



Universidad
Internacional
de Valencia

Desarrollo de un sistema de diagnóstico de enfermedades en hojas de tomate mediante modelos de aprendizaje profundo

Titulación:
Máster en Big Data y
Ciencia de Datos
Curso Académico
2024-2025

Alumno/a: Marín Lucas,
Rubén
DNI: 07272889-J
Director/a del TFT:
Ricardo Lebrón Aguilar

Convocatoria:

SEGUNDA

Índice general

Índice de figuras	1
Índice de cuadros	1
1 Introducción	2
1.1 Motivación	2
1.2 Estructura del resto del documento	4
2 Objetivos	5
2.1 Objetivos específicos	5
3 Estado del arte	6
4 Implementación y desarrollo	7
5 Evaluación y resultados	8
6 Conclusiones	9
A Anexo I: Ejemplo de anexo	10
B Anexo II: Otro ejemplo de anexo	11

Índice de figuras

1.1	<u>Origen del tomate</u>	2
1.2	<u>Tizón tardío en una planta de tomate</u>	3
1.3	<u>Tizón temprano en una planta de tomate</u>	3

Índice de cuadros

1.1	<u>Top 20 países productores de tomates 2022</u>	3
-----	--	---

Resumen

Lorem ipsum (RESUMEN)

Palabras clave: primero, segundo, tercero

Agradecimientos

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1. Introducción

1.1. Motivación

Tomate o tomatara (*Solanum lycopersicum*) es una planta herbácea de la familia Solanaceae cultivada en todo el mundo para el cultivo de su fruto, el tomate o jitomate, uno de los ingredientes más universales de ensaladas y salsas en el mundo entero. [7]

Según los últimos estudios filogenéticos, la planta silvestre de la cual surge el tomate doméstico actual tiene origen en la zona andina del norte de Perú y sur de Ecuador. Su domesticación y diversificación posterior se originó en México.

Los pueblos aztecas y mayas lo usaban en su cocina y fue exportado al resto del mundo a partir de la llegada de los españoles que lo distribuyeron a lo largo de sus colonias en el Caribe y la península ibérica a partir de lo cual pudo llegar al resto de Europa. También lo llevaron a Filipinas y de allí pudo entrar al continente asiático. [2]

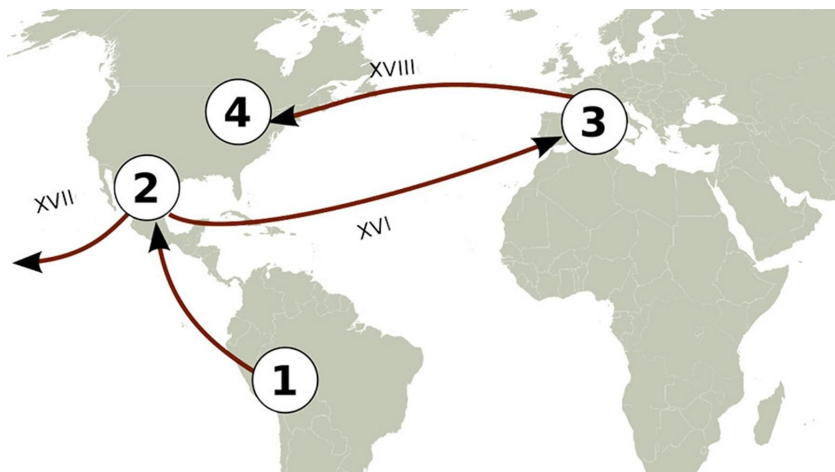


Figura 1.1: Origen del tomate

El tomate es la hortaliza más extendida mundialmente y la de mayor valor económico. Su demanda aumenta continuamente y con ella su producción y comercio.

La producción mundial de tomate ascendió a más de 186 millones de toneladas en 2022 según los datos de la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO). Según esta misma organización esta es la evolución de los 20 países que más han producido hasta 2022: [6]

Como ya se ha mencionado, el cultivo de tomate es uno de los cultivos hortícolas más importantes a nivel mundial. Sin embargo, su producción se ve amenazada por una amplia variedad de enfermedades causadas por hongos, bacterias, virus y nematodos. Estas enfermedades pueden provocar una bajada de rendimiento que van desde reducciones parciales hasta la pérdida completa de la cosecha.

Cuadro 1.1: Top 20 países productores de tomates 2022

Titulo				
País	2000	2010	2020	2022
China	22 200	46 760	64 680	68 242
...

Entre las enfermedades más comunes se encuentran:

- Tizón tardío (*Phytophthora infestans*): Puede destruir por completo una plantación si no se controla a tiempo, especialmente en condiciones húmedas y templadas.



Figura 1.2: Tizón tardío en una planta de tomate

- Tizón temprano (*Alternaria solani*): Produce defoliación progresiva, debilitando la planta y reduciendo el número y calidad de los frutos.



Figura 1.3: Tizón temprano en una planta de tomate

- Fusariosis vascular (*Fusarium oxysporum*): Ataca el sistema vascular, provocando marchitez y muerte de plantas.

- Virus como TYLCV y TSWV: Pueden causar deformaciones severas y reducciones completas en la producción, especialmente cuando se transmiten por vectores como la mosca blanca.

La manifestación simultánea o sucesiva de estas enfermedades es una de las principales causas en la disminución en la productividad del cultivo a escala global. Además, muchas de estas enfermedades no solo viven en la planta sino que persisten en el suelo, semillas o herramientas que hayan interactuado con la planta, lo que dificulta su erradicación y aumenta los costos del tratamiento. [3]

Dada la magnitud del impacto de estas enfermedades, la detección temprana y precisa de las mismas es crucial. Permite una correcta intervención que minimiza las pérdidas, permitiendo la reducción del uso innecesario de los agroquímicos y mejorando la sostenibilidad. En este contexto, las tecnologías basadas en visión por computadora, sensores remotos e inteligencia artificial ofrecen soluciones eficaces para mejorar el seguimiento y el control sanitario de este cultivo clave.

1.2. Estructura del resto del documento

La documentación de este proyecto se ha desarrollado dividiendo el contenido en distintos capítulos con el objetivo de facilitar la comprensión por parte del lector. A continuación se enuncia la información presente en cada capítulo del resto del documento:

- Capítulo 2. Objetivos: se describe el objetivo principal del proyecto junto con los objetivos intermedios necesarios para conseguirlo.
- Capítulo 3. Estado del Arte:
- Capítulo 4. Implementación y desarrollo: se explica el preprocesamiento de los datos y las técnicas ML aplicadas a los mismos.
- Capítulo 5. Evaluación y resultados: se presentan los resultados obtenidos al aplicar las técnicas ML.
- Capítulo 6. Conclusiones

2. Objetivos

El objetivo general de este proyecto consiste en conseguir un clasificador que a partir de imágenes de hojas de plantas de tomate distinga entre estado saludable y 10 enfermedades distintas.

2.1. Objetivos específicos

1. Analizar dataset de hojas de tomate.
2. Implementar y entrenar modelos CNN para la clasificación.
3. Evaluar la precisión de los modelos y comparar resultados.

3. Estado del arte

En los últimos años la aplicación de técnicas de inteligencia artificial en la agricultura ha cobrado un papel relevante, especialmente en tareas de diagnóstico temprano de enfermedades en cultivos. El uso de aprendizaje profundo permite automatizar la detección de patrones en imágenes, lo cual puede ayudar a los agricultores a tomar decisiones más rápidas y eficientes.

Inicialmente, los métodos empleados para esta tarea incluían algoritmos de aprendizaje supervisado como máquinas de vectores de soporte (SVM), kvecinos más cercanos (KNN) y redes bayesianas. Sin embargo, estos enfoques dependían en gran medida de una segmentación previa precisa y de la extracción manual de características, lo que limitaba su capacidad de generalización y precisión en entornos reales.

Con la llegada de las redes neuronales convolucionales (CNN), se ha producido un cambio significativo en la forma de abordar este problema. Las CNN son capaces de aprender representaciones directamente a partir de los datos de imagen, eliminando la necesidad de ingeniería manual de características. Diversos estudios han demostrado su eficacia para la clasificación de enfermedades en hojas de tomate.

Por ejemplo, una revisión publicada en la Revista de Investigación e Innovación de las Ciencias de la Universidad Tecnológica de Bolívar [4], las técnicas tradicionales de aprendizaje supervisado como SVM, KNN y lógica difusa muestran limitaciones significativas en tareas de detección de enfermedades en imágenes de frutas debido a su dependencia de extracción manual de características y segmentación previa. En contraste, las redes neuronales convolucionales han demostrado una precisión superior, mayor robustez frente a la variabilidad y mayor capacidad de generalización. Esta revisión respalda la elección de CNNs como enfoque principal en este trabajo.

Por otra parte, Valeria Maeda Gutiérrez (2019) [5] realizó una comparativa entre varias arquitecturas CNN, incluyendo AlexNet, GoogleNet, InceptionV3, ResNet 18 y ResNet 50 aplicadas al conjunto de datos PlantVillage. Todas las arquitecturas consiguieron más del 98 % de precisión y sensibilidad, lo que confirma la idoneidad de las mismas para la tarea que se pretende hacer. Concretamente con GoogleNet consiguió una precisión del 99,3 % y una sensibilidad del 99,1 %

En otra línea, Eduardo A. Huerta-Mora, Víctor González-Huitrón, Héctor Rodríguez-Rangel y Leonel Ernesto Amabilis-Sosa (2024) [1] emplearon la arquitectura VGG16 con técnicas de fine-tuning para el mismo conjunto de datos PlantVillage, obteniendo alrededor del 90 % de sensibilidad y precisión. Este hecho confirma que esta arquitectura también puede ser interesante para el estudio a realizar.

4. Implementación y desarrollo

5. Evaluación y resultados

6. Conclusiones

A. Anexo I: Ejemplo de anexo

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1. Primer elemento.
2. Segundo elemento
3. Tercer elemento.
 - a) Primer subelemento.
 - b) Segundo subelemento.
 - Primer punto.
 - Segundo punto.

B. Anexo II: Otro ejemplo de anexo

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Bibliografía

- [1] Eduardo A.Huerta-Mora, Víctor González-Huitrón, Héctor Rodríguez-Rangel y Leonel Ernesto Amabilis-Sosa. Detección de enfermedades foliares conarquitecturas de redes neuronales convolucionales. <https://rinderesu.com/index.php/rinderesu/article/view/46/50>, 2024. Accessed: 29 July 2025.
- [2] Ing. Agr. Miguel Silva. Cultivo de tomate: Cómo se realiza, plagas e importancia - agrotendencia tv. https://agrotendencia.tv/agricultura/cultivos/hortalizas/el-cultivo-de-tomate/#Historia_del_tomate_o_jitomate, 2025. Accessed: 28 July 2025.
- [3] Ing. Agr. Miguel Silva. Cultivo de tomate: Cómo se realiza, plagas e importancia - agrotendencia tv. https://agrotendencia.tv/agricultura/cultivos/hortalizas/el-cultivo-de-tomate/#Control_de_plagas_y_enfermedades_del_tomate, 2025. Accessed: 28 July 2025.
- [4] Maday Ynfante Martínez, Minelkis Machado Molina, Neili Machado García y Efraín Velasteguí López. Técnicas de aprendizaje supervisado para la detección y clasificación de enfermedades y defectos en imágenes de frutas: revisiones. <https://revistas.utb.edu.ec/index.php/magazine/article/view/2330/1983>, 2024. Accessed: 30 July 2025.
- [5] Valeria Maeda Gutiérrez. Comparación de arquitecturas de redes neuronales convolucionales para la clasificación de enfermedades en tomate. <https://sedici.unlp.edu.ar/handle/10915/139770>, 2019. Accessed: 29 July 2025.
- [6] Wikipedia contributors. Producción mundial del tomate – wikipedia, the free encyclopedia. https://es.wikipedia.org/wiki/Anexo:Producci%C3%B3n_mundial_de_tomate, 2025. Accessed: 28 July 2025.
- [7] Wikipedia contributors. Solanum lycopersicum – wikipedia, the free encyclopedia. https://es.wikipedia.org/wiki/Solanum_lycopersicum, 2025. Accessed: 28 July 2025.