

# Reporte de Proyecto Grupal U3

## Clasificación de enfermedades en Cítricos

Jonathan Canales Puga\*  
Jorge Luis Charles Torres\*  
Christopher Emmanuel Perez Duque\*  
Humberto Erubiel Ortega Lujano\*  
\*Ingeniería en Tecnologías de la Información  
Universidad Politécnica de Victoria

**Resumen**—En el presente documento se expone de manera clara y ordenada el procedimiento que se llevó a cabo para la realización de un proyecto grupal destinado para la asignatura de Sistemas Inteligentes.

A grandes rasgos, dicho proyecto consta de un programa el cual es capaz de determinar que tipo de enfermedad presenta la hoja (cítrica) que aparece en la imagen que se le cargue. Es importante mencionar que el programa solo puede detectar 12 tipos de enfermedades, las cuales están presentes en el estado de Tamaulipas.

Cabe recalcar que el proyecto se desarrolló en el lenguaje de programación Python y tanto la entrada así como la salida del programa se realiza a través de la interfaz gráfica[1].

### I. INTRODUCCIÓN

Este documento presenta el desarrollo y resultados de un proyecto grupal llevado a cabo como parte de la asignatura de Sistemas Inteligentes. El proyecto consiste en un programa en lenguaje de programación Python usando librerías de pyqt5[2], numpy[3] y openCV que procesan imágenes de hojas de árboles cítricos para determinar qué tipo de enfermedad presentan, con especial atención en las enfermedades comunes en el estado de Tamaulipas las cuales son doce esto se sabe gracias a una investigación realizada por personal de la Universidad Autónoma de Tamaulipas (UAT)[4].

Entre las enfermedades[5] el objetivo de este sistema de clasificación se incluyen el estado saludable de los cítricos, así como diversas condiciones patológicas como la infección por HLB, manchas grasas, deficiencias nutricionales de hierro, magnesio, manganeso, nitrógeno y zinc, así como la presencia de ácaros y otras afecciones.

El objetivo de este proyecto es el de identificar que enfermedad de las 12 que hay en tamaulipas es la que tiene la hoja introducida en el programa. La entrada y salida del programa se realiza a través de una interfaz gráfica

### II. DESARROLLO EXPERIMENTAL

Después de la asignación del proyecto por parte del profesor, lo primero que se realizó fue una investigación con fin de comprender el proyecto asignado, además de comprender las librerías que fueron comentadas con el profesor la librería **opencv** [6] será de gran ayuda para la realización de este sistema.

A su vez se recopilamos imágenes mediante un dataset proporcionado por la UAT que contenía imágenes afectadas por diversas enfermedades presentes en la región de Tamaulipas. También investigamos y agregamos un 30 % mas de imágenes de dichas enfermedades para tener un dataset mas robusto. Estas imágenes se utilizaron para entrenar y evaluar el modelo de clasificación. Las imágenes se organizaron en un conjunto de datos etiquetado que contenía ejemplos de cada categoría de enfermedad.

Se implementó una interfaz gráfica de usuario utilizando la biblioteca PyQt5[7]. La interfaz permitía a los usuarios cargar imágenes desde su sistema de archivos y procesarlas para la clasificación.

Se entrenó un modelo de clasificación utilizando algoritmos de aprendizaje automático. Se utilizó el conjunto de datos etiquetado para entrenar el modelo. El modelo asignaba una categoría de enfermedad a las imágenes cargadas en función de las características extraídas.

Posteriormente, se comenzó a buscar una serie de imágenes[8] las cuales serán las encargadas de comprobar si el sistema es funcional y cumple con las características asignadas.

Cabe mencionar que mientras se iba investigando más y generando posibles avances pero ninguno funciona al adecuadamente se comprendió que pyesseract y opencv tienen un límite de caracteres para procesar en este caso no se muestra correctamente los caracteres.

Se evaluó el rendimiento del modelo utilizando métricas como precisión. Se probaron diferentes configuraciones del modelo y se realizaron ajustes para mejorar la precisión de la clasificación. Se mostraban los resultados al usuario en la interfaz gráfica, indicando la enfermedad detectada y el porcentaje de efectividad

Se realizaron pruebas exhaustivas utilizando conjuntos de datos de prueba separados para verificar la robustez y la generalización del modelo. Se realizaron pruebas con imágenes tanto de cítricos afectados por enfermedades como de cítricos sanos para asegurarse de que el modelo fuera capaz de diferenciar con precisión entre las clases.

El programa cuenta con lo siguiente

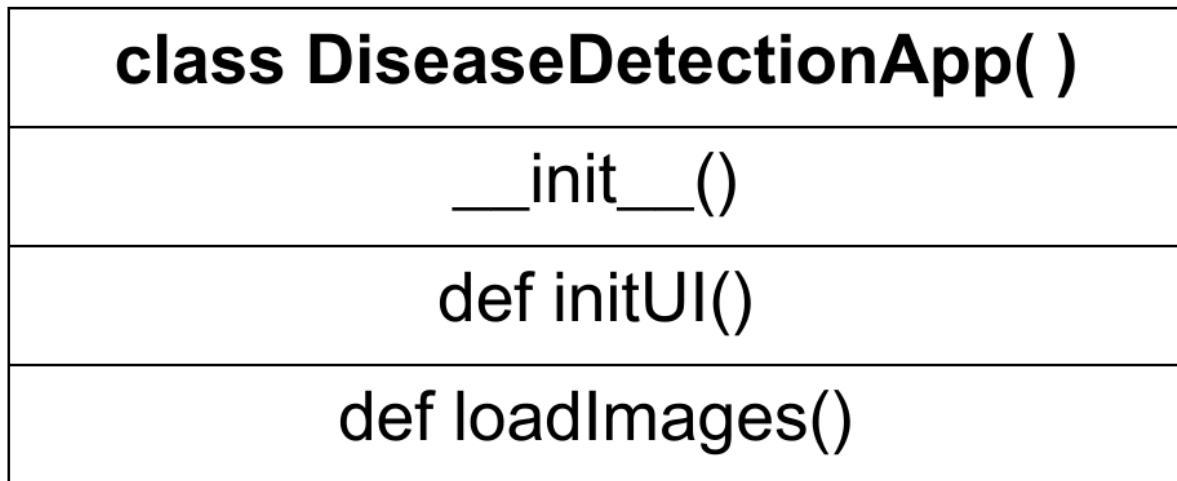


Figura 1: El diagrama de clase del programa.

**class DiseaseDetectionApp:** Utiliza una interfaz gráfica para cargar imágenes de hojas de cítricos, predecir enfermedades en esas imágenes utilizando modelos de aprendizaje automático y mostrar los resultados en una tabla junto con el porcentaje de efectividad de las predicciones.

- **def \_\_init\_\_():** Se encarga de inicializar varios atributos importantes de la clase DiseaseDetectionApp antes de configurar la interfaz de usuario y mostrar la ventana de la aplicación. Estos atributos se utilizan para rastrear estadísticas relacionadas con las predicciones de enfermedades en hojas de cítricos.
- **def initUI():** Configura y organiza la interfaz gráfica de la ventana de la aplicación. Crea un botón para cargar imágenes, una tabla para mostrar resultados y una etiqueta para mostrar el porcentaje de efectividad. Estos elementos se organizan en un diseño vertical que define cómo se mostrarán en la ventana.
- **def loadImages() :** Su función es cargar las imágenes desde la carpeta seleccionada por el usuario, realizar predicciones de enfermedades en esas imágenes utilizando modelos preentrenados y mostrar los resultados en una tabla junto con el porcentaje de efectividad de las predicciones.

### III. RESULTADOS

El desarrollo del programa fue una tarea muy complicada debido a que se solicitó utilizar herramientas que no ofrecen los mejores resultados para este tipo de aplicaciones, a pesar de esto, después de varios intentos se llegó a una versión lo suficientemente buena para este proyecto.

Una vez que el programa ya mostraba resultados alentadores, se realizaron diferentes pruebas para ver su comportamiento y observar si los resultados que mostraba eran buenos. Dichas pruebas se realizaron con el dataset que contiene las enfermedades que existen en los cítricos en

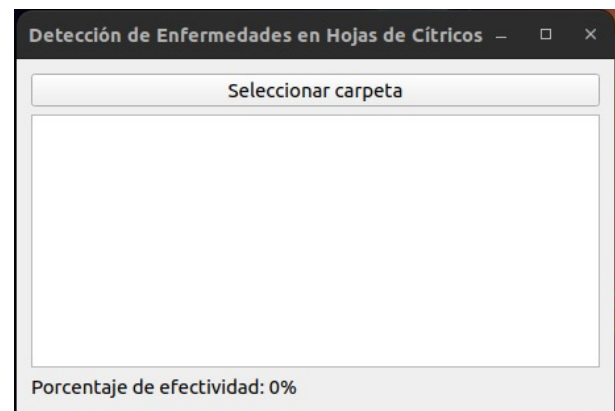


Figura 2: interfaz inicial

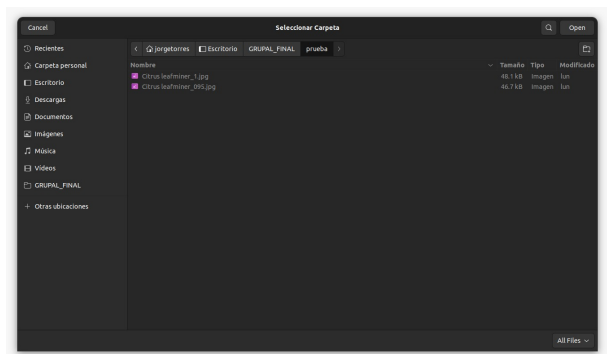


Figura 3: seleccionar carpeta

tamaulipas. Dichas pruebas mostraban muy buenos resultados en la identificación de las enfermedades de los cítricos por dar un ejemplo de 20 imágenes con enfermedades distintas tenía una tasa de error no mayor a 6 erradas.

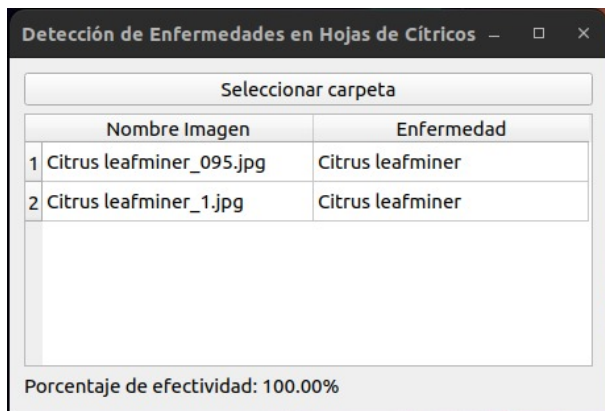


Figura 4: Efectividad

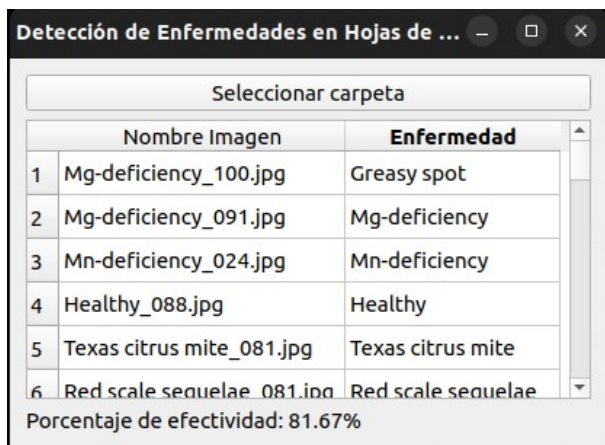


Figura 5: Prueba Con todas las enfermedades

**NOTA:** para que el porcentaje de efectividad se muestre de manera correcta se debe de poner el siguiente formato al nombre de la imagen [enfermedad\_numero]

A continuación, en la figura 2 se muestran algunas capturas de pantalla del proceso que el usuario debe llevar a cabo para utilizar correctamente el programa.

Aquí se muestra la selección de una carpeta para poder cargarla al programa para que así este tenga las imágenes que el usuario quiere procesar.

Después de elegir la carpeta que contiene las imágenes a examinar, la fase de análisis se inicia automáticamente, y los resultados se presentan de manera organizada en forma de tabla mostrando el porcentaje de efectividad.

En esta imagen se muestra el porcentaje de efectividad que obtuvimos al momento de ejecutar el programa cargando imágenes de las 12 enfermedades presentes en Tamaulipas. Obteniendo un porcentaje del 81.67 %.

#### MATRIZ DE CONFUSION

Matriz de confusión de prueba con el 81 % de efectividad

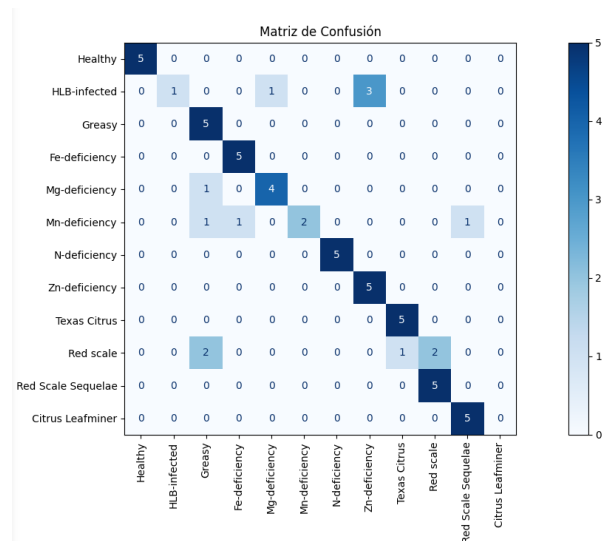


Figura 6: Matriz

#### IV. CONCLUSIÓN

En este trabajo se presentó la implementación de la librería PyQt5 para crear una interfaz gráfica de usuario para detectar en las hojas de los cítricos enfermedades que se encuentran en el estado de Tamaulipas.

Desarrollar este proyecto en específico resultó un poco complicado, ya que este tipo de programas requieren de un buen equipo de cómputo debido a la alta demanda de recursos que se necesitan para llevar a cabo todo el proceso de clasificación. Además, encontrar el dataset con las imágenes de las hojas no fue tarea sencilla, de no ser por el maestro de la asignatura y su ayuda para obtener el dataset empleado para una investigación de la Universidad Autónoma de Tamaulipas, hubiésemos tenido que recolectar imagen por imagen de diferentes fuentes hasta juntar una cantidad considerable. A pesar de todo esto, se llegó a un programa que funciona considerablemente bien bajo óptimas condiciones.

#### REFERENCIAS

- [1] Alberto Cuevas Álvarez. *Interfaces gráficas en Python con PyQt*. Anonimus, 10 de julio de 2020.
- [2] kiranpanigrahi. *Python Binding*. Tutorial Points, 2020.
- [3] numpy. <https://aprendeconalf.es/docencia/python/manual/numpy/>. Consultado el 23-08-2023.
- [4] UAT. <https://www.tamps.cinvestav.mx/~wgomez/uatdata/set.html>. Consultado el 23-08-2023.
- [5] THE AMERICAN PHYTOPATHOLOGICAL SOCIETY. *Plagas y enfermedades de los cítricos*. Ediciones Mundi-Prensa, 10 de julio de 2015.
- [6] opencv. <https://omes-va.com/leerimagen/>. Consultado el 23-08-2023.
- [7] anonymous. *Reference Guide for PyQt5*. Appendix, 2020.
- [8] Citrus Leaves. <https://www.kaggle.com/datasets/dtrilsbek/citrus-leaves-prepared/>. Consultado el 23-08-2023.