizzi and J. R. P. Parra, Eds. São Paulo, Brazil: Manole, 1989.
- [4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, May 2015.
- [5] T.-Y. Lin et al., "Microsoft COCO: Common Objects in Context," in Proc. European Conference on Computer Vision (ECCV), Zurich, Switzerland, Sep. 2014, pp. 740-755.
- [6] D. Tuia, B. Kellenberger, S. Beery, S. R. B. Ferreira, R. D. J. van der GueSt, G. D. L. G. Willemsen, A. C. G. van der Grijp, T. D. H. van der Grijp, and H. L. Vasconcelos, "Perspectives in machine learning for wildlife conservation," *Nature Communications*, vol. 13, no. 1, p. 797, Feb. 2022, doi: 10.1038/s41467-022-28438-8.
- [7] F. D. Sturniolo, "Desarrollo de una herramienta para identificación automática del ritmo de forrajeo de hormigas cortadoras de hojas a partir de registros de video," Tesis de Grado, Fac. de Ing., Univ. Nacional de Entre Ríos, Oro Verde, Argentina, 2021.
- [8] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. New York, NY, USA: Pearson, 2017.
- [9] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [10] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.
- [11] S. Pattanayak, *Pro Deep Learning with TensorFlow: A Mathematical Approach to Advanced Artificial Intelligence in Python*. Berkeley, CA, USA: Apress, 2017.
- [12] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Comput. Surv.*, vol. 38, no. 4, Art. no. 13, Dec. 2006, doi: 10.1145/1177352.1177355.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	9 - Bibliografía

[13] M. Abadi *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467*, 2016. [En línea]. Disponible en: <http://arxiv.org/abs/1603.04467>

[14] G. Jocher, A. Chaurasia, and J. Qiu, *Ultralytics YOLO*. [Software]. Enero de 2023. [En línea]. Disponible en: <https://github.com/ultralytics/ultralytics>

[15] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "You only learn one representation: Unified network for multiple tasks," *arXiv preprint arXiv:2105.04206*, 2021. [En línea]. Disponible en: <http://arxiv.org/abs/2105.04206>

[16] S. A. Soumia, B. Asma, y N. Khaoula, "Comparative Evaluation of YOLOv5 and YOLOv8 Across Diverse Datasets," en *The 2nd Electrical Engineering International Conference (EEIC'23)*, Bejaia, Argelia, Dic. 2023.

[17] Ultralytics, "YOLOv8 vs YOLOv5: comparación detallada". Accedido: 27 de mayo de 2025. [En línea]. Disponible en: <https://docs.ultralytics.com/es/compare/yolov8-vs-yolov5>

[18] Ultralytics, "Ultralytics YOLO Docs". Accedido: 27 de mayo de 2025. [En línea]. Disponible en: <https://docs.ultralytics.com/es/models/yolov8>

[19] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 6, no. 1, pp. 1-48, 2019.

[20] K. Tan, Y. Zhou, Q. Xia, R. Liu, y Y. Chen, "Large Model based Sequential Keyframe Extraction for Video Summarization," *arXiv*, 2024. [En línea]. Disponible en: <https://arxiv.org/abs/2401.04962>

[21] J. A. Sabattini *et al.*, "AntVideoRecord: Autonomous system to capture the locomotor activity of leafcutter ants," ResearchGate, Jan. 2022. [En línea]. Disponible en: [https://www.researchgate.net/publication/358455453\\_AntVideoRecord\\_Autonomous\\_system\\_to\\_capture\\_the\\_locomotor\\_activity\\_of\\_leafcutter\\_ants](https://www.researchgate.net/publication/358455453_AntVideoRecord_Autonomous_system_to_capture_the_locomotor_activity_of_leafcutter_ants)

[22] "Formación de modelos con Ultralytics YOLO," *Ultralytics YOLO Docs*. [En línea]. Disponible en: <https://docs.ultralytics.com/es/modes/train/#train-settings>

[23] "Data Augmentation for Bounding Boxes: Rotation and Shearing," *DigitalOcean Community*. [En línea]. Disponible en: <https://www.digitalocean.com/community/tutorials/data-augmentation-for-object-detection-rotation-and-shearing>

[24] K. Wada, "Labelme: Image polygonal annotation with Python." [Software]. Zenodo, 2021. doi: 10.5281/zenodo.5711226.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	9 - Bibliografía

[25] A. Ariuntuya, "Improve accuracy: Polygon annotations for object detection," *Roboflow Blog*, Jul. 19, 2023. [En línea]. Disponible en: <https://blog.roboflow.com/polygon-vs-bounding-box-computer-vision-annotation/>

[26] "Augment Images," *Roboflow Docs*. [En línea]. Disponible en: <https://docs.roboflow.com/datasets/dataset-versions/image-augmentation>

[27] A. Bernadou, A. Felden, M. Moreau, P. Moretto, and V. Fourcassié, "Ergonomics of load transport in the seed harvesting ant *Messor barbarus*: Morphology influences transportation method and efficiency," *J. Exp. Biol.*, vol. 219, no. 18, pp. 2920–2927, Sep. 2016, doi: 10.1242/jeb.141556.

[28] S. Joshi, "Top 3 Video Annotation Use Cases for Optimizing Computer Vision," *Hitech BPO*, Apr. 29, 2024. [En línea]. Disponible en: <https://www.hitechbpo.com/blog/video-annotation-use-cases.php>

[29] T. Kumar, A. Mileo, R. Brennan, and M. Bendechache, "Image data augmentation approaches: A comprehensive survey and future directions," *arXiv preprint arXiv:2301.02830*, 2023. [En línea]. Disponible en: <https://arxiv.org/abs/2301.02830>

[30] R. Sapkota y M. Karkee, "Comparing YOLOv11 and YOLOv8 for instance segmentation of occluded and non-occluded immature green fruits in complex orchard environment", 26 de enero de 2025, arXiv: arXiv:2410.19869. doi: 10.48550/arXiv.2410.19869.

[31] A. S. Geetha, "YOLOv4: A Breakthrough in Real-Time Object Detection", 6 de febrero de 2025, arXiv: arXiv:2502.04161. doi: 10.48550/arXiv.2502.04161.

[32] M. Yaseen, "What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector", 28 de agosto de 2024, arXiv: arXiv:2408.15857. doi: 10.48550/arXiv.2408.15857.

[33] "YOLOv8 Improvements: Key Architectural Enhancements," *YOLOv8.org*. [En línea]. Disponible en: <https://yolov8.org/yolov8-improvements/>

[34] M. Hussain, "YOLOv5, YOLOv8 and YOLOv10: The go-to detectors for real-time vision," *arXiv preprint arXiv:2407.02988*, 2024. [En línea]. Disponible en: <https://arxiv.org/abs/2407.02988>

[35] Ultralytics, "Visión General de los Conjuntos de Datos de Segmentación de Instancias." [En línea]. Disponible en: <https://docs.ultralytics.com/es/datasets/segment>

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	9 - Bibliografía

- [36] "CustomTkinter Documentation." [En línea]. Disponible en: <https://customtkinter.tomschimansky.com/documentation/>
- [37] "COCO - Common Objects in Context." [En línea]. Disponible en: <https://cocodataset.org/#detection-eval>
- [38] ONNX Runtime." [En línea]. Disponible en: <https://onnxruntime.ai/docs/>
- [39] NVIDIA, "NVIDIA A100 Tensor Core GPU Architecture," White Paper, 2020. [En línea]. Disponible en: <https://images.nvidia.com/aem-dam/en-zz/Solutions/data-center/nvidia-ampere-architecture-whitepaper.pdf>
- [40] "Aprendizaje por Transferencia: Definición, ventajas y aplicaciones," *Ultralytics*. [En línea]. Disponible en: <https://www.ultralytics.com/es/glossary/transfer-learning>
- [41] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: 10.1109/TKDE.2009.191.
- [42] "PyTorch: El marco de aprendizaje profundo," *Ultralytics*. [En línea]. Disponible en: <https://www.ultralytics.com/es/glossary/pytorch>
- [43] "Precision and recall," *Wikipedia, The Free Encyclopedia*. May 24, 2025. [En línea]. Disponible en: [https://en.wikipedia.org/w/index.php?title=Precision\\_and\\_recall&oldid=1292053965](https://en.wikipedia.org/w/index.php?title=Precision_and_recall&oldid=1292053965)
- [44] "Precision and Recall in Machine Learning," *Roboflow Blog*. [En línea]. Disponible en: <https://blog.roboflow.com/precision-and-recall/>
- [45] Ultralytics, "YOLO Métricas de Rendimiento." [En línea]. Disponible en: <https://docs.ultralytics.com/es/guides/yolo-performance-metrics>
- [46] R. Padilla, S. L. Netto, and E. A. B. Da Silva, "A survey on performance metrics for object-detection algorithms," in *Proc. 2020 Int. Conf. Syst., Signals Image Process. (IWSSIP)*, Niterói, Brazil, Jul. 2020, pp. 237–242, doi: 10.1109/IWSSIP48289.2020.9145130.
- [47] R. Padilla, W. L. Passos, T. L. B. Dias, S. L. Netto, and E. A. B. Da Silva, "A comparative analysis of object detection metrics with a companion open-source toolkit," *Electronics*, vol. 10, no. 3, Art. no. 279, Jan. 2021, doi: 10.3390/electronics10030279.
- [48] R. Padilla, *Object-Detection-Metrics*. [Software]. May 28, 2025. [En línea]. Disponible en: <https://github.com/rafaelpadilla/Object-Detection-Metrics>

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	9 - Bibliografía

- [49] S. Abba, A. M. Bizi, J.-A. Lee, S. Bakouri, and M. L. Crespo, "Real-time object detection, tracking, and monitoring framework for security surveillance systems," *Helijon*, vol. 10, no. 15, Art. no. e34922, Aug. 2024, doi: 10.1016/j.helijon.2024.e34922.
- [50] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," in *Proc. 2016 IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 3464–3468, doi: 10.1109/ICIP.2016.7533003.
- [51] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *Proc. 2017 IEEE Int. Conf. on Image Process. (ICIP)*, Beijing, China, Sep. 2017, pp. 3645-3649, doi: 10.1109/ICIP.2017.8296962.
- [52] Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu, and X. Wang, "ByteTrack: Multi-object tracking by associating every detection box," *arXiv preprint arXiv:2110.06864*, 2021. [En línea]. Disponible en: <https://arxiv.org/abs/2110.06864>
- [53] N. Aharon, R. Orfaig, and B.-Z. Bobrovsky, "BoT-SORT: Robust associations multi-pedestrian tracking," *arXiv preprint arXiv:2206.14651*, 2022. [En línea]. Disponible en: <https://arxiv.org/abs/2206.14651>
- [54] J. Jiang, Z. Wang, M. Zhao, Y. Li, and D. Jiang, "SAM2MOT: A novel paradigm of multi-object tracking by segmentation," *arXiv preprint arXiv:2504.04519*, 2024. [En línea]. Disponible en: <https://arxiv.org/abs/2504.04519>
- [55] M. Bashar, S. Islam, K. K. Hussain, M. B. Hasan, A. B. M. A. Rahman, and M. H. Kabir, "Multiple object tracking in recent times: A literature review," *arXiv preprint arXiv:2209.04796*, 2022. [En línea]. Disponible en: <https://arxiv.org/abs/2209.04796>
- [56] "An application of centroid tracking algorithm," *Int. Res. J. Modernization Eng. Technol. Sci.*, vol. 2, no. 4, pp. 162-165, Apr. 2020. [En línea]. Disponible en: [https://www.irjmets.com/uploadedfiles/paper/volume2/issue\\_4\\_april\\_2020/868/1628083008.pdf](https://www.irjmets.com/uploadedfiles/paper/volume2/issue_4_april_2020/868/1628083008.pdf)
- [57] A survey on object detection and tracking algorithms," *Int. J. Eng. Res. Technol.*, vol. 9, no. 5, May 2020. [En línea]. Disponible en: <https://core.ac.uk/download/pdf/53189374.pdf>
- [58] J. Zhu, H. Yang, N. Liu, M. Kim, W. Zhang, and M. H. Yang, "Online multi-object tracking with dual matching attention networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Munich, Germany, Sep. 2018, pp. 366–382.
- [59] A. Bewley et al., "Multiple Object Tracking in Deep Learning Approaches: A Survey," *Electronics*, vol. 10, no. 19, p. 2406, Oct. 2021.

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	9 - Bibliografía

[60] Z. Wang, H. Zhang, Y. Wang, and L. Wang, "Multi-ship tracking by robust similarity metric," *arXiv preprint arXiv:2310.05171*, Oct. 2023. [En línea]. Disponible en: <https://arxiv.org/abs/2310.05171>

[61] J. Huang *et al.*, "DeconfuseTrack: Dealing with confusion for multi-object tracking," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, Jun. 2024.

[62] M. Lanan, "Spatiotemporal resource distribution and foraging strategies of ants (Hymenoptera: Formicidae)," *Myrmecological News*, vol. 20, pp. 53-70, Sept. 2014.

[63] Ultralytics, "Aislar objetos de segmentación." [En línea]. Disponible en: <https://docs.ultralytics.com/es/guides/isolating-segmentation-objects/>

[64] "Contours: Getting Started," *OpenCV 4.x Docs*. [En línea]. Disponible en: [https://docs.opencv.org/4.x/d4/d73/tutorial\\_py\\_contours\\_begin.html](https://docs.opencv.org/4.x/d4/d73/tutorial_py_contours_begin.html)

[65] S. Suzuki and K. Abe, "Topological structural analysis of digitized binary images by border following," *Comput. Vis. Graph. Image Process.*, vol. 30, no. 1, pp. 32–46, Apr. 1985, doi: 10.1016/0734-189X(85)90016-7.

[66] "Structural Analysis and Shape Descriptors," *OpenCV 4.x Docs*. [En línea]. Disponible en: [https://docs.opencv.org/4.x/d3/dc0/group\\_\\_imgproc\\_\\_shape.html#ga2c759ed9f497d4a618048a2f56dc97f1](https://docs.opencv.org/4.x/d3/dc0/group__imgproc__shape.html#ga2c759ed9f497d4a618048a2f56dc97f1)

[67] "Using Green's theorem to find area," *Math Insight*. [En línea]. Disponible en: [https://mathinsight.org/greens\\_theorem\\_find\\_area](https://mathinsight.org/greens_theorem_find_area)

[68] D. Römer, R. Exl, and F. Roces, "Two feedback mechanisms involved in the control of leaf fragment size in leaf-cutting ants," *J. Exp. Biol.*, vol. 226, no. 12, p. jeb244246, Jun. 2023, doi: 10.1242/jeb.244246.

[69] E. Greenwald, A. Friedman, and O. Feinerman, "Emergent regulation of ant foraging frequency through a distributed mechanism," *eLife*, vol. 12, p. e82182, Apr. 2023, doi: 10.7554/eLife.82182.

[70] F. Sabry, *Image Segmentation: Unlocking Insights through Pixel Precision*. One Billion Knowledgeable, 2024.

[71] M. Huang, X. Li, J. Hu, H. Peng, and S. Lyu, "Tracking multiple deformable objects in egocentric videos," in *Proc. 2023 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*,

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	9 - Bibliografía

Vancouver, BC, Canada, Jun. 2023, pp. 1461–1471, doi: 10.1109/CVPR52729.2023.00147.

[72] M. Pichler and F. Hartig, "Machine learning and deep learning—A review for ecologists," *Methods Ecol. Evol.*, vol. 14, no. 4, pp. 994–1016, Apr. 2023, doi: 10.1111/2041-210X.14061.

[73] A. Deviyani, "Assessing dataset bias in computer vision," ResearchGate, Prepr., 2021, doi: 10.13140/RG.2.2.19950.89924.

[74] W. Luo *et al.*, "Multiple object tracking: A review," *arXiv preprint arXiv:1409.7618*, 2014. [En línea]. Disponible en: <https://arxiv.org/abs/1409.7618>

[75] J. Ferryman and A. Ellis, "PETS2010: Dataset and Challenge," in *2010 7th IEEE International Conference on Advanced Video and Signal Based Surveillance*, 2010, pp. 143–150.

[76] C. R. Castro, "Seguimiento de múltiples personas considerando oclusión parcial y total en escenarios estacionarios no controlados," M.S. thesis, Inst. Nacional de Astrofísica, Óptica y Electrónica, Puebla, Mexico, 2014. [En línea]. Disponible en: <https://inaoe.repositorioinstitucional.mx/jspui/bitstream/1009/154/1/RetaCaC.pdf>

[77] L. Zhang *et al.*, "An anti-occlusion optimization algorithm for multiple pedestrian tracking," *PLoS One*, vol. 19, no. 1, p. e0291538, Jan. 2024, doi: 10.1371/journal.pone.0291538.

[78] B. Ristic, B.-N. Vo, D. Clark, and B.-T. Vo, "A metric for performance evaluation of multi-target tracking algorithms," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3452–3457, Jul. 2011, doi: 10.1109/TSP.2011.2140111.

[79] A. Milan, L. Leal-Taixé, I. Reid, S. Roth, and K. Schindler, "MOT16: A benchmark for multi-object tracking," *arXiv preprint arXiv:1603.00831*, 2016. [En línea]. Disponible en: <https://arxiv.org/abs/1603.00831>

[80] K. Bernardin and R. Stiefelhagen, "Evaluating multiple object tracking performance: The CLEAR MOT metrics," *EURASIP J. Image Video Process.*, vol. 2008, no. 1, p. 1-10, Jan. 2008, doi: 10.1155/2008/246309.

[81] Videmo Intelligente Videoanalyse GmbH & Co. KG, *pymot*. [Software]. Dec. 8, 2024. [En línea]. Disponible en: <https://github.com/Videmo/pymot>

[82] J. A. Sabattini, F. Sturniolo, M. Bollazzi, and L. A. Bugnon, "AntTracker: A low-cost and efficient computer vision approach to research leaf-cutter ants behavior," *Smart Agric. Technol.*, vol. 5, p. 100252, Oct. 2023, doi: 10.1016/j.atech.2023.100252.

 <b>UTN Regional Paraná</b>	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	9 - Bibliografía

[83] Y. Song, Z. Hu, T. Li, and H. Fan, "Performance evaluation metrics and approaches for target tracking: A survey," *Sensors*, vol. 22, no. 3, p. 793, Jan. 2022, doi: 10.3390/s22030793.

[84] E. P. Blasch y P. Valin, "Track Purity and Current Assignment Ratio for Target Tracking and Identification Evaluation".

[85] N. Imirzian *et al.*, "Automated tracking and analysis of ant trajectories shows variation in forager exploration," *Sci. Rep.*, vol. 9, no. 1, p. 13246, Sep. 2019, doi: 10.1038/s41598-019-49655-3.

[86] A. Godil, R. Bostelman, W. Shackleford, T. Hong, and M. Shneier, "Performance metrics for evaluating object and human detection and tracking systems," *Nat. Inst. Stand. Technol.*, Gaithersburg, MD, USA, NIST IR 7972, Jul. 2014, doi: 10.6028/NIST.IR.7972.

[87] M. Staniszewski *et al.*, "Application of crowd simulations in the evaluation of tracking algorithms," *Sensors*, vol. 20, no. 17, p. 4960, Sep. 2020, doi: 10.3390/s20174960.

[88] W. Ren, D. Kang, Y. Tang, and A. B. Chan, "Fusing crowd density maps and visual object trackers for people tracking in crowd scenes," in *Proc. 2018 IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 5353–5362, doi: 10.1109/CVPR.2018.00561.

[89] X. Cao, S. Guo, J. Lin, W. Zhang, and M. Liao, "Online tracking of ants based on deep association metrics: Method, dataset and evaluation," *Pattern Recognit.*, vol. 103, p. 107233, Jul. 2020, doi: 10.1016/j.patcog.2020.107233.

[90] J. Torres, "How to evaluate YOLOv8 model?" *YOLOv8.org*, Sep. 4, 2024. [En línea]. Disponible en: <https://yolov8.org/how-to-evaluate-yolov8-model/>

[91] T. Lucas, "Understanding the F1 score in machine learning: The harmonic mean of precision and recall," *Picsellia Blog*, May 25, 2023. [En línea]. Disponible en: <https://www.picsellia.com/post/understanding-the-f1-score-in-machine-learning-the-harmonic-mean-of-precision-and-recall>

[92] J. Torres, "Unveiling the secrets of YOLOv8 metrics," *YOLOv8.org*, Mar. 20, 2024. [En línea]. Disponible en: <https://yolov8.org/unveiling-the-secrets-of-yolov8-metrics/>

[93] Encord Team, "F1 score in machine learning," *Encord Blog*, Jul. 18, 2023. [En línea]. Disponible en: <https://encord.com/blog/f1-score-in-machine-learning/>

[94] L.Cheng, G.Li, S.Zhang YOLO-SAG: An improved wildlife object detection algorithm based on YOLOv8n

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	9 - Bibliografía

[95] A. Hernandez *et al.*, "Pytorch-wildlife: A collaborative deep learning framework for conservation," *arXiv preprint arXiv:2405.12930*, 2024. [En línea]. Disponible en: <https://arxiv.org/abs/2405.12930>

[96] S. D. Krishnan, N. Kavya, J. K. Mishra, G. S. Kant, K. Bhushan, and M. V. Sheela Devi, "Automating bird species identification with convolutional neural networks," *Nanotechnol. Percept.*, vol. 20, no. S16, pp. 529-536, 2024.

[97] IBM, "¿Qué es la segmentación de imágenes?" [En línea]. Disponible en: <https://www.ibm.com/es-es/topics/image-segmentation>

[98] N. O'Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. V. Hernandez, L. Krpalkova, D. Riordan, and J. Walsh, "Deep learning vs. traditional computer vision," in *Advances in Computer Vision*, L. C. Jain, H. H. C. Howlett, and R. J. Howlett, Eds. Cham, Switzerland: Springer, 2020, pp. 128–144. [En línea]. Disponible en: <https://arxiv.org/pdf/1910.13796.pdf>

[99] J. Ludwig, "Image convolution," Dept. of Elect. and Comput. Eng., Portland State Univ., Portland, OR, USA, Tech. Rep., 2007.

[100] "A complete walkthrough of convolution operations - viso.ai" \*Viso.ai\*, [En línea]. Disponible en: <https://viso.ai/deep-learning/convolution-operations/>. [Accedido: 5 de junio de 2025].

[101] G. A. Requena, "Sistema de seguimiento y análisis de comportamiento de hormigas mediante técnicas de visión por computador" [En línea]. Disponible en <https://hdl.handle.net/10630/30427>

[102] I. Apeinans, V. Tārauds, L. Litavniece, S. Kodors, y I. Zaremba, "ANT DETECTION USING YOLOV8: Evaluation of dataset transfer impact" \*Engineering for Rural Development (ETR)\*, vol. 2, pp. 34–37, Jun. 2024, doi: 10.17770/etr2024vol2.8040.

[103] M. Wu, X. Cao, y S. Guo, "Accurate detection and tracking of ants in indoor and outdoor environments" \*Animal Behavior and Cognition\*, Nov. 30, 2020, doi: 10.1101/2020.11.30.403816.

[104] N. Wang, S. Fu, Q. Rao, G. Zhang, y M. Ding, "Insect-Yolo: A New Method on Crop Insect Detection" \*SSRN\*, 2024, doi: 10.2139/ssrn.4813498.

[105] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, real-time object detection" in \*Proc. IEEE Conf. Comput. Vis. Pattern Recognit.\*, 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.

 UTN Regional Paraná	Utilización de Inteligencia Artificial para la estimación de la actividad forrajera de las hormigas cortadoras de hojas.
Ingeniería en Electrónica	9 - Bibliografía

- [106] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection" \*arXiv preprint\*, arXiv:2004.10934, 2020.
- [107] J. Delgado-Mata, I. Rojas, A. Guillén, and E. R. Miranda, "Real-time tracking system for ant foraging behavior analysis" \*J. R. Soc. Interface\*, vol. 19, no. 189, p. 20210923, 2022, doi: 10.1098/rsif.2021.0923.[108] A. Wilson, F. Arce, and J. Halloy, "Closed-loop experimental systems for insect behavior studies" \*Front. Robot. AI\*, vol. 8, p. 645283, 2021, doi: 10.3389/frobt.2021.645283.
- [109] A. Milan, L. Leal-Taixé, I. Reid, S. Roth, and K. Schindler, "Online multi-target tracking using recurrent neural networks" \*IEEE Trans. Pattern Anal. Mach. Intell.\*, vol. 39, no. 4, pp. 677–689, Apr. 2017, doi: 10.1109/TPAMI.2016.2592910.
- [110] N. Wojke and A. Bewley, "Deep cosine metric learning for person re-identification" in \*Proc. IEEE Winter Conf. Appl. Comput. Vis.\*, 2018, pp. 748–756, doi: 10.1109/WACV.2018.00087.
- [111] D. P. Mersch, A. Crespi, and L. Keller, "Tracking individuals shows spatial fidelity is a key regulator of ant social behavior" \*Science\*, vol. 340, no. 6136, pp. 1090–1093, May 2013, doi: 10.1126/science.1234316.
- [112] M. Das \*et al.\* , "ANT DETECTIVE: An automated approach for counting ants in densely populated images and gaining insight into ant foraging behavior" \*ResearchGate\*, Oct. 2024.
- [113] RAE y ASALE, "Imagen" \*Diccionario de la lengua española\* - Edición del Tricentenario. [En línea]. Disponible en: <https://dle.rae.es/imagen>.
- [114] "Specify ROI as Binary Mask - MATLAB & Simulink" \*MathWorks\*, [En línea]. Disponible en: <https://www.mathworks.com/help/images/specify-roi-as-binary-mask.html>.
- [115] J. Russ and F. B. Neal, \*The Image Processing Handbook\*, 7th ed. 2016.
- [116] J. E. Chiang Cruz, I. V. González, y J. R. Beltrán, "Estado del arte de algoritmos de Machine Learning para la detección de rupturas súbitas," *Ingeniería Hidráulica y Ambiental*, vol. 45, no. 2, pp. 78-87, Abr.-Jun. 2024.
- [117] PyTorch, "PyTorch," GitHub [Software]. [En línea]. Disponible en: <https://github.com/pytorch/pytorch>
- [118] P. Querejeta Simbeni, "Procesamiento digital de imágenes," Universidad Nacional del Sur, Bahía Blanca, Argentina, Julio 2015.