



Reconocimiento de elementos de laboratorio mediante algoritmos de visión por computadora

Alejandro Castaño Rojas
Angélica Arroyave Mendoza

Trabajo de grado presentado para optar al título de Ingenieros de Sistemas

Asesor

Luis Hernando Silva Flórez, Magíster (MSc) en Ingeniería de Sistemas.

Universidad de Antioquia
Facultad de Ingeniería
Ingeniería de Sistemas
Medellín, Antioquia, Colombia
2023

Referencia

Estilo IEEE (2020)

- [1] A. Arroyave Mendoza y A. Castaño Rojas, "Reconocimiento de elementos de laboratorio mediante algoritmos de visión por computadora", Trabajo de grado profesional, Ingeniería de Sistemas, Universidad de Antioquia, Medellín, Antioquia, Colombia, 2023.



Centro de Documentación Ingeniería (CENDOI)

Repositorio Institucional: <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - www.udea.edu.co

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

Dedicatoria

A nuestros familiares que nos dieron apoyo durante tantas horas de traspasar, esfuerzo, frustración y gozo; a nuestros amigos que nos resolvieron tantas dudas y nos acompañaron a través de un tema que no entendíamos, a nuestros docentes y tutores que hicieron crecer nuestro conocimiento mostrándonos nuevos temas y maneras de comprenderlos, edificando siempre sobre conocimientos previos y retando nuestra capacidad de entender el mundo con perspectivas cada vez más ingeniosas; a nuestras mascotas que estuvieron siempre a nuestro lado, acompañándonos en silencio. Este trabajo de grado, actuando como la cúspide de nuestro recorrido académico, va dedicado a todos y todas aquellas que lo hicieron posible y que hoy por hoy ven cómo estamos a tan solo un par de pasos de obtener nuestro título de pregrado.

Agradecimientos

Agradecimientos especiales al profesor Luis Hernando Silva que nos acompañó, entendió y nos guió con paciencia durante la concepción de esta idea con propuestas y maneras de ejecutarla.

TABLA DE CONTENIDO

SIGLAS, ACRÓNIMOS Y ABREVIATURAS	6
RESUMEN.....	7
ABSTRACT	8
I. INTRODUCCIÓN	9
II. PLANTEAMIENTO DEL PROBLEMA	10
III. JUSTIFICACIÓN.....	10
IV. OBJETIVOS	11
A. Objetivo general	11
B. Objetivos específicos	11
V. PROBLEMA DE INVESTIGACIÓN.....	12
VI. HIPÓTESIS.....	12
VII. MARCO TEÓRICO	12
VIII. METODOLOGÍA	14
VIII-A. Fases del proyecto	14
VIII-A1. Diseño y análisis	14
VIII-A2. Codificación y desarrollo.....	15
VIII-A3. Integración	15
VIII-A4. Validación.....	15

VIII-A5. Despliegue 16

IX RESULTADOS..... 16

X. DISCUSIÓN..... 18

XI. CONCLUSIONES 18

REFERENCIAS 19

LISTA DE FIGURAS

Figura 1: Arquitectura para una aplicación de reconocimiento de imágenes 12

Figura 2: Ciclo en la metodología SCRUM 14

Figura 3: Aplicación Inventory Recognition en versión web. 16

Figura 4: Aplicación Inventory Recognition en versión móvil..... 17

SIGLAS, ACRÓNIMOS Y ABREVIATURAS

YOLOv5	Red neuronal <i>You Only Look Once version 5</i>
UdeA	Universidad de Antioquia
Backend	Sistema que no es directamente accedido por el usuario y que es responsable de la manipulación y almacenamiento de los datos.
Frontend	Sistema con el que el usuario interactúa directamente desde un navegador de escritorio, una aplicación móvil o cualquier programa que pueda ser ejecutado en un dispositivo periférico.
OpenCV	Librería para problemas relacionados con visión por computadora, la manipulación de imágenes es su fuerte. Significa <i>Open Computer Vision</i> .
ML-Ops	<i>Machine Learning-Operations</i> . Paradigma usado en ingeniería de software para automatizar la creación, aseguramiento de la calidad y despliegue de artefactos de <i>Machine Learning</i> o Aprendizaje de Máquina.

RESUMEN

El laboratorio Movil-IS es un laboratorio adscrito al departamento de Ingeniería de Sistemas de la Universidad de Antioquia, es común hacer uso de sus instalaciones e implementos dadas sus características, sin embargo, el proceso de préstamo y devolución de implementos se puede dificultar al tener que disponer de una variedad tan amplia de insumos con los que los auxiliares administrativos no necesariamente están familiarizados, por lo que el objetivo de este trabajo de grado es automatizar el reconocimiento de objetos del inventario de Móvil IS de tal modo que el proceso de préstamo de implementos sea más ágil y menos propenso a errores. Por ende, el trabajo de facilitar el proceso de préstamo de insumos se propone que sea resuelto mediante un sistema que será dividido en dos: el frontend, que será la capa de interacción con el usuario, y el backend con el modelo integrado, que será la capa de procesamiento y transaccionalidad.

Palabras clave — Reconocimiento de objetos, visión por computadora, YOLO, YOLOv5, Móvil-IS, inventario, red neuronal.

ABSTRACT

The Movil-IS laboratory is affiliated with the Systems Engineering department at the University of Antioquia. It is common to use its facilities and equipment due to their characteristics, but the process of borrowing and returning equipment can be difficult due to the large variety of supplies that administrative assistants are not necessarily familiar with. Therefore, the goal of this graduation project is to automate the recognition of objects in the Movil-IS inventory in order to make the borrowing process more efficient and less prone to errors. To facilitate the supply lending process, it is proposed to be resolved through a system that will be divided into two: the frontend, which will be the user interaction layer, and the backend with the integrated model, which will be the processing and transactional layer.

***Keywords* — Object recognition, computer vision, YOLO, YOLOv5, Móvil-IS, inventory, computer vision, neural network.**

I. INTRODUCCIÓN

El laboratorio Móvil-IS es un laboratorio adscrito al departamento de Ingeniería de Sistemas de la Universidad de Antioquia en el cual se proveen espacios e implementos para la experimentación y desarrollo de sistemas de comunicación, de aplicaciones para dispositivos móviles y desarrollo de soluciones tecnológicas para problemas relacionados con el internet de las cosas (IoT). El personal autorizado para uso de los implementos del laboratorio son todos los miembros de la comunidad universitaria pertenecientes al departamento de Ingeniería de Sistemas, Ingeniería Electrónica y de Ingeniería de Telecomunicaciones. Este laboratorio cuenta con procesos operativos del día a día que podrían ser automatizados para facilitar la rotación de personal administrativo sin que requieran de conocimiento especializado para llevar a cabo sus funciones de préstamo de insumos del laboratorio a estudiantes. Este trabajo de grado propone una forma de automatizar el reconocimiento de objetos para préstamo.

II. PLANTEAMIENTO DEL PROBLEMA

El funcionamiento del laboratorio depende de la labor de los auxiliares administrativos y auxiliares de programación. El proceso de acceso a los implementos disponibles en el laboratorio es a través de un protocolo bastante sencillo que consta en solicitar el implemento necesario al auxiliar administrativo en turno, verificar la disponibilidad de los objetos solicitados, y si éstos están disponibles se pueden prestar con la condición de dejar la TIP (Tarjeta de Identificación Personal) durante todo el tiempo que se usen; al devolver los implementos, el auxiliar en turno debe verificar que se entregue la totalidad de los objetos y el estado de los mismos, luego, registrar que están disponibles. Todo este proceso se hace de manera manual, ayudados por un registro en una hoja de cálculo (por ejemplo, en Google Sheets) donde se anota toda la información relacionada con el préstamo y si fue devuelto lo prestado o no. Con lo anterior expuesto, se puede pensar que hay pasos propensos a errores humanos de los auxiliares administrativos, se pueden formular algunas preguntas que ayudan a clarificar la problemática, por ejemplo: ¿Es claro para un auxiliar saber cuántos y cuáles objetos deben ser devueltos si fueron prestados en un turno diferente al suyo? ¿Sabe un auxiliar novato cómo se ve un objeto que le soliciten?

III. JUSTIFICACIÓN

Se puede pensar en un procedimiento más riguroso para el proceso de préstamo, añadiendo una capa de automatización que pueda ayudar con el reconocimiento de objetos y el inventario de los mismos. Por tanto, este trabajo de grado se enfoca en resolver estos retos de automatización presentados con ayuda de técnicas de visión computacional orientadas a reconocer objetos y de construcción de sistemas tradicionales de inventarios.

IV. OBJETIVOS

A. Objetivo general

Construir un sistema que permita a los auxiliares del laboratorio Movil-IS facilitar el proceso de entrega y recuperación de elementos o accesorios adscritos al mismo, mediante un proceso automatizado apoyado por técnicas de reconocimiento de objetos.

B. Objetivos específicos

- Analizar un método o técnica que permita reconocer elementos adscritos al laboratorio Movil-IS: por ejemplo, podría ser por comparación o algoritmo de machine learning.
- Diseñar un sistema de reconocimiento de los objetos de Movil-IS con base en el resultado obtenido del objetivo anterior.
- Implementar un sistema backend y un sistema frontend con el que puedan interactuar los auxiliares y hacer uso de este modelo.
- El sistema desarrollado puede ser probado.

V. PROBLEMA DE INVESTIGACIÓN

Encontrar en la literatura uno o varios algoritmos que puedan solucionar el problema de reconocimiento de objetos específicos en fotografías, además de buscar si existen implementaciones aptas para realizar pruebas de concepto en el laboratorio Móvil-IS.

VI. HIPÓTESIS

Existe un algoritmo que es capaz de reconocer objetos del laboratorio Móvil-IS y este puede ser utilizado desde una aplicación web.

VII. MARCO TEÓRICO

Para crear un sistema capaz de cumplir con los objetivos propuestos, hicimos uso de algoritmos de detección o reconocimiento de objetos[1] que utilizan conceptos de visión por computadora[2] y procesamiento digital de imágenes [3] para detectar un objeto en una imagen digital. Estos algoritmos de detección suelen ir acompañados por técnicas de machine learning utilizadas para clasificación como la técnica de k-means o redes neuronales. Finalmente, para que este sistema cumpla su objetivo de ser usado por personal no técnico, se hizo uso de una integración de un sistema frontend (interfaz de cara al usuario) con un sistema backend (capa de procesamiento), ver Figura 1.

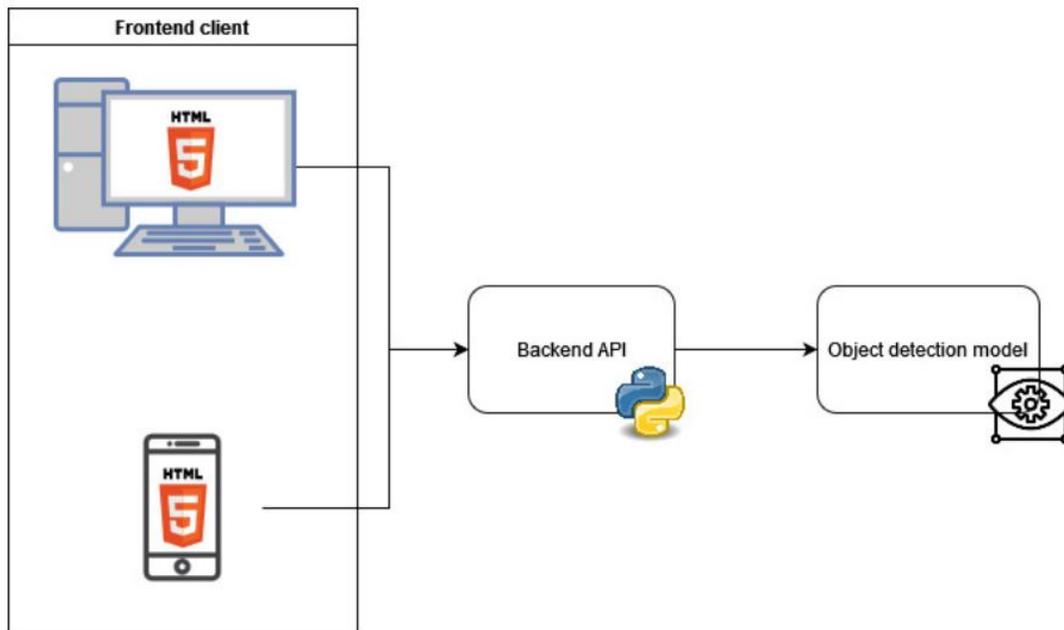


Figura 1: Arquitectura para una aplicación de reconocimiento de imágenes

Para el reconocimiento de los elementos del laboratorio se usó un algoritmo de detección de objetos llamado YOLOv5 (You Only Look Once) [4] entrenado con imágenes tomadas de los elementos de Movil-IS; los instrumentos involucrados en el proceso fueron los servidores del LIS, que es donde es posible actualmente desplegar el backend con el respectivo modelo de detección de objetos y desde donde se pueden servir los contenidos estáticos para que el frontend pueda ser cargado y utilizado en los dispositivos de los auxiliares, bien sea el computador administrativo o sus celulares conectados a la red de Movil-IS.

VIII. METODOLOGÍA

Usamos bases de la metodología ágil para iterar sobre el desarrollo del sistema, la metodología escogida es scrum [5] (ver Figura 2); para el desarrollo del trabajo usaremos instrumentos como computadores y celulares para la implementación y pruebas del sistema. Además de que usamos imágenes de objetos que actualmente se prestan en el laboratorio Movil-IS para entrenar el modelo de detección de objetos.



Figura 2: Ciclo en la metodología SCRUM

VIII-A. Fases del proyecto

Este proyecto contó con 5 fases dentro de cada sprint (o ciclo de desarrollo), siendo la fase inicial la fase de “análisis y diseño” que permitió estimar los tiempos necesarios para desarrollar una funcionalidad particular y qué nivel de prioridad se tuvo para alcanzar los objetivos propuestos en este trabajo.

VIII-A1. Diseño y análisis

- Definición de la arquitectura que permita una alta escalabilidad y un bajo acoplamiento (ver Figura 1)
- Elicitación e implementación de los requisitos funcionales y no funcionales del sistema.

- Requisitos funcionales:
 - El usuario puede cargar una imagen desde su computadora.
 - El sistema procesa la imagen ingresada por el usuario y devuelve la misma con los respectivos elementos del laboratorio detectados en esta.
- Requisitos no funcionales:
 - El sistema tiene una disponibilidad de más del 90\% de las veces en que un usuario intenta usarlo.
 - El sistema tiene una interfaz de usuario amigable para facilidad de uso por personal no técnico.
 - El sistema hace uso de contenedores para ejecutarse en cualquier servidor independiente del sistema operativo.
- Implementación de patrones arquitectónicos que permitan la mantenibilidad del sistema.
 - El patrón arquitectónico utilizado fue un modelo de dos capas siendo cada una de estas el backend y el frontend.

VIII-A2. Codificación y desarrollo

- Implementación de una interfaz amigable con el usuario que le permita realizar el correcto uso de la plataforma intuitivamente.
- Definición e implementación de criterios de aceptación.
 - El usuario debe poder subir una imagen.
 - El usuario no debe poder subir otro archivo diferente a una imagen.
 - El sistema debe recibir la imagen del usuario y entregar otra imagen como respuesta.
- Creación de una estructura del proyecto basada en la modularización por componentes.
- Entrenamiento de la red neuronal con un conjunto de imágenes provisto por personal del laboratorio.
 - Se recolectaron imágenes con ayuda del personal del laboratorio.
 - El conjunto de imágenes tuvo que ser etiquetado y aumentado con técnicas de sobremuestreo para alcanzar una amplia variedad de imágenes para el entrenamiento de la red.

VIII-A3. Integración

Integración del modelo con la estructura del backend y este a su vez con el frontend. Esto es posible mediante protocolo HTTP [6], dado que el Frontend solo debe hacer una sola petición para obtener sus resultados esto puede ser logrado con un método POST.

VIII-A4. Validación

Realizar pruebas en el sistema para garantizar en todo momento el correcto funcionamiento del mismo.

VIII-A5. Despliegue

El despliegue y aprovisionamiento del sistema es agnóstico al sistema operativo por la naturaleza de desarrollarlo en contenedores que permitan su fácil distribución.

IX RESULTADOS

Se desarrolló un sistema para el apoyo de los auxiliares del laboratorio Móvil-IS en su labor de entrega, recuperación e inventario de implementos del laboratorio con los cuales no necesariamente están familiarizados. Para esto se desarrolló una aplicación frontend en React, un backend en Python con FastAPI y un modelo de aprendizaje de máquina (machine learning) con la red neuronal YOLOv5 que hace uso de OpenCV para procesar las imágenes.

- La aplicación desarrollada en su interfaz web, ver Figura 3.

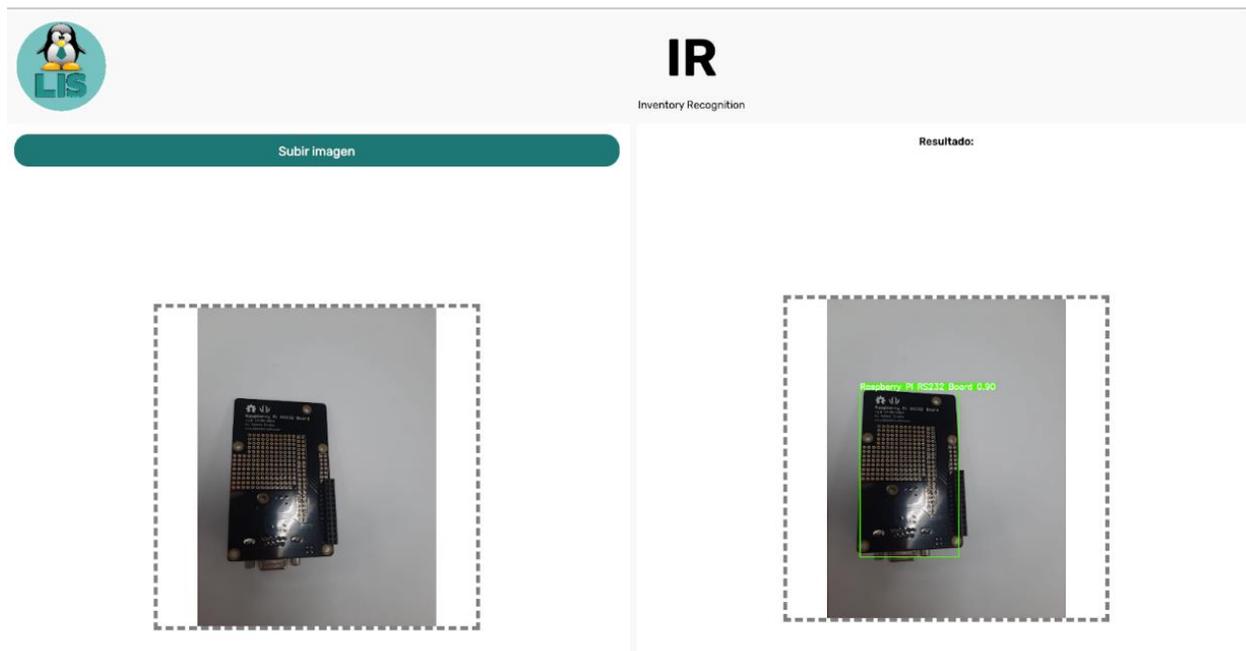


Figura 3: Aplicación Inventory Recognition en versión web.

- La aplicación desarrollada en su interfaz móvil, ver Figura 4.

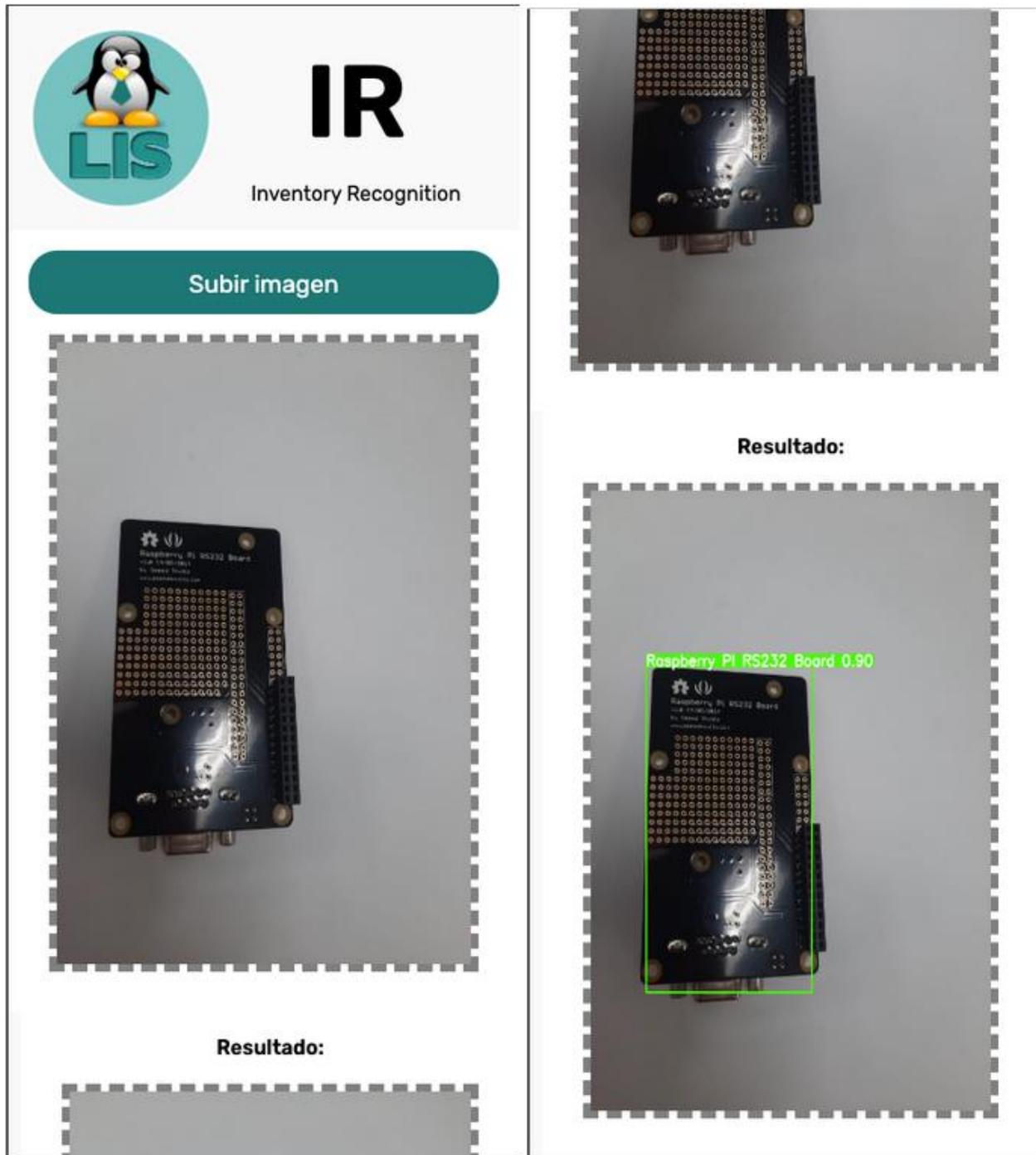


Figura 4: Aplicación Inventory Recognition en versión móvil.

X. DISCUSIÓN

Los resultados logrados en este proyecto muestran que es posible construir un sistema de apoyo para inventarios con objetos específicos y poco comunes como lo son los insumos de un laboratorio de casi cualquier índole, en nuestro caso de un laboratorio de insumos electrónicos para estudiantes de ingeniería de sistemas, telecomunicaciones y electrónica. La tecnología desarrollada para el reconocimiento de imágenes a menudo se muestra como un logro científico que es difícil de replicar para uso general, de tal modo que dicha tecnología no interactúa con más sistemas más allá de los usados por los investigadores para publicar sus resultados, sin embargo, esta aplicación hace uso de un modelo llamado YOLOv5 para identificar objetos. Esta red neuronal es poderosa y los usos más comunes que se encontraron para esta red fue identificando objetos en videos en tiempo real, no obstante requiere de una vasta cantidad de imágenes para poder reconocer un objeto con alta precisión sin confundirse, en el futuro pueden explorarse otros modelos que requieran una menor cantidad de imágenes para identificar un objeto, además de extender las funcionalidades de la aplicación provista para que se puedan agregar objetos del inventario de este y otros laboratorios de forma simple, de tal modo que esto pueda convertirse en una herramienta de uso general entre diversos laboratorios de la facultad o la universidad.

XI. CONCLUSIONES

- Es posible construir un sistema que sea capaz de reconocer objetos en imágenes mediante redes neuronales y que sea de fácil acceso a usuarios no técnicos.
- El entrenamiento del modelo no es trivial; el personal no técnico podría esperar que el modelo sea capaz de reconocer un objeto adicional tan solo adicionando a su base de datos una imagen y esto no es posible dado el avance tecnológico en este campo al momento de escribir este trabajo, ya que en la literatura no existen modelos maduros basados en “*One-shot learning*” que puedan dar los resultados que son obtenibles con este tipo de modelos como el usado previamente (YOLOv5)
- Aún faltan labores inherentes al ciclo de vida de desarrollo de software para que el sistema esté desplegado en un ambiente comercial, tales como definir cuál sería el flujo de datos que debería tener el sistema para que siempre esté actualizado con todos los objetos disponibles en un inventario de magnitud comercial. Este tipo de objetivos requieren un trabajo de ML-Ops.
- El presente sistema puede servir como piedra angular para apoyar la automatización al interior de la Universidad de Antioquia y demás laboratorios académicos en el territorio nacional.

REFERENCIAS

- [1] K. a. L. B. Grauman, «Visual object recognition,» *Synthesis lectures on artificial intelligence and machine learning*, vol. 5, pp. 1--181, 2011.
- [2] L. G. a. S. G. C. a. o. Shapiro, *Computer vision*, Prentice Hall New Jersey, 2001.
- [3] K. R. Castleman, *Digital image processing*, Prentice Hall Press, 1996.
- [4] Y. a. S. Y. a. W. Z. Zhao, «The Improved YOLOV5 Algorithm and Its Application in Small Target Detection,» de *International Conference on Intelligent Robotics and Applications*, 2022.
- [5] P. a. B. G. a. L. C. a. V. B. Deemer, «Información básica de SCRUM,» *California: Scrum Training Institute*, 2009.
- [6] T. a. F. R. a. F. H. Berners-Lee, «Hypertext transfer protocol--HTTP/1.0,» 1996.