



MODELO PARA DETECCIÓN DE XANTHOMONAS CAMPESTRIS APLICANDO TÉCNICAS DE MACHINE LEARNING MEJORADAS MEDIANTE ALGORITMOS DE OPTIMIZACIÓN

Daniel David Leal Lara

**UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS
FACULTAD DE INGENIERÍA
MAESTRÍA EN CIENCIAS DE LA INFORMACIÓN Y LAS COMUNICACIONES
ÉNFASIS EN INGENIERÍA DE SOFTWARE
BOGOTÁ D.C. – COLOMBIA
2022**

MODELO PARA DETECCIÓN DE XANTHOMONAS CAMPESTRIS APLICANDO TÉCNICAS DE MACHINE LEARNING MEJORADAS MEDIANTE ALGORITMOS DE OPTIMIZACIÓN

Daniel David Leal Lara

Director:

Ing. Julio Barón Velandia Ph.D

Grupo de Investigación:

Interoperabilidad Tecnológica y Semántica

INTECSE

UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS

FACULTAD DE INGENIERÍA

MAESTRÍA EN CIENCIAS DE LA INFORMACIÓN Y LAS COMUNICACIONES

ÉNFASIS EN INGENIERÍA DE SOFTWARE

BOGOTÁ D.C. – COLOMBIA

2022

Dedicatoria

A mis padres porque ellos han sido el pilar fundamental que ha orientado mi vida.

A mi familia que me han guiado y apoyado.

A todos mis amigos y compañeros.

A todas las personas que me han iluminado en el proceso.

AGRADECIMIENTOS

A los profesores de la facultad y especialmente de la maestría que acompañaron el proceso de formación posgradual, y demostrarme que con empeño y dedicación, es posible lograr las metas propuestas.

Particularmente a mi Director de tesis, Ingeniero Julio Barón Velandia Ph.D., por creer en la investigación y darme el acompañamiento y seguimiento en todas las etapas del proceso investigativo.

TABLA DE CONTENIDO

RESUMEN	11
PALABRAS CLAVE	11
INTRODUCCIÓN	12
1. PROBLEMA DE INVESTIGACIÓN.....	13
1.1 Planteamiento del problema	13
1.2 Formulación del problema	14
1.3 Sistematización del problema	14
2. OBJETIVOS.....	15
2.1 Objetivo General	15
2.2 Objetivos Específicos.....	15
3. JUSTIFICACIÓN	16
3.1 Justificación Práctica.....	16
3.2 Justificación Metodológica	16
3.3 Justificación Teórica.....	17
4. HIPÓTESIS.....	18
5. MARCO DE REFERENCIA.....	19
5.1 Marco Teórico	19
5.1.1 Diagnóstico de enfermedades en plantas.....	19
5.1.2 Métodos clásicos de identificación de enfermedades en las plantas	20
5.1.3 Métodos alternativos para identificar enfermedades en plantas	23
5.2 Marco Conceptual	34
5.2.1 Medición del Error.....	34
5.2.2 Optimización mediante algoritmos genéticos.....	36
5.2.3 Optimización mediante algoritmo de punto interior.....	38

5.2.4	Lógica Difusa	39
5.2.5	Modelos de Color	42
5.3	Marco espacial	45
5.3.1	Enfermedad Xanthomonas Campestris.	45
6.	LIMITACIONES	46
7.	DESARROLLO DE LA INVESTIGACIÓN.....	47
7.1	Análisis e investigación.....	47
7.1.1	Alcances y limitaciones del modelo.....	51
7.1.2	Herramientas y aplicaciones en la agricultura	52
7.1.3	Adquisición y pre-procesamiento de datos	53
7.1.4	Selección y procesamiento de características.....	55
7.2	Diseño y Modelado	56
7.2.1	Construcción del modelo	56
7.2.2	Sistema de inferencia difusa ANFIS.....	58
7.2.3	Configuración de los sistemas difusos tipo Sugeno	59
7.2.4	Implementación de agrupamiento difuso	60
7.3	Optimización y Pruebas.	61
7.3.1	Implementación de técnicas de optimización.....	61
7.3.2	Validación.....	63
7.4	Resultados.....	73
7.4.1	Entrenamiento.....	73
7.4.2	Prueba.....	76
7.4.3	Validación del mejor modelo	79
7.4.4	Composición del modelo	80
7.4.5	Desempeño general del modelo.....	86
7.4.6	Prototipo Software	88

8.	CONCLUSIONES.....	96
8.1	Principales Aportes	97
9.	PROSPECTIVA DEL TRABAJO DE GRADO	99
9.1	Lineas de investigación futuras.....	99
9.2	Trabajos de investigación futuros	99
9.3	Productos Académicos	100
9.3.1	Artículos	100
9.3.2	Registro de Software y propiedad intelectual.....	100
9.3.3	Libros.....	100
9.3.4	Ponencias en Eventos Científicos	101
10.	BIBLIOGRAFIA	102

LISTA DE FIGURAS

Figura 5.1. Estructura de un algoritmo de aprendizaje automático	26
Figura 5.2. Estructura de la optimización mediante algoritmo genético	37
Figura 5.3. Funcionamiento del sistema de inferencia Tsukamoto	40
Figura 5.4. Sistema de razonamiento difuso tipo 1.....	41
Figura 5.5. Equivalente ANFIS del sistema de razonamiento difuso tipo 1.	41
Figura 5.6. Plano RGB.....	43
Figura 5.7. Componentes del Modelo HSB.....	44
Figura 7.1. Infección de <i>Xanthomonas Campestris</i> en hojas de cultivo <i>Calendula officinalis</i>	47
Figura 7.2. Reconocimiento de patrones por pixel de acuerdo a una escala de color asociada en una planta enferma.....	48
Figura 7.3. Transformación de color a modelo HSB de hojas con significado enfermo.	50
Figura 7.4. Transformación de color a modelo HSB de hojas con significado sano.....	51
Figura 7.5. Muestra del conjunto de datos de plantas enfermas adquiridas en distintas fuentes para desarrollar el modelo.....	54
Figura 7.6. Muestra del conjunto de datos de plantas sanas adquiridas en distintas fuentes para desarrollar el modelo.	55
Figura 7.7. Comparación de valores reales y simulados de en fase de entrenamiento del primer sistema neurodifuso.....	74
Figura 7.8. Error de valores reales del primer sistema neurodifuso en fase de entrenamiento .	75
Figura 7.9. Comparación de valores reales y simulados de en fase de entrenamiento del sistema neurodifuso.....	75
Figura 7.10. Error de valores reales del primer sistema neurodifuso en fase de entrenamiento	76
Figura 7.11. Comparación de valores reales y simulados de en fase de prueba del primer sistema neurodifuso.....	77
Figura 7.12. Error de valores reales del primer sistema neurodifuso en fase de prueba	78
Figura 7.13. Comparación de valores reales y simulados de en fase de prueba del segundo sistema neurodifuso.....	78
Figura 7.14. Error de valores reales del segundo sistema neurodifuso en fase de prueba.	79
Figura 7.15. Resultados de prueba Kruskal-Wallis para segundo sistema con clustering	80
Figura 7.16. Configuración general primer sistema neurodifuso con clustering.	81
Figura 7.17. Visualización del conjunto de reglas del primer sistema de inferencia neurodifuso con clustering	82

Figura 7.18. Superficie de salida del primer sistema neurodifuso con clustering.	82
Figura 7.19. Estructura ANFIS para primer sistema con clustering.	83
Figura 7.20. Configuración general del segundo sistema neurodifuso con clustering.....	84
Figura 7.21. Visualización del conjunto de reglas del segundo sistema de inferencia neurodifuso con clustering.	85
Figura 7.22. Estructura ANFIS para el segundo sistema con clustering.....	85
Figura 7.23. Matriz de confusión del modelo con mejor desempeño.....	86
Figura 7.24. Aplicación en servidor web.....	88
Figura 7.25. Interfaz de usuario DetXantCamp.....	89
Figura 7.26. Página Principal.....	90
Figura 7.27. Selección de imagen a diagnosticar.....	91
Figura 7.28. Diagnóstico del estado de la planta dado por la aplicación.....	92
Figura 7.29. Documentación del software.....	92
Figura 7.30. Créditos Software DetXantCamp.....	93
Figura 7.31. Interfaz de entorno.....	94
Figura 7.32. Instalación de la aplicación.....	94
Figura 7.33. Instalación y términos y condiciones de la aplicación.....	95
Figura 7.34. Ícono en escritorio de aplicación.....	95

LISTA DE TABLAS

Tabla 5.1. Resumen de trabajos realizados a partir de técnicas de aprendizaje automático para la detección de enfermedades en las plantas a partir de imágenes	27
Tabla 7.1. Relación de HSB Normalizado por píxel de acuerdo a significado enfermo	49
Tabla 7.2. Relación de HSB Normalizado por píxel de acuerdo a significado sano	50
Tabla 7.3. Configuración de parámetros de inferencia difusa Sugeno sin agrupación de clúster.	57
Tabla 7.4. Configuración de parámetros de inferencia difusa Sugeno con agrupación de clúster.	57
Tabla 7.5. Configuración de parámetros de sistema de inferencia difusa ANFIS.....	58
Tabla 7.6. Parámetros de agrupamiento difuso FCMClustering	61
Tabla 7.7. Parámetros para implementación de Algoritmo Genético	62
Tabla 7.8. Parámetros para implementación de Algoritmo de Punto Interior	63
Tabla 7.9. Valores de error del primer sistema neurodifuso con clustering para la detección de Xanthomonas Campestris	64
Tabla 7.10. Valores de error del primer sistema neurodifuso sin clustering para la detección de Xanthomonas Campestris	66
Tabla 7.11. Valores de error del segundo sistema neurodifuso con clustering para la detección de Xanthomonas Campestris	68
Tabla 7.12. Valores de error del segundo sistema neurodifuso sin clustering para la detección de Xanthomonas Campestris	70
Tabla 7.13. Resultados de fase de entrenamiento	73
Tabla 7.14. Resultados fase de prueba	76
Tabla 7.15. Comparación del modelo propuesto respecto a los encontrados en literatura relacionada.....	87

RESUMEN

Esta propuesta se centra en la elaboración de un modelo que permita la detección temprana de la enfermedad *Xanthomonas Campestris* aplicando técnicas de Machine Learning, caracterizadas por su alta interpretabilidad, mejoradas mediante algoritmos de optimización, permitiendo identificar de manera precisa el estado de una planta (Sana o enferma), con el objeto de que los agricultores puedan tomar acciones tempranas reduciendo el impacto que genera la enfermedad en la presentación y rendimiento del cultivo.

PALABRAS CLAVE

Algoritmos de optimización; Análisis de imágenes; Enfermedad *Xanthomonas Campestris*; Optimización; Machine Learning.

INTRODUCCIÓN

La pérdida de cultivos debido a enfermedades es un riesgo importante para la seguridad alimentaria y una amenaza para el sostenimiento de los agricultores que dependen en gran medida del rendimiento saludable de sus cultivos (Yadav, Kumar, and Nagpal 2019). Una de las disciplinas con mayor campo de aplicación para la tecnología es la agricultura, en la cual, se han desarrollado numerosas investigaciones sobre el proceso de siembra, cuidado y recolección de cultivos. Adicionalmente, se ha evidenciado que aún existen amplias posibilidades para realizar aportes significativos a este campo (Zhuang et al. 2015); dentro del cual, se destaca el trabajo sustancial en el campo de la visión por computadora, abordando diversos problemas de detección y clasificación de plantas, utilizando reconocimiento a partir de imágenes de hojas (Elhariri, El-Bendary, and Hassanien 2014).

Tradicionalmente, la identificación de enfermedades en las plantas es realizada a través de un monitoreo constante mediante la observación directa por parte de un experto, ocasionando en su contratación un importante gasto adicional en recursos para la mayoría de los agricultores, y por lo cual, en muchos casos terminan desestimando esta opción (Yadav, Kumar, et al. 2019). Esto ocasiona grandes pérdidas por la falta de control de plagas y enfermedades en los cultivos, ya que, por lo general, los consumidores finales de productos agrícolas descartan aquellos con una presentación física por debajo de los estándares (Francis, Anto Sahaya Dhas D, and Anoop B K 2016). Es por esto que detectar si un cultivo presenta alguna enfermedad en el momento oportuno, puede evitar grandes pérdidas económicas relacionadas con la presentación del producto.

En este trabajo se propone el desarrollo de un modelo basado en técnicas de Machine Learning, mejorado mediante algoritmos de optimización, como solución alternativa para la detección temprana de la enfermedad *Xanthomonas Campestris* con un alto nivel de precisión e interpretabilidad.

Este documento se divide en los siguientes capítulos: problema de investigación, conformado por el planteamiento, formulación y sistematización del problema, objetivo general y objetivos específicos, justificación y finalmente, referencias bibliográficas.

1. PROBLEMA DE INVESTIGACIÓN

1.1 PLANTEAMIENTO DEL PROBLEMA

Los procesos de siembra y cosecha de cultivos demandan una gran cantidad de tiempo en el cuidado y observación por parte del agricultor, siendo la detección temprana de enfermedades una de las tareas críticas en cuanto a tiempo y recursos se refiere. Tradicionalmente para llevar a cabo esta actividad, se requiere la intervención humana constante por parte de un experto a través de la observación directa, siendo esto para el agricultor un proceso difícil, lento y costoso, lo cual implica que muchas veces no se realice. Esto se traduce en general en que un 50% de rendimiento del cultivo se pierda a causa de plagas y enfermedades (Yadav, Kumar, et al. 2019). Para intentar suplir esto, en los últimos años, las investigaciones se han esforzado en automatizar el proceso de detección de enfermedades en lugar de adoptar una técnica convencional (Kaur, Pandey, and Goel 2018), resaltando que los avances en los dominios de la visión artificial y el reconocimiento de patrones han motivado a los investigadores a emplear la tecnología en el dominio de la agricultura para la detección automática de enfermedades foliares.

Teniendo en cuenta que los procedimientos convencionales de observación y cuidado del cultivo no se sustituyen de manera sencilla, es común que se presenten los mismos tipos de problemas de décadas pasadas, como los observados en la detección tardía de la enfermedad *Xanthomonas Campestris* (Conselleria d'Agricultura and Ecològica 2007), los cuales perjudican a los agricultores al ver disminuidas sus respectivas ventas. La detección tardía de enfermedades afecta la apariencia externa de los productos agrícolas, el cual es uno de los atributos de calidad de mayor consideración por parte de los consumidores (Francis et al. 2016), satisfacer estas necesidades a nivel mundial implica aumentar la calidad y rendimiento de los cultivos (S and R 2019).

La intervención humana en actividades de monitoreo constante para la detección visual de enfermedades de los cultivos por parte de personal experto, puede ser mitigada mediante técnicas basadas en el análisis de imagen, las cuales, al obtener información sustancial, permiten una correcta identificación (Ab Jabal et al. 2013). Existen métodos basados en Machine Learning orientados a la clasificación, caracterizados por su alta precisión pero baja interpretación de la dinámica del sistema, como los basados en redes neuronales (Petrellis 2019), y métodos

diferenciados por su alta interpretabilidad pero bajo grado de precisión como los sistemas de lógica difusa (Baquero Hernández, Rodríguez Valdés, and Ciudad Ricardo 2016). Pero no existen modelos que permitan detectar, de manera temprana, la enfermedad *Xanthomonas Campestris* encontrada en las hojas de las plantas, con un alto grado de precisión e interpretabilidad. Por lo cual, se propone el diseño de un modelo basado en técnicas altamente interpretables de Machine Learning, mejoradas mediante algoritmos de optimización con el objeto de permitir la detección en una etapa inicial de la enfermedad.

1.2 FORMULACIÓN DEL PROBLEMA

A partir de los diferentes aspectos establecidos en el planteamiento del problema relacionado con la detección de enfermedades en las plantas, y específicamente asociado a la bacteria *Xanthomonas Campestris*, se plantea la siguiente pregunta de investigación:

¿Cómo detectar de manera temprana la enfermedad *Xanthomonas Campestris* encontrada en las hojas de las plantas, con alto grado de interpretabilidad y precisión?

1.3 SISTEMATIZACIÓN DEL PROBLEMA

Con base en la formulación del problema, se establecen las siguientes subpreguntas orientadas a resolver la pregunta de investigación.

- ¿Cómo caracterizar el comportamiento que tiene la enfermedad *Xanthomonas Campestris* encontrada en las hojas de las plantas?
- ¿Qué técnica basada en Machine Learning permite la detección de la enfermedad *Xanthomonas Campestris* facilitando a los usuarios entender la dinámica y funcionamiento del modelo?
- ¿Cómo aumentar la precisión, sin perder la interpretabilidad, del modelo basado en Machine Learning?
- ¿Cómo evaluar el desempeño del modelo alcanzado con respecto a los métodos similares desarrollados en la literatura?

2. OBJETIVOS

2.1 OBJETIVO GENERAL

Diseñar un modelo que permita la detección de la enfermedad *Xanthomonas Campestris* con alto nivel de precisión e interpretabilidad, aplicando técnicas de Machine Learning mejoradas mediante algoritmos de optimización.

2.2 OBJETIVOS ESPECÍFICOS

- Determinar las variables que permitan la identificación y caracterización de la enfermedad *Xanthomonas Campestris* encontrada en las hojas de las plantas.
- Estructurar el modelo que permita la detección de la enfermedad *Xanthomonas Campestris* con un alto grado de interpretabilidad, permitiendo a los usuarios entender la dinámica del modelo conceptual planteado con base en técnicas de Machine Learning.
- Aplicar técnicas de optimización a la técnica de Machine Learning establecida, con la intención de obtener una mayor precisión sin afectar la interpretabilidad del modelo.
- Evaluar el desempeño del modelo planteado, contrastando los resultados obtenidos, con los reportados en literatura especializada, para determinar el modelo que proporcione el mayor nivel de interpretabilidad y precisión en relación al escenario de aplicación.

3. JUSTIFICACIÓN

3.1 JUSTIFICACIÓN PRÁCTICA

Los métodos convencionales de detección de enfermedades demandan una gran cantidad de recursos representados en tiempo y dinero para los agricultores, dado que usualmente, un agricultor debe viajar grandes distancias para contactar a un experto, el cual, a partir de su observación directa, logra identificar y tratar a tiempo una posible enfermedad, evitando secuelas notorias en el producto final, para conseguir que la presentación de los productos sea acorde a los estándares de calidad de los consumidores (N. H. Reddy, E. R. Kumar, M. V. Reddy 2019). El avance en el campo de técnicas de visión artificial ofrece una oportunidad para realizar una detección temprana de enfermedades con una alta precisión, en comparación con los métodos convencionales (Yadav, Kumar, et al. 2019), esto brinda una alternativa a los agricultores para reducir costos asociados con la contratación del personal especializado, permitiendo disminuir las posibles pérdidas económicas a raíz de la detección tardía de las enfermedades y el tiempo empleado en la actividad, logrando aumentar la rentabilidad de los cultivos.

3.2 JUSTIFICACIÓN METODOLÓGICA

Existen distintas técnicas usadas para obtener modelos que permitan la detección temprana de enfermedades en plantas, dentro de los cuales, se destacan las basadas en redes neuronales por su alta precisión (Petrellis 2019), o las basadas en lógica difusa las cuales se centran en tener una alta interpretabilidad (Baquero Hernández et al. 2016). El uso de un sistema automatizado que extrae las características taxonómicas de las plantas encontradas en una imagen (VijayaLakshmi and Mohan 2016), puede aplicarse a la detección temprana de la enfermedad *Xanthomonas Campestris* y permite reducir la necesidad de monitoreo constante a través de la observación directa por parte de un experto. Constituyendo un aporte desde la visión artificial que mejora los resultados en el campo de la agricultura.

3.3 JUSTIFICACIÓN TEÓRICA

Varios autores han propuesto diversos modelos que permiten realizar una caracterización temprana de enfermedades encontradas en las hojas de las plantas. Sin embargo, se encuentra que la mayoría de los casos, la caracterización realizada en estos modelos resulta ser ineficiente en el momento de ser aplicados al mundo real, ya que en este, factores como la intensidad de luz, sombra, entre otros, varían y dificultan la medición y cuantificación estable y reproducible del color de la hoja (Matsunaga et al. 2017). Con la elaboración de un modelo, que a partir de la fundamentación teórica de variables como intensidad de luz y sombra, actualmente no incluidos en los estudios, es posible obtener de manera temprana la correcta identificación y caracterización de patrones de manchas de la enfermedad *Xanthomonas Campestris* encontrada en las hojas de las plantas, con un alto grado de precisión e interpretabilidad, permitiendo mediante métodos de optimización, mejorar los resultados obtenidos al usar técnicas basadas en Machine Learning. El uso conjunto de estas dos herramientas permitiría en futuras investigaciones, la definición de técnicas metodológicas y mejores prácticas de aplicación.

4. HIPÓTESIS

La obtención de un modelo basado en técnicas de Machine Learning mejoradas mediante algoritmos de optimización, que tiene en cuenta factores medioambientales como la luz, sombra, fondo y reflexión; y que, a su vez, caracteriza mediante un modelo de color las tonalidades específicas asociadas al objeto de estudio, a partir de imágenes tomadas del cultivo afectado, permite detectar con alta precisión e interpretabilidad, la enfermedad *Xanthomonas Campestris* presentada en las hojas de las plantas.

5. MARCO DE REFERENCIA

5.1 MARCO TEÓRICO

En esta sección se realiza una revisión teórica de aspectos que se tuvieron en cuenta en la investigación y el desarrollo del modelo. En la primera subsección se presentan aspectos relacionados con el diagnóstico de enfermedades en plantas y las características propias de la bacteria *Xanthomonas Campestris*, posteriormente se describen los algunos de los métodos clásicos más utilizados para la identificación de enfermedades en las plantas y se finaliza la sección con los métodos alternativos.

5.1.1 Diagnóstico de enfermedades en plantas

En la agricultura el diagnóstico de enfermedades en las plantas se da a partir de la fitopatología, la cual se define como la ciencia que estudia las enfermedades en las plantas y a su vez propone cómo combatirlas (Luis Espinoza Pineda and Urbina Chavarría 2011). El nacimiento de esta disciplina se ubica alrededor del siglo XIX bajo la necesidad de identificar los agentes causales del daño de algunos cultivos y la dificultad que presentaban los agricultores para recuperarlos, ya que, a pesar de tener conocimientos empíricos dados desde la antigüedad, no fue sino hasta el siglo XIX, que a raíz de la gran hambruna irlandesa provocada por el tizón tardío de la papa (*Phytophthora infestans*) entre 1845 y 1849, se empezaron a realizar estudios científicos puntuales acerca de las enfermedades de las plantas, sus causas, su desarrollo y posibles métodos para combatirlas (ARAUZ CAVALLINI n.d.).

La bacteria *Xanthomonas Campestris* es causante de gran variedad de fitopatologías ya que puede llegar a afectar a más de 10 especies, incluso aquellas dispuestas para la alimentación humana, de modo que un diagnóstico oportuno y acertado permite optimizar el desarrollo de la agricultura de precisión y mejorar los cultivos alimenticios en gran cantidad, favoreciendo la productividad del sector agrícola.

De acuerdo con la literatura, los síntomas de esta enfermedad se evidencian principalmente en las hojas y en algunos casos en el tallo de las plantas, de modo que es de especial interés parametrizar los síntomas de esta enfermedad para realizar el diagnóstico. Para esto, en muchos

casos basta con realizar una comparación entre los síntomas descritos para una enfermedad en la literatura destinada a ello y los observados en la planta que se pretende diagnosticar como se observa en numerosos artículos.

Para definir los criterios bajo los cuales se dan los distintos diagnósticos de enfermedades en las plantas, en el texto de (Agrios 1995) se dispone de un apartado donde se describen los diferentes agentes causales de enfermedades en las plantas y a su vez se propone una clasificación para estudiar los mismos la cual se muestra a continuación:

- **Enfermedades infecciosas:**
 - Ocasionadas por plantas parásitas superiores.
 - Ocasionadas por nematodos.
 - Ocasionadas por hongos.
 - Ocasionadas por bacterias.
 - Ocasionadas por virus y viroides.
 - Ocasionadas por micoplasmas.
- **Enfermedades no infecciosas:**
 - Ocasionadas por déficit o exceso de algún factor abiótico del medio.

Por su parte, otros textos como la guía de campo para la identificación y manejo de enfermedades de corpoica, tienen en cuenta las plagas en los cultivos como parte del problema, por lo cual se podría considerar como una clasificación adicional en materia de enfermedades.

Cada uno de estos grupos tiene propiedades comunes, lo cual permite manejar la complejidad inmersa en la identificación de los patógenos. Pese a ello, hay métodos que son específicos para un solo agente patógeno perteneciente a alguno de esos grupos. Tomando como referencia el texto de Introducción a la fitopatología de (Agrios 1995) y el texto sobre técnicas para el diagnóstico de enfermedades en plantas (IICCA 1997) del Instituto Interamericano para la cooperación de la agricultura IICA de México, se hace la siguiente recopilación de técnicas clásicas para la identificación de enfermedades usando la clasificación realizada anteriormente.

5.1.2 Métodos clásicos de identificación de enfermedades en las plantas

Los métodos clásicos para la detección de enfermedades han orientado por décadas a los agricultores, permitiéndoles identificar con cierto grado de éxito, la presencia o ausencia de

enfermedades en las plantas. A nivel de clasificación de enfermedades y su método clásico de identificación, se pueden encontrar las siguientes:

- **Enfermedades infecciosas:** Para Agrios, las enfermedades infecciosas o aquellas que son ocasionadas por patógenos, se caracterizan por la presencia de estos en la superficie de las plantas, allí hospedan y se transfieren con otras generando el contagio. Si se encuentran en la superficie, el diagnóstico puede lograrse por observación directa. Si no se encuentran en la superficie, el diagnóstico se realiza buscando otros síntomas adicionales e intentando localizar al patógeno en alguna parte del tejido infectado para aislarlo.
- **Enfermedades ocasionadas por plantas superiores parásitas:** De los 2 textos base para este apartado (Agrios 1995) (IICCA 1997), este grupo solamente está catalogado en el texto de Agrios y afirma que para realizar esta identificación basta con comprobar que la planta parásita se encuentra creciendo sobre la planta de interés.
- **Enfermedades ocasionadas por nematodos:** Este grupo está catalogado en el texto de Agrios, pero no indica un método de identificación. Por otro lado, están bien documentados en el texto del (IICCA 1997), , donde se indica que de todas las especies de nematodos fitoparásitos solo algunas de ellas deben considerarse como patógenos graves y aun así solo causan grandes pérdidas cuando habitan en condiciones favorables y su población crece de forma desmedida. Se puede realizar identificación por observación en campo, siempre y cuando el nematodo genere agallas o quistes en raíces, hojas o semillas. En otros casos las plantas no presentan síntomas exclusivos en presencia de los nematodos y necesita realizarse un proceso de laboratorio para su identificación. Se debe realizar una toma de muestras, un proceso de extracción y un proceso de identificación.

En primer lugar, las muestras dependen del cultivo y pueden ser aleatorias o aleatorias dirigidas, seguidamente, la extracción depende de la precisión deseada, por ejemplo, uno de los métodos que se utilizan para el diagnóstico de estas enfermedades es el embudo de Baermann, el cual consiste en la coloración diferencial de raíces y centrifugación-flotación en solución azucarada, una vez extraídos los nematodos, se procede a la identificación la cual se realiza comparando con claves o alguna otra herramienta.

- **Enfermedades causadas por hongos:** Para (Agrios 1995), debe tenerse especial cuidado al momento de diagnosticar este tipo de afectación en las plantas, ya que puede haber presencia de hongos saprófitos y hongos patógenos, sólo estos últimos son causantes de enfermedades. La identificación de un determinado tipo de hongo se hace mediante el estudio microscópico de la morfología de su micelio, aunque la mayoría de enfermedades fúngicas que presentan síntomas característicos están bien documentadas en la literatura y se pueden identificar observando los patrones característicos.
- **Enfermedades causadas por bacterias:** En el texto de (Agrios 1995) se enuncia que, por lo general, es posible dar este diagnóstico mediante el análisis de los síntomas característicos de enfermedades bacterianas y la presencia de miles de bacterias en la superficie de los órganos infectados, es decir que sus síntomas se pueden evidenciar a simple vista en la superficie de la planta como en las hojas y el tallo. De igual manera se debe tener cuidado de no confundir una bacteria patógena con una saprófita. El texto del (IICCA 1997) relaciona las bacterias con muchos de los trastornos en las plantas tales como necrosis, manchas, pudriciones, marchitez y agallas, para la identificación precisa de las bacterias por lo general se requiere de aislamiento del agente y retransmisión en plantas. La toma de las muestras y la identificación en laboratorio se debe realizar lo más rápido posible ya que se pueden desarrollar otras bacterias saprófitas en el tejido muerto.
- **Enfermedades causadas por micoplasmas:** Los micoplasmas son similares a las bacterias, pero carecen de pared celular y están presentes en todas las partes de la planta al mismo tiempo. Para (Agrios 1995) el diagnóstico de estas enfermedades también se realiza por sintomatología y asocia sus síntomas característicos a alteraciones morfológicas tales como el enanismo, amarillamiento o enrojecimiento de las hojas, proliferación de raíces y flores anormales. En el texto del (IICCA 1997) presenta la sintomatología como un primer paso en el diagnóstico, pero la identificación precisa requiere de la retransmisión del agente causal a partir de plantas sospechosas mediante injertos.
- **Enfermedades causadas por virus:** Finalmente en la caracterización de agentes causales se describen los virus, donde para (Agrios 1995) se utiliza fundamentalmente la sintomatología para realizar el diagnóstico, pero cuando no es posible realizarlo de esta

forma, se sugiere retransmisión a plantas huésped seleccionadas, serología o pruebas con microscopio electrónico. Por su parte, en el texto del (IICCA 1997) se destaca la dificultad para trabajar con virus y también recomienda técnicas como la retransmisión

5.1.3 Métodos alternativos para identificar enfermedades en plantas

A pesar de la gran variedad de métodos científicos utilizados para detectar enfermedades en las plantas de manera oportuna, no todas resultan ser accesibles para los agricultores, ya sea por su costo, tiempo o conocimiento sobre ellos. Por ello, a lo largo del tiempo y con el desarrollo de la fitopatología, se ha dado la necesidad de conseguir otros métodos para asignar diagnósticos acertados de enfermedades en las plantas y que de la misma forma permitan mejorar la calidad de los cultivos y a su vez la productividad de los mismos, algunos de estos métodos son la detección basada en técnicas moleculares, la microscopía, el bioensayo para la detección de virus y la serología.

Así mismo, más allá de la biología otras áreas del conocimiento se han interesado en desarrollar trabajos que fortalezcan el diagnóstico de enfermedades en las plantas, por ejemplo, en la literatura se datan trabajos recientes como el de espectroscopía para la detección de *Fusarium oxysporum* (Marín-Ortiz, Hoyos-Carvajal, and Botero Fernandez 2018) o el trabajo sobre Dot Blot para la detección temprana de hongos en la soya de (Mattio et al. 2018a) , los cuales se encuentran descritos a continuación:

Espectroscopía: En el trabajo titulado *Detección de plantas asintomáticas de Solanum lycopersicum L. infectadas con Fusarium oxysporum usando espectroscopía de reflectancia* de los autores Marín, Hoyos y Botero de la Universidad Nacional de Colombia sede Medellín (Marín-Ortiz et al. 2018), se propone un método basado en la tecnología de espectroscopía de reflectancia, la cual se basa en la emisión de un haz de luz sobre la muestra a analizar, la cual, en función de su composición, o mejor aún, de la naturaleza de los enlaces presentes en sus moléculas, interaccionará con ellos absorbiendo una determinada cantidad de radiación. Es una técnica no destructiva y de rápida implementación. En este caso se realizó el análisis sobre las hojas para detectar anomalías en la planta, demostrando que este método funciona para detectar enfermedades en plantas asintomáticas de manera acertada.

Dot Blot: Es una prueba de diagnóstico económica y de fácil ejecución en laboratorios de baja complejidad, que se encarga de generar un precipitado coloreado para la detección de bacterias,

moléculas o proteínas puntuales sin la necesidad de disponer de equipos específicos (Mattio et al. 2018a). Un ejemplo de este método alternativo para la detección de enfermedades, es el trabajo titulado *Validación de una técnica de Dot blot para la detección de Cercospora kikuchi en plantas de soya* (Mattio et al. 2018b), en el que se utiliza esta técnica como una herramienta para detectar la enfermedad en la soya precozmente ofreciendo resultados positivos y negativos que permiten confirmar o descartar la presencia de *Cercospora kikuchi* con un elevado nivel de confianza en etapas tempranas de infección.

Sin embargo, ambos trabajos concuerdan que los métodos mencionados en sus investigaciones pueden tomar mucho tiempo, llegar a ser costosos o requieren inicialmente de personal altamente experimentado para la realización de tales procedimientos, de modo que, aunque favorecen la detección de enfermedades en plantas, no son una vía más sencilla para algunos agricultores. En respuesta a una optimización en tiempo y costo de los procedimientos de identificación, de la mano de la tecnología han surgido otras propuestas alternativas para la detección de enfermedades. Algunas de ellas son variaciones de las técnicas enunciadas anteriormente y otras se basan en tecnologías que implementan inteligencia artificial (Sistemas expertos y detección mediante imágenes), como los siguientes:

Sistemas expertos: Se trata una propuesta tecnológica que pretende simular el diagnóstico que realizaría una persona experta basada en la observación de los síntomas en la planta que se pretende diagnosticar. Hay diversos trabajos de este tipo, uno de ellos es el trabajo titulado *Sistema experto para el diagnóstico presuntivo de enfermedades fúngicas en los cultivos* (Quintero-Domínguez, Fajardo, and Spíritus 2019) , desarrollado en la Universidad de ciencias informáticas de la Habana Cuba, donde se realizó un sistema experto usando SWI prolog para la base de conocimientos y Java para la interfaz gráfica. Este sistema realiza una serie de preguntas y basado en datos descriptivos sobre las partes de la planta y sobre las manifestaciones de la enfermedad, determina uno o varios hongos que posiblemente afectan a uno de 8 cultivos. Tiene todas las ventajas y desventajas de cualquier sistema experto, adicionalmente tiene la desventaja de poder diagnosticar únicamente cuando hay manifestaciones de la enfermedad, lo cual deja fuera de este método las plantas asintomáticas. No obstante, los métodos anteriores requieren de horas para realizarse, mientras que este requiere de tan solo algunos minutos.

Identificación de imágenes: El método de detección de enfermedades en las plantas a partir de imágenes, se remonta en la literatura desde los años ochenta aproximadamente (Kerkech, Hafiane, and Canals 2020b), donde en búsqueda de evitar pérdidas en la industria agrícola se han propuesto muchos métodos rápidos y eficaces para la detección de enfermedades (Deng et al. 2016), los cuales han llevado a varios investigadores a desarrollar aplicaciones activas de detección remota y diagnóstico técnico, basadas en el procesamiento de imágenes para crear algoritmos de detección mediante aprendizaje automático (Shin et al. 2020).

Actualmente, existen diferentes formas de adquirir imágenes que faciliten el reconocimiento de regiones enfermas en las plantas, las cuales, generalmente están disponibles para que, junto con el avance de la visión por computadora, se logre descifrar síntomas específicos de enfermedades que afectan a las mismas (Trivedi, Shamnani, and Gajjar 2020), un ejemplo de ello, son las imágenes capturadas bajo condiciones controladas, las cuales han permitido detectar lesiones a partir de las características de la imagen como lo son la textura y el color.

De igual manera, a medida que pasa el tiempo ha crecido el interés por investigar la aplicación de algoritmos de aprendizaje automático para la detección de enfermedades en las plantas, siguiendo una estructura similar al de la figura 5.1, se han empleado y discutido varias técnicas de estas, mientras que parece que no existe una técnica universal para todo tipo de circunstancias (Nturambirwe and Opara 2020). Es por ello que dentro de la bibliografía extensa que recopila las distintas formas de reconocer enfermedades a partir de imágenes, se destaca que las imágenes adquiridas son la mayoría de cultivos que tienen mayor sensibilidad a contraer enfermedades que afectan principalmente sus hojas.

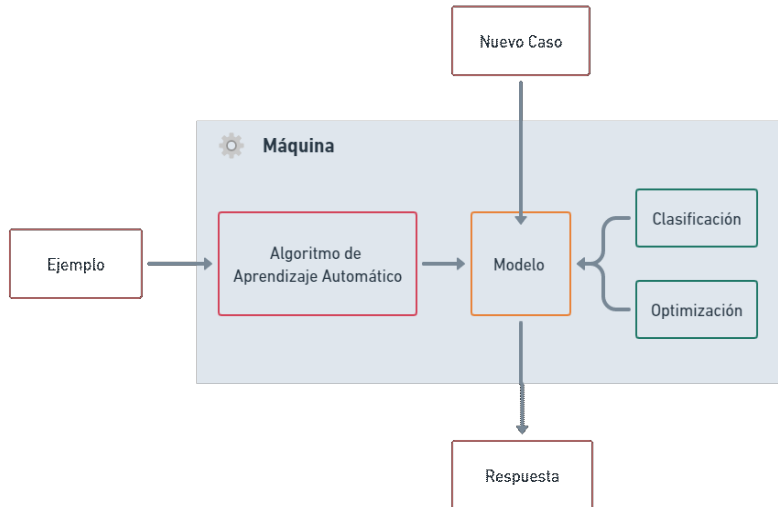


Figura 5.1. Estructura de un algoritmo de aprendizaje automático
Fuente: Elaboración Propia

De algunos trabajos ya mencionados y otros que han sido investigados, se ha desarrollado la tabla 5.1, donde se resume la forma en la que los investigadores han logrado detectar enfermedades de las plantas a partir de imágenes con ayuda de un modelo propuesto, en esta, se tiene en cuenta la técnica de aprendizaje automático utilizada, el conjunto de datos (el cultivo en estudio, las enfermedades en específico, la fuente de adquisición de las imágenes y la cantidad necesaria para desarrollar el modelo), la arquitectura utilizada, la técnica de optimización en caso de ser necesaria y finalmente el rendimiento obtenido por cada uno.

Tabla 5.1. Resumen de trabajos realizados a partir de técnicas de aprendizaje automático para la detección de enfermedades en las plantas a partir de imágenes

Fuente: Elaboración Propia

Referencia	Técnica de Aprendizaje Automático	Conjunto de Datos				Arquitectura	Técnica de Optimización	Rendimiento
		Tipo de Cultivo	Enfermedad	Fuente de adquisición de imagen	Cantidad de imágenes			
(Yadav, Kumar Rana, and Nagpal 2019)	Red Neuronal Convolutacional	7 especies	26 enfermedades	Plant Village	8750	AlexNet	Particle Swarm Optimization (PSO)	97,39
(Mustafa et al. 2020)	Red Neuronal Convolutacional	Hierbas	No especifica	Manual	300	Algoritmo inteligente híbrido	No aplica	98,4
(Sharma, Berwal, and Ghai 2020)	Red Neuronal Convolutacional	Tomate	9 enfermedades	Plant Village	17929	Generalized Softmax Perceptron (GSP)	Optimizador de gradiente	98,6
(Yamamoto, Guo, and Ninomiya 2016)	KNN y Bosque Aleatorio	Tomate	No especifica	Scale-invariant feature transform y Bolsa de palabras visuales (BoVWs)	30000 pixeles	No especifica	No aplica	91

(Zhu et al. 2017)	Máquina de vectores de soporte (SVM) y Árbol de decisión	Tabaco (Nicotiana tabacum L)	Tobacco mosaic virus	No especifica	180	No especifica	Algoritmo de proyecciones sucesivas (SPA)	98,33
(Cao et al. 2018)	Máquina de Vectores de Soporte (SVM), KNN y Naïve Bayes	Brassica napus o Colza	Podredumbre Blanca	Scale-invariant feature transform	240	No especifica	No aplica	93,55
(Hernández and López 2020)	Red Neuronal Convolutacional	Banano	Enfermedad del racimo del banano (BBTD), Xanthomonas Wilt de banano (BXW)	RetinaNet	7013	Grupo de Geometría Visual (VGG-16) y un modelo personalizado	Adam Optimization Algorithm y descenso de gradiente estocástico	92
(Tan et al. 2018)	Máquina de Vectores de Soporte (SVM) y KNN	Cacao	Mazorca Negra (Phytophthora spp)	Cámara de teléfono móvil	60	No especifica	Algoritmo de cortes de gráficos	86,96
(Alsuwaidi, Grieve, and Yin 2018)	Máquina de Vectores de Soporte (SVM) y KNN	Hoja de Azúcar	Cercospora, roya y mildiú polvoriento	Espectroradiómetro de mano	2516	No especifica	Algoritmo Relief-F	91,58
(Arribas et al. 2011)	Red Neuronal	Helianthus annuus	Maleza	Cámara digital	1434	Generalized Softmax Perceptron (GSP)	Algoritmo genético	90
(Santoso et al. 2019)	Máquina de Vectores de Soporte (SVM) y Árbol de decisión	Helianthus annuus	Enfermedad de la pudrición del tallo basal (BSR)	WorldView-3 (Imágenes satelitales)	1923	No especifica	Stepclass algoritmo	54,1

(Shin et al. 2020)	Máquina de Vectores de Soporte (SVM) y Red Neuronal Artificial	Fresa	Sphaerotheca macularis	Cámara digital	1016	No especifica	Algoritmos HOG, SURF y GLCM	94,34
(Gu et al. 2019)	Máquina de Vectores de Soporte (SVM) y Árbol de decisión	Tabaco	Virus del marchitamiento manchado del tomate	Cámara espectral e HyperScanner	320	No aplica	Algoritmo de proyecciones sucesivas (SPA) y algoritmo genético (GA)	85,2
(Hernández and López 2020)	Naïve Bayes y Red Neuronal Convolutacional	14 especies	26 enfermedades	Base de datos PlantVillage e Imagenet	54306	Grupo de Geometría Visual (VGG-16), ResNet y DenseNet	Descenso de gradiente estocástico	96
(Ali et al. 2017)	KNN y Máquina de vectores de soporte (SVM)	Mandarina Kinnow	Enfermedades vellosas y Antracnosis	Cámara digital	199	No especifica	Algoritmo Delta E y Bagged Tree Classifier	99,7
(Wspanialy and Moussa 2020)	Red Neuronal Convolutacional	Tomate de invernadero	9 enfermedades	Base de datos PlantVillage	18005	Arquitectura de aprendizaje profundo ResNet y U-Net	Descenso de gradiente estocástico	94 - 97
(Poblete et al. 2020)	Red Neuronal Artificial, KNN, Árbol de decisión y Máquina de vectores de soporte (SVM)	Olivo	Xylella fastidiosa	Un generador de imágenes hiperespectrales y una cámara térmica	1543	No especifica	Algoritmo Adaboost y RUSBoost	80

(Ozguven and Adem 2019)	Red Neuronal Convolutacional	Remolacha azucarera	Cercospora beticola Sacc	No especifica	155	Faster R-CNN	Descenso de gradiente estocástico	95,48
(Omrani et al. 2014)	Máquina de vectores de soporte (SVM)	Manzano	Alternaria, la mancha negra de la manzana y la plaga del minador	Cámara digital y de computador	320	No especifica	Sistema adaptativo neuro-difuso y Levenberg-Marquardt	89 - 98
(Xia et al. 2019)	Máquina de Vectores de Soporte (SVM) y KNN	Brassica napus o Colza	Estrés por anegamiento	Cámara Hiperespectral	180	No especifica	Algoritmo de proyecciones sucesivas (SPA)	95,9
(Kerkech et al. 2020b)	Red Neuronal Convolutacional	Vid o Parra	Mildiú en la vid	El dron quadcopter UAV y el sensor Survey-2 de alta resolución	17640	SegNet	Algoritmo iterativo basado en un detector de puntos de interés	92,81
(Sharif et al. 2018)	Máquina de vectores de soporte (SVM)	Plantas de cítricos	Antracnosis, mancha negra, cancro, sarna, enverdecimiento y melanosa	Plant Village y Citrus Images Database of Infested with Scale	5632	No especifica	High Dimensional Color Transform y Análisis de componentes principales	95,8
(Van De Vijver et al. 2020)	Máquina de vectores de soporte (SVM)	Solanum tuberosum L.	Alternaria solani	Sensor de imágenes	1000	No especifica	Optimizador de gradiente	92
(Deng et al. 2016)	Máquina de vectores de soporte (SVM)	Plantas de cítricos	Citrus Huanglongbing	Cámara digital	896	No especifica	Particle Swarm Optimization (PSO)	88,2 - 91,92

(Behmann, Steinrücken, and Plümer 2014)	Máquina de vectores de soporte (SVM)	Hordeum vulgare	Estrés por sequía	Sensor generador de imágenes hiperespectrales	312	No especifica	Simple Linear Iterative Clustering	67,9
(Zhang and Meng 2011)	Red Neuronal, Máquina de vectores de soporte y KNN	Plantas de cítricos	Xanthomonas citri	Cámara digital	9000	Algoritmo SceBoost	Algoritmo de unión de ventanas	88
(Cruz et al. 2019)	Red Neuronal Convolutacional	Vid o Parra	Grapevine yellows	Imágenes de laboratorio y base de datos PlantVillage	2680	AlexNet, Google-Net, Inception V3, SqueezeNet, ResNet-50 y ResNet-101	Descenso de gradiente estocástico	98,51
(Kerkech, Hafiane, and Canals 2020a)	Red Neuronal Convolutacional	Vid o Parra	Mildiú en la vid	Software Agisoft Metashape	35820	Red de detección de enfermedades de la vid (VddNet)	No aplica	93,72
(Kim, Lee, and Kim 2020)	Red Neuronal Convolutacional	Allium cepa	Mildiú veloso	Sistema de monitoreo de campo con escáner de alta resolución	12813	Grupo de Geometría Visual (VGG-16)	Optimización a partir de Umbral	87,2
(Sladojevic et al. 2016)	Red Neuronal Convolutacional	6 especies	13 enfermedades	ImageNet	33469	CaffeNet	Particle Swarm Optimization (PSO)	96,3
(Chouhan et al. 2021)	Red Neuronal Convolutacional y KNN	Jatropha Curcas L. y Pongamia Pinnata L.	No especifica	Cámara digital	288	ADALINE (ADAPtative LINear Element)	Optimizador de gradiente	95 -97

(Yang et al. 2020)	Red Neuronal Convolutacional	14 especies	21 enfermedades	Base de datos PlantVillage e ImageNet	58200	NASNetLarge	Descenso de gradiente aleatorio	93,05
(Xiong et al. 2020)	Red Neuronal Convolutacional	6 especies	27 enfermedades	Base de datos PlantVillage	38072	MobileNet	No aplica	84,83
(Sambasiva m and Opiyo 2021)	Red Neuronal Convolutacional y KNN	Manihot esculenta	5 enfermedades	Kaggle	10000	Personalizada	Learning Rate Finder	92 - 96
(Singh and Misra 2017)	Máquina de Vectores de Soporte (SVM) y KNN	5 especies	Enfermedades bacterianas, quemaduras solares y tempranas y enfermedad fúngica	Cámara digital	106	No especifica	Algoritmo genético	97,6
(Kasinathan , Singaraju, and Uyyala 2021)	Red Neuronal Artificial, KNN, Naïve Bayes y Máquina de vectores de soporte (SVM)	Beta Vulgaris, Soja y Arroz	Plagas provocadas por insectos	Conjunto de datos de Wang y Xie dentro de cultivos de campo	1784	Modelo CNN	Descenso de gradiente estocástico	93,9
(Khan and Narvekar 2020)	Máquina de Vectores de Soporte (SVM), Árbol de decisión, KNN y Naïve Bayes	Solanum lycopersicum	Tizón temprano, el tizón tardío y la mancha foliar por Septoria	Base de datos PlantVillage y otras plataformas web	733	No especifica	Optimización basada en superpíxeles en imágenes naturales	93,12

Como se puede evidenciar en la tabla 5.1, es posible capturar imágenes a partir de vehículos aéreos no tripulados, cámaras digitales, sitios web o bases de datos como PlantVillage, imágenes de laboratorio en ambientes controlados o aleatorios, imágenes hiperespectrales, e incluso de forma más sencilla con cámaras fotográficas simples como la de los teléfonos móviles, lo que permite la facilidad de generar gran variedad de estudios a partir de imágenes para la agricultura de precisión.

Sin embargo, varios autores han mostrado que los modelos basados en técnicas de aprendizaje automático reflejan óptimos resultados, entre los cuales se destaca el uso de Máquina de Vectores de Soporte (SVM), Vecino más cercano (KNN), Redes Neuronales Convolucionales y Artificiales, Bosque Aleatorio, Naïve Bayes y Árbol de decisión, que en algunos casos, con ayuda de arquitecturas como algoritmos inteligentes tipo híbrido, ResNet, U-Net, Grupo de geometría visual (VGG-16), Google-Net, Inception V3, SqueezeNet y Generalized Softmax Perceptron (GSP) e incluso técnicas de optimización como Descenso de gradiente estocástico, Optimización por enjambre de partículas, Optimizador de descenso de gradiente y Algoritmos de selección de funciones, han alcanzado rendimientos entre el 54,1% y 99,7%.

Lo anterior, demuestra que este método alternativo para identificar enfermedades en las plantas tiene grandes capacidades, ha resultado ser una buena respuesta para los agricultores ya que con el transcurso del tiempo se ha convertido en una de las mejores alternativas para el control y seguimiento de los cultivos de gran y poca extensión, asignando un diagnóstico rápido y preciso al momento de detectar enfermedades determinadas en las plantas, a su vez, se ha demostrado que es posible tener un diagnóstico oportuno a partir de la implementación de un sistema de aprendizaje automático que sea capaz de detectar los primeros síntomas de una enfermedad de forma más económica y sin la necesidad de tener grandes conocimientos.

5.2 MARCO CONCEPTUAL

En la presente sección se describen conceptos y variables que al ser tenidas en cuenta en la ejecución del presente proyecto permitieron el óptimo desarrollo del modelo propuesto. En primer lugar, se detallan las medidas estadísticas utilizadas para la valoración de éxito del modelo, seguidamente se abordan las técnicas de optimización que permitieron mejorar los resultados, para luego abordar los sistemas de lógica difusa y culminar con los diferentes modelos de color.

5.2.1 Medición del Error

La valoración de éxito se realiza sobre una predicción numérica, teniendo en cuenta que en las fórmulas de cada medida, las instancias de prueba son denominadas $p_i, 0 \leq i \leq n$, mientras que los valores actuales son $a_i, 0 \leq i \leq n$.

Algunas medidas para la medición del error son las siguientes (Witten et al. n.d.):

- Error Cuadrático Medio (MSE): Es la medida más usada ya que compara las diferencias entre los valores esperados del modelo o estimador puntual y los valores verdaderos generados por el mismo como se muestra en la ecuación 1, en ocasiones también es denominada raíz del error cuadrático medio (RMSE) cuando se usa la raíz cuadrada para expresar el error en las mismas dimensiones que los valores obtenidos como se muestra en la ecuación 2. Diversas técnicas la utilizan porque es matemáticamente sencillo de manipular. No obstante, todas las medidas de rendimiento son fáciles de calcular por lo que no ofrece una ventaja particular.

$$MSE = \frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{n} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (p_i - a_i)^2}{n}} \quad (2)$$

- Error Absoluto Medio (MAE): Es una medida la cual realiza la media de la magnitud de los diversos errores individuales obviando el signo como se muestra en la ecuación 3. Este error tiende a resaltar los efectos causados por los valores atípicos en el resultado final.

$$MAE = \frac{|p_1 - a_1|^2 + \dots + |p_n - a_n|^2}{n} \quad (3)$$

- Error Cuadrático Relativo (RSE): Este error al ser relativo es el indicador de la calidad de las medidas tomadas durante la prueba, es decir, el error es relativo a lo que hubiese podido suceder si un predictor se hubiese aplicado. Este predictor promedia todos los valores de los datos actuales de entrenamiento denotados por \underline{a} como se muestra en la ecuación 4. Consecuentemente el error cuadrático toma el error cuadrático total y lo normaliza dividiéndolo entre el error cuadrático total del predictor.

$$RSE = \frac{(p_1 - \underline{p})^2 + \dots + (p_n - \underline{p})^2}{(a_1 - \underline{a})^2 + \dots + (a_n - \underline{a})^2} \quad (4)$$

- Error Relativo Absoluto (RAE): Esta medida se encarga del rendimiento especialmente en los modelos de aprendizaje automático, la descripción es similar al error cuadrático relativo con la diferencia de que este es el error absoluto total dividido entre el error absoluto total del predictor que predice valores promedio como se muestra en la ecuación 5.

$$RAE = \frac{|p_1 - \underline{p}| + \dots + |p_n - \underline{p}|}{|a_1 - \underline{a}| + \dots + |a_n - \underline{a}|} \quad (5)$$

- Coeficiente de Correlación (ρ): En el aprendizaje automático esta medida es utilizada para visualizar y evaluar los modelos según las diferencias entre los valores esperados y reales durante la prueba (Nagano et al. 2019). Mide la correlación estadística entre el conjunto de datos entrenamiento y los datos obtenidos por la

predicción como se muestra en la ecuación 6, si este valor llega a valer 0, no existe correlación entre los datos y si vale 1, existe correlación perfecta entre los datos.

$$\rho = \frac{\frac{\sum(p_i - \bar{p})(a_i - \bar{a})}{n-1}}{\sqrt{\frac{\sum(p_i - \bar{p})^2}{n-1} \frac{\sum(a_i - \bar{a})^2}{n-1}}} \quad (6)$$

5.2.2 Optimización mediante algoritmos genéticos

Los algoritmos genéticos son algoritmos basados en los principios de la genética y de la selección natural. Los elementos de la genética natura –Reproducción, cruzamiento y mutación – son utilizados para el procedimiento de búsqueda genética (Rao 2009).

Estos algoritmos han demostrado un buen comportamiento y ofrecen un enfoque atractivo para la optimización de múltiples criterios que no puede ser manejado por la mayoría de los otros métodos (Verma and Zhang 2007) y difieren de otros métodos de optimización en los siguientes aspectos (Rao 2009):

1. Una población de puntos es utilizada por empezar el procedimiento de un único punto de diseño. Si el número de variables de diseño es n , usualmente el tamaño de la población se toma como $2n$ hasta $4n$. Debido a que muchos puntos son utilizados como soluciones candidatas, estos algoritmos no suelen caer atrapados en un mínimo local.
2. Usa solo los valores de la función objetivo. Las derivadas no se utilizan.
3. Las variables de diseño son representadas como cadenas de variables binarias que corresponden a cromosomas de la genética. Por ello, el método de búsqueda es naturalmente aplicable para resolver problemas de programación entera y discreta; para el caso continuo, la cadena puede variar para llegar a la solución deseada.
4. La función objetivo corresponde a un vector de diseño que toma el rol de “aptitud” (“*fitness*” en inglés) en la genética.
5. En cada generación, un nuevo conjunto de cadenas se produce utilizando selección aleatoria de los padres y cruzamiento de las generaciones anteriores (conjuntos de cadenas anteriores). No son técnicas de búsqueda aleatoria, sino que exploran

eficientemente las nuevas combinaciones con el conocimiento disponible para hallar una nueva generación con mejor “fitness” o valor de la función objetivo.

En la figura 5.2, se evidencia el proceso de optimización utilizando algoritmos genéticos permitiendo una mejor selección de características.

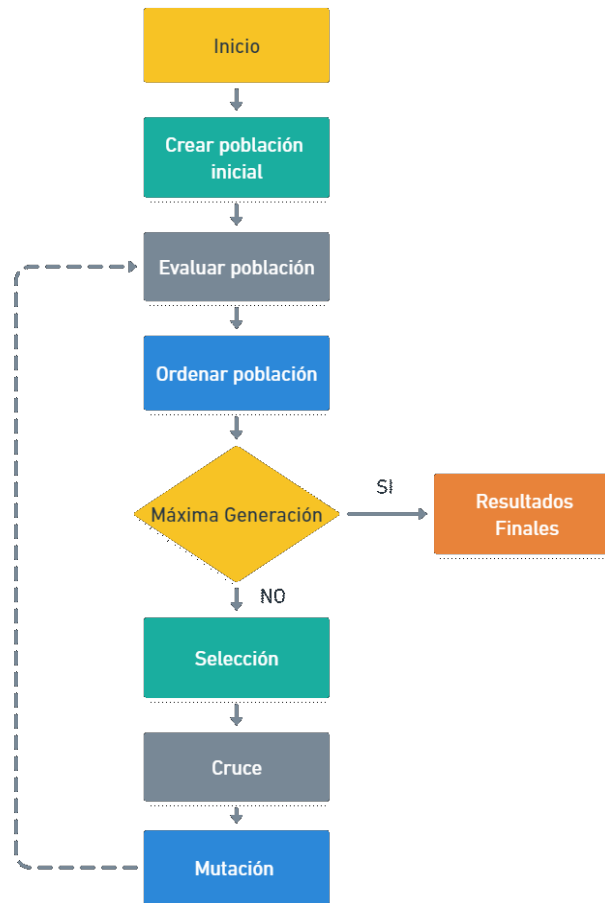


Figura 5.2. Estructura de la optimización mediante algoritmo genético
Fuente: Elaboración Propia

A continuación se presentan algunas implementaciones de algoritmos genéticos (Guo, Wang, and Han 2010):

- Algoritmo Genético Híbrido (HGA): Este algoritmo surge de la necesidad de encontrar los valores iniciales adecuados para el algoritmo genético, ya que por lo general estos son seleccionados por la intuición y diferentes valores iniciales derivan en diferentes resultados óptimos. La debilidad del algoritmo genético tradicional

reside en su dificultad cuando se establecen restricciones al problema, en especial a los problemas con restricciones exactas. Para ello, el HGA consiste en aplicar un algoritmo genético para proveer un conjunto inicial de variables de diseño y así evitar el proceso de prueba; posteriormente, se aplica un algoritmo de optimización tradicional para hallar el resultado óptimo.

- Algoritmo Genético de Intervalo (Interval Genetic Algorithm - IGA): En ingeniería no es fácil de manufacturar variables de diseño exactas por los errores de medición y también porque es más costoso. En el proceso tradicional de optimización de intervalos, el análisis es presentado como una parte del proceso y su propósito es establecer las cotas superiores e inferiores en el efecto de todos los errores en una cantidad computada. Este análisis sirve para comprender la relación entre los parámetros del sistema y su rendimiento, no obstante, la formulación del intervalo no es fácil de construir. En IGA, los parámetros optimizados se pueden obtener y, además, el análisis de intervalo se puede excluir del proceso de optimización.
- Algoritmo Genético Híbrido de Intervalo (Hybrid Interval Genetic Algorithm - HIGA): En el algoritmo anterior, el error objetivo es calculado de la ecuación formulada. En los sistemas complicados, formular un sistema de ecuaciones que le describa resulta complejo y aun si se puede formular el sistema, es difícil determinar soluciones precisas. Para resolver esto, se combina IGA con software de método de elementos finitos para la optimización de intervalos.

5.2.3 Optimización mediante algoritmo de punto interior

Los algoritmos de punto interior son algoritmos iterativos de optimización aplicado a problemas de muchas variables cuyo principio básico es la transformación del problema original en una secuencia de sistemas no lineales que se aproximan de forma sucesiva a las condiciones de Karush-Kuhn-Tucker, este algoritmo asegura que las restricciones se encuentren en el interior de las restricciones de desigualdad siguiendo una dirección entre centralidad y gradiente (Garcés Ruiz and Gómez Carmona 2008). En general, este algoritmo se basa en comenzar el proceso de minimización con un vector de variables que satisface las restricciones del problema, moviéndose por direcciones de descenso que permanezcan siempre en la región admisible (Queizan 2006).

5.2.4 Lógica Difusa

El lenguaje natural es posiblemente la más poderosa forma de comunicación que poseen los seres humanos para cualquier problema o situación que requiera solución o razonamiento. Este poder no se ha explotado en los paradigmas matemáticos modernos, no tanto como en la lógica difusa (Ross 2010).

Un sistema difuso consiste en un sistema en el cual los rangos de las variables sobre los cuales representan sus estados (o, al menos, algunas de ellas) se tratan de conjuntos difusos. Para cada variable, los conjuntos difusos son definidos en algún conjunto difuso que a menudo es un intervalo de números reales y las variables asociadas son variables lingüísticas (Klir and Yuan 1995).

A continuación, se presentan los sistemas de inferencia difusos de tipo Tsukamoto y neurodifuso (ANFIS).

- *Sistema de Inferencia Difuso Tsukamoto*

Este sistema de inferencia, a diferencia de otros, cada consecuente de las reglas condiciones (IF-THEN) debe ser representada por un conjunto difuso con una función de pertenencia monótona. Por lo cual las consecuencias de cada regla son concretas (crisp) presentado con un predicado α . El resultado final es obtenido como un promedio ponderado (Ariani and Endra 2013). El valor α es el valor mínimo del grado de pertenencia (Robert Czabanski 2017),(Perangin-angin et al. 2016).

Los pasos que se siguen para el funcionamiento del sistema de inferencia Tsukamoto se muestran en la figura 5.3:



*Figura 5.3. Funcionamiento del sistema de inferencia Tsukamoto
Fuente: Elaboración propia basada de (Adriyendi 2018)*

Las ventajas que presenta este modelo de inferencia están especialmente en las situaciones de incertidumbre y vagancia, donde este modelo es muy flexible y posee tolerancia a los datos existentes (Adriyendi 2018).

- *Sistema de Inferencia Neurodifuso Adaptativo (ANFIS)*

Este sistema utiliza un procedimiento de aprendizaje híbrido basado en el marco de trabajo de las redes neuronales adaptativas, el ANFIS construye un mapa de entrada-salida basado en conocimiento humano y de los pares de entrada-salida estipulados (Jang 1993). Es utilizado en la predicción de series de tiempo caóticas así como de sistemas de control automático y procesamiento de señales (Jang 1993).

En la figura 5.4 se observa un sistema de razonamiento difuso de tipo 1.

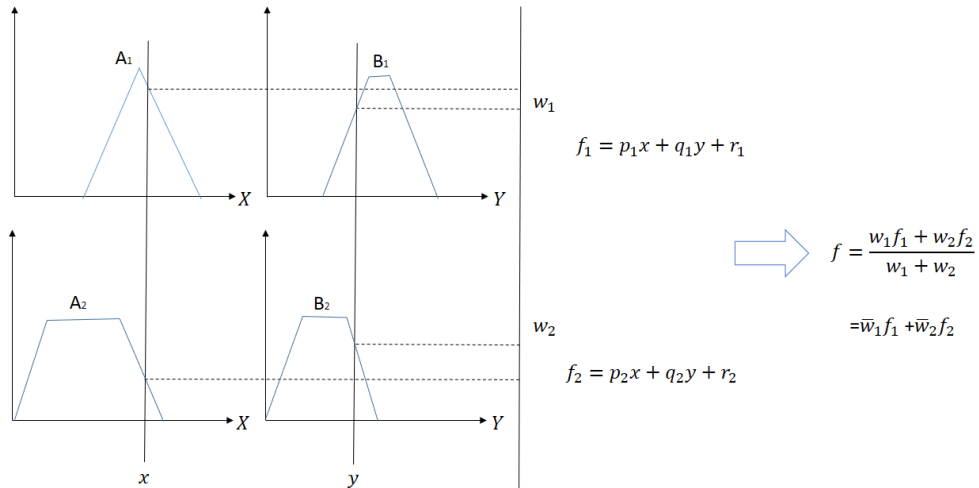


Figura 5.4. Sistema de razonamiento difuso tipo 1
 Fuente: Elaboración y traducción propia basado de (Jang 1993)

La anterior arquitectura de un sistema de razonamiento difuso tipo 1 se traduce en una arquitectura tipo ANFIS en lo observado en la figura 5.5.

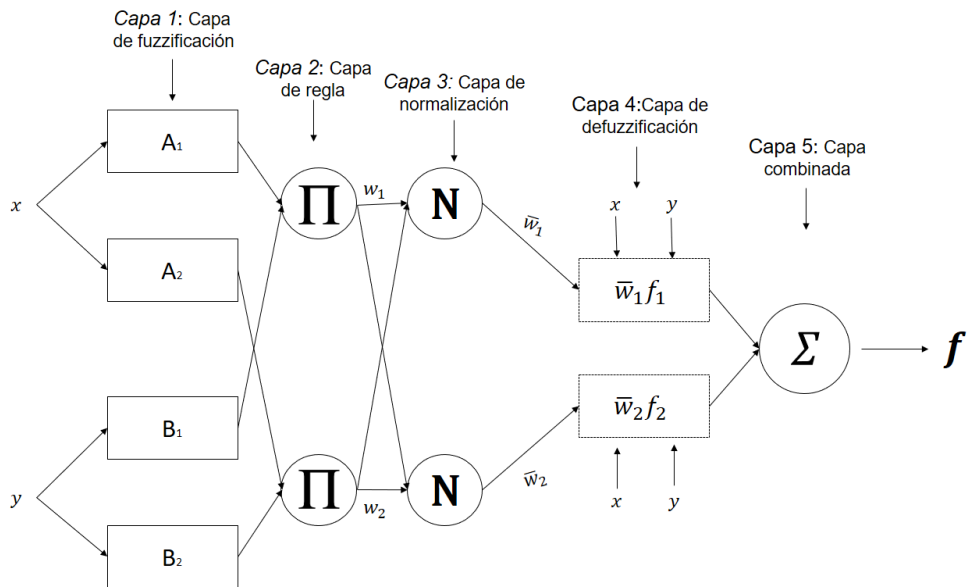


Figura 5.5. Equivalente ANFIS del sistema de razonamiento difuso tipo 1.
 Fuente: Elaboración y traducción propia basado de (Jang 1993)

Los parámetros que conforman el sistema de inferencia difuso se pueden optimizar de manera automática utilizando un algoritmo (Knapp 2007b). El Sistema de Inferencia Neurodifuso adaptativo es un algoritmo que toma los parámetros del sistema de inferencia difuso y los adapta utilizando redes neuronales (Knapp 2007a).

Los siguientes son los pasos para construir un sistema ANFIS basados en lo afirmado por (Knapp 2007a):

1. Diseñar un sistema de inferencia de tipo Sugeno con el fin de realizar la clasificación.
2. Optimizar de manera manual el sistema de inferencia difuso con base a los datos de clasificación de entrada.
3. Preparar los datos de entrenamiento y de prueba.
4. Ejecutar el algoritmo ANFIS sobre los datos de entrenamiento.
5. Probar los resultados utilizando datos de prueba.

Dentro de las ventajas que presenta este sistema son las siguientes (Isanta Navarro 2013)

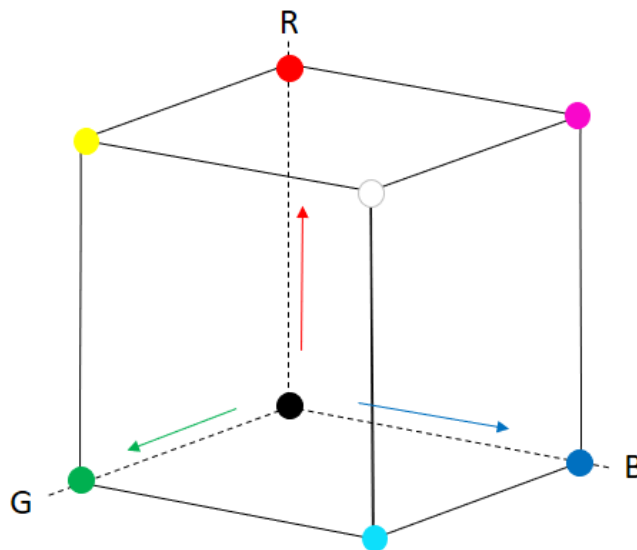
- Su estructura permite el uso de la computación en paralelo.
- La representación del conocimiento es bien estructurada.
- Integración con otros métodos de diseño de control.

5.2.5 Modelos de Color

Los modelos de color describen los colores que se ven en las imágenes digitales y análogas, permitiendo así no solo establecer un espacio único común a todos los equipos que forman parte de la cadena de adquisición de imágenes y reproducción de color, sino que también permiten simular cómo lucirá la imagen y su color en otro dispositivo (Calvo 2016). En el terreno del análisis visual, se han establecido algunos modelos para identificar la composición de color que ofrece una mejor discriminación de cubiertas para un fin determinado, como lo puede ser la detección de enfermedades a partir de los síntomas visibles de la misma, pese al interés de estas composiciones, es importante las combinaciones de color disponibles en los sensores de interés medioambiental, además de garantizar una buena discriminación de las cubiertas vegetales independientemente de las bandas del espectro que se incluyan en la composición (Chuvieco 1990).

- *Modelo RGB*

RGB es un modelo de color aditivo utilizado en los sistemas electrónicos de visualización, proceso habitual en un equipo de tratamiento digital (Chuvieco 1990), cada color aparece en sus componentes espectrales primarias: rojo, verde y azul. Este modelo está basado en el sistema de coordenadas cartesianas. Como se muestra en la figura 5.6, los valores de R, G, y B se encuentran a lo largo de tres ejes, de modo que la suma de dos colores primarios permite lograr un color complementario.



*Figura 5.6. Plano RGB.
Fuente: Elaboración propia*

Todos los valores de R, G y B están en el intervalo $[0,1]$. En el caso de imágenes digitales los valores de R, G y B son números enteros y van de 0 a 255, lo cual permite generar 16 777 216 colores.

- *Modelo HSB*

Otro modelo de definición del color es el modelo HSV, también llamado HSB, el cual está basado en el trabajo de Albert Munsell y sus estudios de la percepción humana del color, definiendo los colores en función de las tres propiedades del color (matiz, luminosidad y saturación).

La figura 5.7 representa este modelo de color donde el matiz o tonalidad, son un grado de ángulo en sentido contrario a las manecillas del reloj alrededor del cono de color cuyos

valores posibles van de 0 a 359° (aunque para algunas aplicaciones se normalizan del 0 al 100%). Cada valor corresponde a un color. Ejemplos: 0° es rojo, 60° es amarillo y 120° es verde. (Calvo 2016).

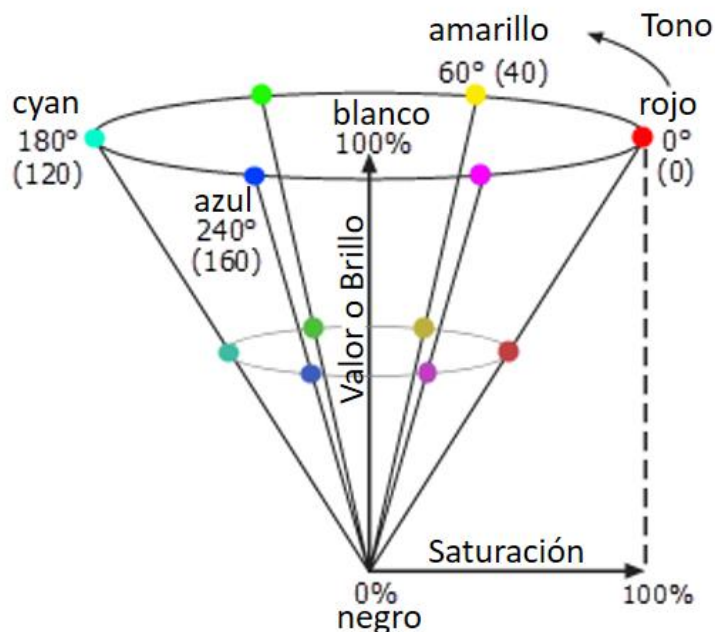


Figura 5.7. Componentes del Modelo HSB.
Fuente: Elaboración y traducción propia basado de (Desktop 2021)

La saturación se representa como la distancia al eje de brillo negro-blanco, es decir, la saturación de un color hace referencia al grado de extensión con respecto a un color neutro como el gris. Los valores posibles van del 0 al 100%. A este parámetro también se le suele llamar “pureza” por la analogía con la pureza de excitación y la pureza de color. Cuanto menor sea la saturación de un color, mayor tonalidad grisácea habrá y más decolorado estará (Calvo 2016).

Finalmente, el brillo o luminosidad representa la altura en el eje negro-blanco y especifica la intensidad del blanco en el color. Los valores posibles van del 0 al 100%. 0 siempre es negro. Dependiendo de la saturación, 100 podría ser blanco o un color más o menos saturado. Es común que deseemos elegir un color adecuado para alguna de nuestras aplicaciones; cuando es así, resulta muy útil usar la ruleta de color HSB (Calvo 2016).

5.3 MARCO ESPACIAL

El marco espacial del presente proyecto de investigación y a su vez la limitación que tendrá el modelo se basa en que únicamente se podrá detectar la enfermedad *Xanthomonas Campestris*.

5.3.1 Enfermedad *Xanthomonas Campestris*.

El género *Xanthomonas* es una de las bacterias de mayor cuidado e importancia epidemiológica en las zonas tropicales (Zapata, Mildred; Gaud 2001). *Xanthomonas campestris* es una bacteria patógena, la cual es capaz de desarrollarse sobre las hojas de los diferentes vegetales que ella contagia, está considerada de cuarentena y sumo cuidado por la unión europea (Conselleria d'Agricultura and Ecològica 2007), por lo cual la clasificación temprana de esta enfermedad ha sido motivo de diversos estudios taxonómicos (Zapata, Mildred; Gaud 2001).

De acuerdo la investigación realizada por (Conselleria d'Agricultura and Ecològica 2007), las plantas infectadas por *Xanthomonas Campestris* se caracterizan porque:

- Los síntomas inicialmente aparecen en las hojas con pequeñas manchas entre 1 y 5 mm de diámetro, posteriormente se puede expandir a frutos y ramas.
- Los primeros síntomas sobre el fruto usualmente aparecen entre la tercera y quinta semana después de la infección.

Así mismo, esta enfermedad se caracteriza por no tener un medio de lucha directa y efectiva contra ella, por lo cual la mejor forma de actuar es de forma preventiva, y en caso de detectar la enfermedad es recomendable eliminar las ramas afectadas, desinfectar los utensilios que se usaron para la poda y quemar los restos de la poda (Conselleria d'Agricultura and Ecològica 2007).

6. LIMITACIONES

Para el presente proyecto se tienen identificadas las siguientes limitaciones:

- El modelo se centrará únicamente en la detección de la enfermedad *Xanthomonas Campestris* cuando esta se manifiesta con síntomas que afectan la apariencia de las hojas de las plantas.
- El prototipo final que hace uso del modelo desarrollado requerirá que imágenes sean tomadas previamente a su entrenamiento y evaluación, debido a que evaluarla en tiempo real requeriría un módulo dedicado a la detección y segmentación de las hojas de la planta, incrementando la carga computacional, el tiempo de respuesta y aumentando el alcance del presente proyecto.

7. DESARROLLO DE LA INVESTIGACIÓN

7.1 ANÁLISIS E INVESTIGACIÓN.

Xanthomonas Campestris es una bacteria causal de varias fitopatologías, que al momento de infectar una planta genera lesiones o manchas foliares amarillas en forma de V, irregular o angulosa con venas negras y generalmente con centros necróticos en la superficie de sus hojas e incluso el tallo (Vicente et al. 2001), en estudios como el de (Matsunaga et al. 2017) se observaron otros síntomas, como distribuciones características en halo amarillo junto con lesiones marrones en canales rojos y verdes desde los primeros días del contagio como se puede observar en la figura 7.1.



Figura 7.1. Infección de Xanthomonas Campestris en hojas de cultivo Calendula officinalis
Fuente: Elaboración propia

Al evidenciarse síntomas notorios al poco tiempo del contagio de *Xanthomonas Campestris* en las hojas, es posible proponer una solución alternativa a las biológicas para asignar un diagnóstico rápido y preciso a partir de imágenes, por ejemplo, en (Ali et al. 2017) los investigadores adquirieron imágenes de hojas en diferentes huertos utilizando una cámara digital para capturar diversidad de muestras y obtener datos que les permitieran identificar y clasificar enfermedades en los cítricos de acuerdo con los síntomas de cada patógeno, usando reconocimiento de patrones, procesamiento de imágenes y técnicas de aprendizaje automático, logrando así una precisión del 99,7%.

Del mismo modo, para el desarrollo del modelo propuesto, se utilizó un conjunto de datos que consta de 1471 imágenes aproximadamente, las cuales fueron tomadas de 15 especies distintas que se destacan por su facilidad de contraer la enfermedad causada por

Xanthomonas Campestris y cuyos cultivos por lo general son masivos, como por ejemplo Mangifera indica, Brassica oleracea o Corylus avellana.

En la selección de imágenes se tuvo en cuenta hojas de plantas sanas y enfermas, con las cuales el algoritmo aprenderá las principales características visibles de la enfermedad a partir de reconocimiento de patrones, es decir, el modelo tendrá la capacidad de reconocer una escala de color que se asocie a los síntomas presentados en las hojas de las plantas que contengan la enfermedad, evaluando cada pixel de la imagen como se muestra en la figura 7.2. Para el desarrollo del modelo, las imágenes fueron adquiridas en RGB, el cual comprende una escala de color de 0 a 255 colores posibles y posteriormente se determina el color de acuerdo a las características propias de la enfermedad en condiciones naturales.

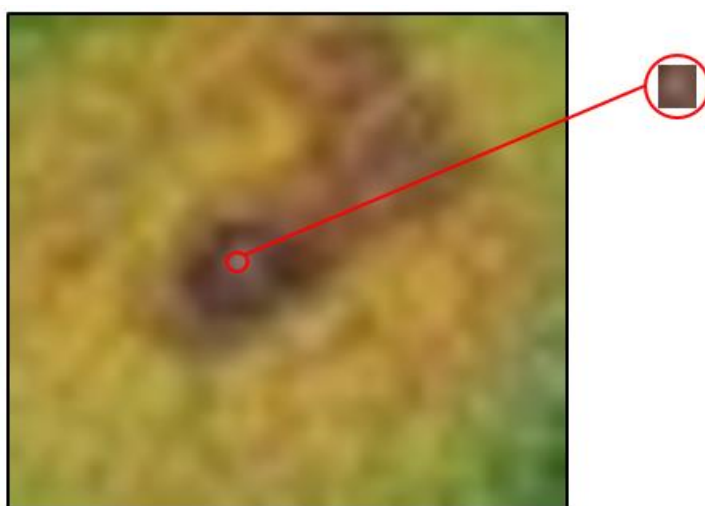


Figura 7.2. Reconocimiento de patrones por pixel de acuerdo a una escala de color asociada en una planta enferma.

Fuente: Elaboración Propia

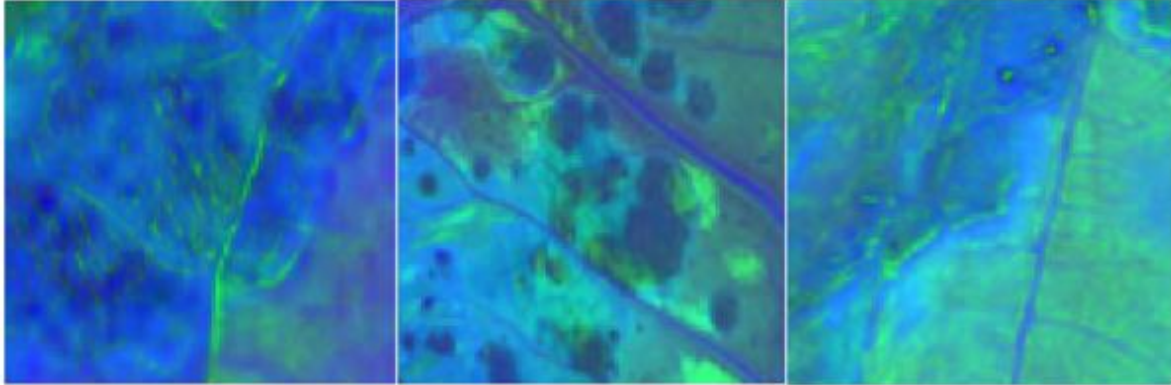
De acuerdo con (Sharif et al. 2018), si las imágenes que se utilizan para entrenar un modelo de aprendizaje automático son procesadas previamente, hay una gran posibilidad que al mejorar la calidad visual de la imagen de entrada se optimice el análisis del color de los síntomas de la enfermedad, y a su vez los resultados del diagnóstico. Con base a esto y a lo documentado por (Barbedo, Koenigkan, and Santos 2016) donde indica que las características de color son más importantes para la clasificación de la región infectada en la imagen porque cada punto de lesión tiene un color diferente, en el modelo propuesto se realizó transformación del color por imagen de RGB a HSB normalizado estableciendo valores de umbral y asignando una escala de color de amarillos, rojos y marrones la cual se muestra en la tabla 7.1, determinando los síntomas de la enfermedad en las distintas

especies, lo cual permite una mejoría en el análisis de las características que se obtienen mediante el uso de diferentes tipos de espacios de color y parámetros, en (Matsunaga et al. 2017) este proceso resultó ser bien perceptivo para detectar *Xanthomonas Campestris* en la Soya, y proporcionó mejores resultados en la detección de marrones y halos amarillos en hojas verdes, sin importar aspectos medioambientales como la luz y reflexión de la imagen.

*Tabla 7.1. Relación de HSB Normalizado por píxel de acuerdo a significado enfermo
Fuente: Elaboración propia*

H	S	B	Significado
0.170833	0.3125	0.501961	Enfermo
0.251773	0.319728	0.576471	Enfermo
0.0897436	0.619048	0.658824	Enfermo
0.151852	0.335821	0.52549	Enfermo
0.108247	0.433036	0.878431	Enfermo
0.154472	0.347458	0.462745	Enfermo
0.191406	0.618357	0.811765	Enfermo
0.140562	0.491124	0.662745	Enfermo

Posteriormente a la transformación del color a modelo HSB, las imágenes de hojas que se encuentren enfermas son perceptibles al algoritmo como se muestra en la figura 7.3, en ella se puede observar cómo el algoritmo recibirá información de las manchas características de la enfermedad causada por *Xanthomonas Campestris* de acuerdo a la escala de color dada anteriormente.



*Figura 7.3. Transformación de color a modelo HSB de hojas con significado enfermo.
Fuente: Elaboración Propia*

Para entrenar el modelo, al igual que para la identificación de síntomas se asigna al algoritmo una escala de color correspondiente para la detección de plantas sanas con una escala HSB normalizado como se muestra en la tabla 7.2, con el fin de fortalecer el análisis del algoritmo y la precisión del diagnóstico.

*Tabla 7.2. Relación de HSB Normalizado por píxel de acuerdo a significado sano
Fuente: Elaboración propia*

H	S	B	Significado
0.309524	0.234899	0.584314	Sana
0.243986	0.76378	0.498039	Sana
0.241935	0.659574	0.552941	Sana
0.231183	0.849315	0.572549	Sana
0.274691	0.421875	0.501961	Sana
0.289683	0.617647	0.266667	Sana
0.281046	0.389313	0.513725	Sana
0.291667	0.358974	0.305882	Sana

Al completar el entrenamiento a partir de la escala de color asignada para cada posible caso, el algoritmo tendrá la capacidad de discriminar las plantas que no presenten manchas características de la enfermedad y que a su vez su color se encuentre dentro del rango de

color asignado para hojas sanas, como se muestra en la figura 7.4. Culminada la fase de entrenamiento con los rangos de color asignados, el modelo propuesto tendrá mayor capacidad para el reconocimiento de patrones asociados a la enfermedad, teniendo en cuenta factores medioambientales como la luz y sombra, y por lo tanto una mejor toma de decisión al establecer si se enfrenta a una hoja sana o enferma.

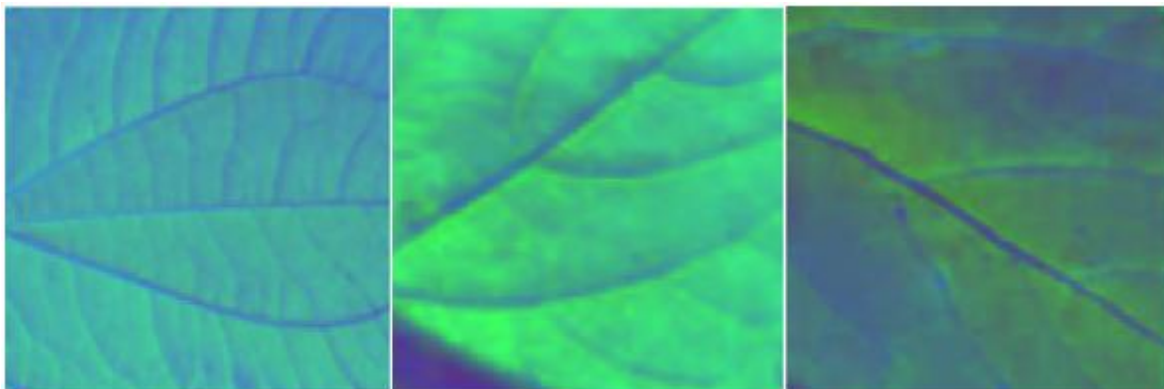


Figura 7.4. Transformación de color a modelo HSB de hojas con significado sano.

Fuente: Elaboración Propia

7.1.1 Alcances y limitaciones del modelo

La detección temprana de enfermedades en las plantas, ha sido un avance tecnológico muy importante para la agricultura, debido al crecimiento de distintos modelos que buscan favorecer el rendimiento de los cultivos y a su vez la productividad de los mismos en el mercado. Sin embargo, aunque parece no existir una técnica que funcione para todos los cultivos y que sea capaz de diagnosticar cualquier enfermedad en diferentes circunstancias (Nturambirwe and Opara 2020), el modelo propuesto para la detección de enfermedades causadas por *Xanthomonas Campestris* tiene la capacidad de reconocer los síntomas de estas por medio de imágenes sin importar el dispositivo con el que fueron tomadas, siempre y cuando se detalle correctamente la hoja, ya sea una cámara digital, una cámara de dispositivo móvil o un scanner, ya que el modelo recibe información de diferentes resoluciones espaciales, es decir, le permite al usuario conocer si una planta está enferma o no sin importar que tan clara esté la imagen o que tan pequeña sea esta, lo cual beneficia a aquellos agricultores que cuenten con dispositivos de adquisición de imagen sencillos.

Por otra parte, el modelo también está en capacidad de detectar la enfermedad *Xanthomonas Campestris* a partir de una imagen de la hoja del cultivo con la posible enfermedad, teniendo en cuenta factores medioambientales como la luz y sombra de la imagen, tomada previamente, de modo que, si en caso dado no se presenta en la imagen algún síntoma definido en el algoritmo, la implementación debe dar como resultado que el estado de la planta refleja que se encuentra sana.

Finalmente, una de las características biológicas más importantes de la bacteria *Xanthomonas Campestris* es que esta tiene tendencia a implantarse en cultivos en su mayoría vegetales, lo cual abarca muchas especies y tipos de hoja, el modelo propuesto a pesar de ser entrenado inicialmente con 15 especies aproximadamente, también tiene la capacidad de diagnosticar plantas de diferentes especies mientras estas presenten síntomas visibles similares a lesiones o manchas foliares dentro del rango de color con el que el modelo fue entrenado principalmente, favoreciendo a muchos agricultores no reconocen o detectan a tiempo la enfermedad.

De igual manera, a nivel de limitación, se establece que el modelo debe recibir de entrada imágenes que previamente hayan sido preprocesadas por un sistema de detección de bordes, a fin de que esté bien definido el contorno de la hoja y que en lo posible no tengan fondo, esto debido a que objetos de más a la región de interés pueden llegar a confundir el algoritmo o crear información redundante que posteriormente genere un diagnóstico erróneo del estado de la planta.

Por último, que inicialmente el requiere que las imágenes sean tomadas previamente a su evaluación, debido a que evaluar la imagen en tiempo real requeriría un módulo dedicado a la detección y segmentación de las hojas de la planta, incrementando la carga computacional y tiempo de respuesta.

7.1.2 Herramientas y aplicaciones en la agricultura

La agricultura es uno de los sectores económicos más importantes para la producción interna de un país, en ella se desarrollan actividades como la siembra, cuidado y recolección de vegetales para su posterior consumo, lo cual depende en gran medida de los pequeños agricultores, quienes son responsables del 80% del rendimiento total de los cultivos, y de los cuales el 50% de este se pierde debido a las plagas y enfermedades en los mismos (Yadav, Kumar Rana, et al. 2019). Por ello, una de las problemáticas relevantes en el cuidado de las plantas es la detección temprana de enfermedades, ya que por lo general

los agricultores poseen cultivos de distintas especies y cuidan de estos por medio de conocimientos empíricos, reduciendo la posibilidad de detectar enfermedades a partir de síntomas apreciables planta por planta y eventualmente ampliando la probabilidad de propagación y destrucción del cultivo, lo que implica pérdidas económicas y de tiempo en la producción.

El modelo propuesto es una respuesta rápida, precisa e interpretable para la detección temprana de enfermedades causadas por la bacteria *Xanthomonas Campestris* en cultivos principalmente vegetales a partir de imágenes de hojas de las plantas y técnicas de aprendizaje automático mejoradas mediante algoritmos de optimización, cuyo objetivo es brindar un diagnóstico acertado de la condición de una planta a partir de los síntomas que ésta presente por medio de sus hojas, de modo que la implementación de este modelo permita que el agricultor aisle a tiempo la planta enferma de las sanas y reduzca las pérdidas en sus cultivos.

7.1.3 Adquisición y pre-procesamiento de datos

De acuerdo con la literatura, para el desarrollo de algoritmos de aprendizaje automático se necesita de un conjunto de datos muy grande, teniéndose en cuenta un porcentaje para su entrenamiento y prueba. En este caso, donde la cuestión clave es el reconocimiento de las regiones que presenten síntomas asociados a la aparición de *Xanthomonas Campestris* en las plantas, se han adquirido más de 1000 imágenes de hojas de plantas enfermas y sanas por medio de tres fuentes diferentes: una cámara digital en escenarios aleatorios de luz, ángulo, sombra o brillo, ImageNet y el conjunto de datos públicos PlantVillage, este último se utilizó más debido a que las imágenes contenidas en este tienen las características de un entorno de fondo único y condiciones de iluminación estables (Xiong et al. 2020), a su vez allí se han publicado imágenes que se han usado para fomentar la investigación en esta área y comparar los resultados entre diferentes técnicas (Hernández and López 2020).

Posteriormente, al recopilar la mayor cantidad de datos, se realizaron recortes específicos de diferentes tamaños y almacenados en modelo de color RGB, este preprocesamiento se hace con el fin de que las imágenes de entrada al modelo no tuviesen fondo que pudiera confundir el modelo o saturarlo de información innecesaria.

En la figura 7.5, se muestran algunas de las imágenes con síntomas visibles de *Xanthomonas Campestris* en sus hojas, las cuales fueron adquiridas por distintas fuentes y seguidamente recortadas.

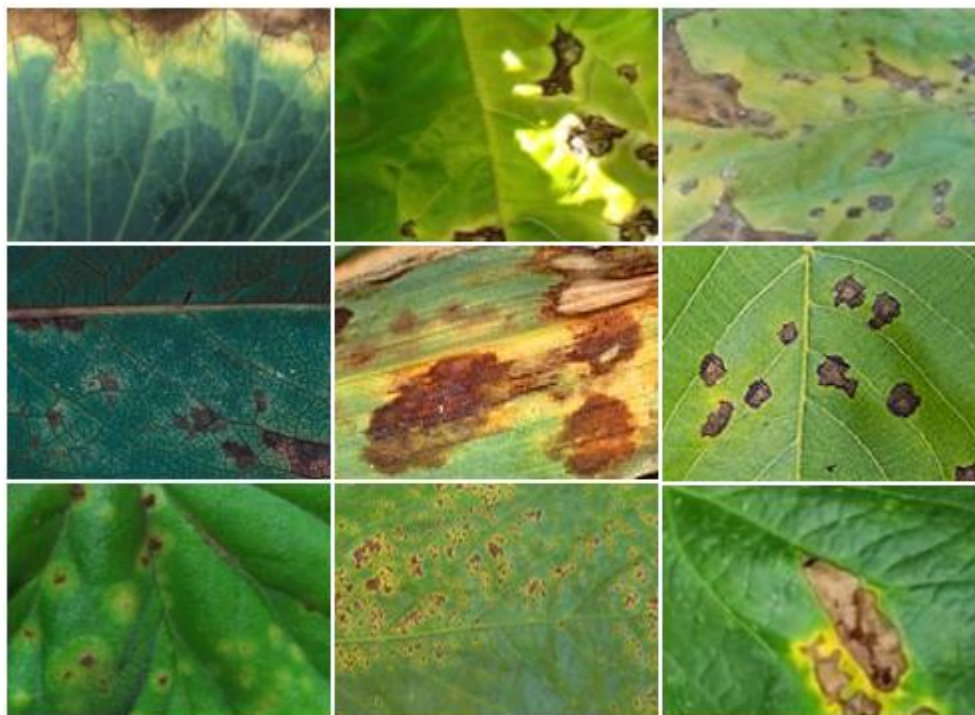


Figura 7.5. Muestra del conjunto de datos de plantas enfermas adquiridas en distintas fuentes para desarrollar el modelo.

Fuente: Elaboración propia a partir de datos adquiridos en Plant village, Imagenet, y fotos propias.

En el caso de las hojas sanas la mayor cantidad de píxeles se encuentran en el rango verde presentando variaciones de luz, contraste, ángulo y sombra como se muestra en la figura 11, esto se realizó con el fin de que el modelo tenga la capacidad de reconocer las diferentes características de las hojas sanas en diferentes momentos del día y pueda dar un diagnóstico más preciso.



Figura 7.6. Muestra del conjunto de datos de plantas sanas adquiridas en distintas fuentes para desarrollar el modelo.

Fuente: Elaboración propia a partir de datos adquiridos en PlantVillage, Imagenet, y fotos propias.

7.1.4 Selección y procesamiento de características

Tras la adquisición y preprocesamiento inicial de las imágenes, se requiere de la transformación del color, el análisis de los datos y el desarrollo de un algoritmo neurodifuso que finalmente permitirá dar un diagnóstico del estado de la planta, para estas tareas se seleccionó la plataforma MATLAB 2019b el cual brinda un mejor desarrollo a comparación de otros softwares para modelos de lógica difusa. Por ejemplo, en (D. Guzmán 2006) se utilizan herramientas como “control difuso” para desarrollar una simulación idéntica a la de la arquitectura Mamdani con la diferencia de que el editor FIS se elige la arquitectura Sugeno y las funciones de membresía de salida son las constantes seleccionadas. Adicionalmente MATLAB tiene una comunidad que presta el soporte en caso de ser requerido.

7.2 DISEÑO Y MODELADO

Como se evidencia en (Leal 2021), con respecto a los modelos dispuestos para la detección de enfermedades en las plantas a partir de imágenes, hay un vacío en el campo de la interpretabilidad, ya que los autores que han realizado modelos similares aplicados en distintas enfermedades, se han enfocado principalmente en obtener resultados cada vez más precisos en sus modelos, más allá de brindar al usuario una explicación clara de las características que se reconocen en su algoritmo o de qué es lo que afecta realmente los resultados que se obtienen en la clasificación. Por ello, se destaca la importancia de tener en cuenta este factor dentro la presente investigación sin dejar a un lado la relevancia de la precisión dentro del mismo conformando así una contribución a la agricultura y la inteligencia artificial, en esta sección se describe la propuesta y la implementación del modelo para la detección de *Xanthomonas Campestris* de manera no destructiva a partir de la aparición de síntomas de la enfermedad en hojas.

7.2.1 Construcción del modelo

El modelo propuesto se compone de dos sistemas neurodifusos los cuales, trabajando de manera conjunta permiten determinar por medio de una imagen de la hoja de una planta, si en esta se encuentra la bacteria *Xanthomonas Campestris*. El primer sistema neurodifuso permite detectar, a nivel de pixel en escala HSB perteneciente a una imagen de una hoja, los rasgos específicos relacionados con la tonalidad del color de la enfermedad *Xanthomonas Campestris*, para que, con esta salida, el segundo sistema pueda con base en los resultados totales de una imagen tomar una decisión de acuerdo al estado de la hoja.

Para cada una de las propuestas se realiza la construcción de varios modelos usando sistemas caracterizados por su precisión, pero también por su interpretabilidad como lo son los sistemas neurodifusos junto con un mecanismo de optimización híbrido que implementa algoritmos de optimización y el algoritmo de punto interior. Al mismo tiempo, se realiza una comparación con la utilización de mecanismos de agrupamiento difuso para cada una de las configuraciones.

La configuración que se realizó para los sistemas de inferencia Sugeno sin agrupación de Clúster se describen en la tabla 7.3, donde se relaciona el valor dado a cada uno. Estos valores fueron utilizados en la función propia de MATLAB *sugfis*.

*Tabla 7.3. Configuración de parámetros de inferencia difusa Sugeno sin agrupación de clúster.
Fuente: Elaboración propia*

Parámetro	Valor sin Cluster Difuso
Método de operador Difuso And (And method)	El producto de valores de entrada difusos
Método de operador Difuso OR (Or method)	El máximo de valores de entrada difusos.
Método de Defuzzificación para calcular valores de salida nítidos (Defuzzification)	Promedio ponderado de todas las salidas de reglas
Método de implicación para calcular el conjunto difuso consecuente (Implication)	Escala la función de pertenencia consecuente por el valor del resultado antecedente
Método de agregación para combinar consecuentes de reglas (Aggregation)	Suma de conjuntos difusos consecuentes

Por otro lado, para aquellos sistemas que estén implementados utilizando Clustering Difuso, la configuración que se siguió es la observada en la tabla 7.4. Estos parámetros a su vez fueron ingresados en la función de MATLAB *genfis*.

*Tabla 7.4. Configuración de parámetros de inferencia difusa Sugeno con agrupación de clúster.
Fuente: Elaboración propia*

Parámetro	Valor con Cluster Difuso
Método de operador Difuso And (And method)	El producto de valores de entrada difusos
Método de operador Difuso OR (Or method)	OR probabilístico de valores de entrada difusos
Método de Defuzzificación para calcular valores de salida nítidos (Defuzzification)	Promedio ponderado de todas las salidas de reglas

Método de implicación para calcular el conjunto difuso consecuente (Implication)	Escala la función de pertenencia consecuente por el valor del resultado antecedente
Método de agregación para combinar consecuentes de reglas (Aggregation)	Suma de conjuntos difusos consecuentes

7.2.2 Sistema de inferencia difusa ANFIS

Posteriormente a la elaboración de los sistemas de inferencia difusa Sugeno, se realiza la creación de los sistemas de inferencia difusa tipo ANFIS, el cual tiene la capacidad de reducir considerablemente el tiempo de modelamiento y los requerimientos de los conocimientos de expertos (Correa-Henao and Montoya-Suárez 2013). Estos sistemas toman como entrada los sistemas tipo Sugeno para crear el sistema neurodifuso. Los parámetros necesarios para su ejecución se indican en la tabla 7.5. Estos valores se utilizan en la función de MATLAB denominada ***anfisOption*** para posteriormente crear el sistema ANFIS utilizando la función ***anfis*** son ingresados en la función de MATLAB ***anfisOptions***.

Tabla 7.5. Configuración de parámetros de sistema de inferencia difusa ANFIS
Fuente: Elaboración propia

Parámetro	Configuración
Número máximo de épocas de entrenamiento	600
Tamaño del paso de entrenamiento inicial	0.1
Método de optimización	Para la entrada, Backpropagation y para la salida, mínimos cuadrados.
Datos de validación	Los diferentes valores de prueba tanto de entrada como la salida esperada de manera tal que se evite el sobreajuste del modelo.

7.2.3 Configuración de los sistemas difusos tipo Sugeno

El modelo a implementar se divide en dos sistemas que permiten, de manera conjunta, permiten identificar la enfermedad a partir de una imagen. La configuración de estos sistemas denominados sistema 1 y sistema 2, se muestra a continuación.

Sistema 1:

La configuración del Sistema 1, fue diseñada para recibir una entrada individual a nivel de pixel convertido a escala HSB, dado lo anterior, se proponen tres conjuntos difusos para la entrada del sistema Sugeno que corresponden a:

H: Matiz.

S: Saturación.

B: Brillo.

A su vez, cada conjunto difuso lo conforman tres funciones de pertenencia que representan los valores de bajo, medio y alto para cada valor de la escala HSB. Para la salida, por defecto, se configura una función tipo lineal.

Teniendo en cuenta que se van a elaborar varios tipos de sistemas, para los sistemas que utilizan Cluster difuso, las reglas se definen automáticamente, de tal manera que se relacione cada una de las entradas con las salidas de manera uniforme, con un límite de reglas de 27 como máximo.

Para el sistema implementado sin uso de agrupamiento difuso, utilizando análisis exploratorio de los diferentes valores HSV y su significado a nivel de píxel enfermo o sano, se definen las siguientes reglas:

- Si el valor de H es bajo y S es bajo y B bajo es entonces Sano.
- Si el valor de H es bajo y S es medio y B bajo es entonces Sano.
- Si el valor de H es bajo y S es medio y B medio es entonces Sano.
- Si el valor de H es bajo y S es medio y B es alto entonces Enfermo.
- Si el valor de H es medio y S es medio y B es alto entonces Enfermo.
- Si el valor de H es bajo y S es alto y B es alto entonces Enfermo.

- Si el valor de H es medio y S es alto y B es alto entonces Enfermo.

Sistema 2:

La configuración realizada para el Sistema 2, permite determinar con base en el promedio de los resultados obtenidos a nivel de píxel de toda una imagen del primer sistema, si una imagen se puede evidenciar la enfermedad Xanthomonas Campestris o no, la entrada del sistema es:

Prom: Promedio de valores de los resultados del sistema 1.

El conjunto difuso lo conforman cinco funciones de pertenencia que representan los valores de bajo, medio bajo, medio, medio alto y alto. Para la salida, por defecto, se configura una función tipo linear.

Teniendo en cuenta que se van a elaborar varios tipos de sistemas, para los sistemas que utilizan Cluster difuso, las reglas se definen automáticamente, de tal manera que se relacione cada una de las entradas con las salidas de manera uniforme, con un límite de reglas de 27 como máximo.

Para el sistema implementado sin uso de agrupamiento difuso, utilizando análisis exploratorio de los diferentes valores HSB y su significado a nivel de pixel enfermo o sano, se definen las siguientes reglas:

- Si el valor de prom es alto entonces Sana.
- Si el valor de prom es medio alto entonces Sana.
- Si el valor de prom es medio entonces Sana.
- Si el valor de prom es medio bajo entonces Enferma.
- Si el valor de prom es bajo entonces Enferma.

7.2.4 Implementación de agrupamiento difuso

En la literatura se considera que el agrupamiento difuso es la técnica más adecuada para la superposición de tareas de agrupamiento en la obtención de los sistemas difusos, en ella el procesamiento de iteración se realiza en la imagen de la hoja de la enfermedad completa como se desarrolló en (Zhang, You, and Wu 2019) obteniendo buenos resultados.

En Matlab, esta clase de agrupamiento se utiliza usando la función **genfisOptions**, cuya configuración se establece en la siguiente tabla:

*Tabla 7.6. Parámetros de agrupamiento difuso FCMClustering
Fuente: Elaboración propia*

Parámetro	Configuración
Número de Clusters	5 para el sistema 2 3 para el sistema 1
Type	FCMClustering

7.3 OPTIMIZACIÓN Y PRUEBAS.

Con el fin de mejorar la precisión de los diferentes sistemas configurados anteriormente, se procede a realizar la optimización utilizando técnicas híbridas, que mezclan métodos tanto heurísticos como exactos, junto con su respectiva validación que permita establecer la configuración que mejor resultados de para conformar el modelo final.

7.3.1 Implementación de técnicas de optimización

Para reducir las pérdidas de información suministrada en las imágenes y proporcionar los resultados más precisos posibles a partir de un diagnóstico rápido, fue esencial optimizar el modelo para tener un control eficaz sobre el mismo, esta optimización se llevó a cabo con una implementación de algoritmo híbrido que tiene la capacidad de determinar puntos candidatos a partir de píxeles sanos obtenidos dilatando el punto candidato dos veces sin utilizar los píxeles obtenidos a partir de un algoritmo de punto interior al terminar la ejecución del algoritmo genético.

Los valores de los parámetros del algoritmo genético se encuentran descritos en la tabla 7.7, los cuales fueron utilizados para crear una opción de optimización en Matlab usando la función **optimoptions**. Estos parámetros se establecieron teniendo en cuenta que, las recomendaciones encontradas en la literatura por los diferentes expertos. En cuanto al tamaño de la población, teniendo en cuenta que cuando se tiene un tamaño de población grande, permite que algoritmo genético realice una búsqueda más amplia en todo el espacio de la solución, reduciendo de esta manera la probabilidad de que el algoritmo devuelva un

mínimo local que no sea un mínimo global. En (“Genetic Algorithm Options - MATLAB & Simulink,” n.d.), se recomienda que:

- Para el tamaño de la población, el valor recomendado debido el número de variables es de 50, sin embargo, para ampliar el rango de búsqueda en el campo solución, este valor se configura en 100.
- Se establece que, para los algoritmos genéticos, la fracción de cruce que produce los mejores resultados, es de 0.8.
- Con respecto a los individuos elites que sobreviven para la siguiente iteración, se define que este valor es el 0.05 del tamaño de la población, por ende, se asigna el valor de 5.

Con respecto a los criterios de detención del algoritmo, se estableció un límite de 300 segundos para detener una iteración. Y, teniendo en cuenta que se va a implementar una técnica híbrida, y que la salida del algoritmo genético será igualmente optimizada por un algoritmo de punto interior, se establece que el número máximo de iteraciones es de 100.

*Tabla 7.7. Parámetros para implementación de Algoritmo Genético
Fuente: Elaboración propia*

Parámetro	Valor
Tamaño de la población	100
Número máximo de iteraciones antes de que el algoritmo se detenga	100
Individuos de la generación actual que tienen garantizado sobrevivir hasta la próxima generación	5
Fracción de la población en la próxima generación, sin incluir a los que se les garantizo la supervivencia, que crea la función cruce	0.8
Tiempo para que el algoritmo se detenga después de ejecutarse en segundos	300

Por su parte, Matlab también tiene la función **hybridFcn** cuyos parámetros se encuentran relacionados en la tabla 7.8, y la cual se ha elegido para ser implementada junto con el algoritmo de punto interior, esto debido a que este algoritmo permite continuar la optimización contando con la solidez del modelo, el aumento de las capacidades de resolución de problemas y las mejoras de rendimiento. Debido a esto, según lo

recomendado en (“Tolerances and Stopping Criteria - MATLAB & Simulink,” n.d.), se establece un valor de 3000 con respecto al número máximo de evaluaciones de las funciones, y, teniendo en cuenta que se está implementando en el marco de una técnica híbrida, la cantidad de iteraciones se establece en 30.

*Tabla 7.8. Parámetros para implementación de Algoritmo de Punto Interior
Fuente: Elaboración propia*

Parámetro	Valor
Número máximo de evaluaciones de funciones	3000
Número máximo de iteraciones antes de que el algoritmo se detenga	30

7.3.2 Validación

Una vez realizado el proceso de definición de variables y configuración de los sistemas neurodifusos, se desarrollan un total de 4 modelos, 2 de ellos para el sistema que permite, a partir de una entrada a nivel de Pixel en escala HSV, obtener un valor sobre el estado del pixel, mediante el uso de Clustering difuso para un modelo, y realizando la definición propia de las reglas y la configuración de los conjuntos difusos para el segundo modelo. Los otros 2 modelos faltantes se realizaron para el sistema que permite, a partir de la media aritmética obtenida por todos los pixeles de una imagen pasados por el primer sistema, el estado de la hoja que representa dicha imagen, teniendo definido un rango de valores para el estado enfermo entre 0.3884 y 0.54, y para el estado sano entre 0.55 y 0.70.

Cada modelo es ejecutado en 50 instantes de tiempo con el fin de comparar los diferentes resultados a nivel de los valores de sus errores máximo, mínimo, MSE, RMSE y MAE, tanto para los conjuntos de datos de entrenamiento como de validación, debido a, que algunos de ellos convergen hacia los mínimos locales y no al objetivo que es el mínimo global. En la tabla 7.9 y 7.10 se encuentran los valores de error del primer sistema con y sin clustering, mientras que en las tablas 7.11 y 7.12 se encuentran los valores de error del segundo sistema con y sin clustering. En las tablas de cada modelo se encuentra el promedio del error cuadrático medio obtenido en fases de prueba y validación, y a su vez se destaca la ejecución que muestra mejores resultados dentro de cada modelo.

Tabla 7.9. Valores de error del primer sistema neurodifuso con clustering para la detección de Xanthomonas Campestris
Fuente: Elaboración propia

Modelo ANFIS con Clustering											
Ejecución	Entrenamiento					Validación					\overline{MSE}
	MSE	RMSE	MAE	MIN	MAX	MSE	RMSE	MAE	MIN	MAX	
1	4,12E-03	6,42E-02	2,32E-05	1,58E-07	9,53E-01	9,04E-03	9,51E-02	2,99E-05	5,95E-07	7,91E-01	6,58E-03
2	9,43E-03	9,71E-02	2,85E-05	4,05E-07	1,04E+00	1,45E-02	1,20E-01	3,54E-05	1,21E-07	6,73E-01	1,20E-02
3	2,55E-02	1,60E-01	4,46E-05	1,84E-07	9,84E-01	2,40E-02	1,55E-01	4,49E-05	4,03E-07	1,28E-01	2,47E-02
4	4,03E-03	6,35E-02	2,31E-05	2,53E-07	9,35E-01	2,57E-03	5,07E-02	2,34E-05	1,51E-07	1,08E+00	3,30E-03
5	8,92E-03	9,45E-02	2,80E-05	1,31E-07	9,38E-01	2,62E-03	5,12E-02	2,35E-05	2,61E-07	8,91E-01	5,77E-03
6	1,62E-03	4,02E-02	1,91E-05	2,11E-08	9,70E-01	8,30E-04	2,88E-02	2,09E-05	5,94E-08	2,77E-01	1,22E-03
7	9,78E-03	9,89E-02	2,89E-05	3,40E-07	9,71E-01	7,69E-03	8,77E-02	2,86E-05	5,20E-07	3,83E-01	8,73E-03
8	4,10E-02	2,03E-01	6,01E-05	4,14E-07	9,74E-01	2,46E-02	1,57E-01	4,55E-05	6,28E-07	6,23E-01	3,28E-02
9	9,11E-03	9,54E-02	2,82E-05	3,46E-08	9,75E-01	9,47E-03	9,73E-02	3,03E-05	5,33E-07	5,90E-01	9,29E-03
10	7,08E-03	8,41E-02	2,62E-05	2,46E-07	9,74E-01	1,03E-02	1,01E-01	3,12E-05	6,57E-07	1,07E+00	8,68E-03
11	5,58E-03	7,47E-02	2,47E-05	7,94E-08	9,75E-01	2,02E-03	4,49E-02	2,29E-05	2,06E-07	2,83E-01	3,80E-03
12	8,53E-03	9,24E-02	2,76E-05	7,10E-08	9,77E-01	3,03E-03	5,51E-02	2,39E-05	5,61E-07	3,64E-01	5,78E-03
13	2,40E-03	4,90E-02	2,15E-05	2,91E-07	9,75E-01	5,48E-03	7,40E-02	2,64E-05	6,41E-07	3,31E-01	3,94E-03
14	9,11E-02	3,02E-01	1,10E-04	4,53E-07	9,77E-01	4,09E-03	6,39E-02	2,50E-05	1,35E-07	5,45E-01	4,76E-02

15	9,93E-03	9,97E-02	2,90E-05	7,15E-08	9,74E-01	9,97E-03	9,98E-02	3,08E-05	9,54E-08	9,65E-01	9,95E-03
16	6,65E-03	8,16E-02	2,57E-05	3,54E-07	9,67E-01	9,95E-04	3,15E-02	2,19E-05	1,67E-07	6,70E-01	3,82E-03
17	5,43E-02	2,33E-01	7,34E-05	3,75E-07	9,71E-01	2,47E-03	4,97E-02	2,33E-05	2,35E-07	6,34E-01	2,84E-02
18	1,06E-02	1,03E-01	2,97E-05	4,59E-07	9,75E-01	6,68E-03	8,18E-02	2,76E-05	3,26E-07	7,24E-01	8,66E-03
19	8,63E-03	9,29E-02	2,77E-05	2,31E-07	9,69E-01	7,56E-02	2,75E-01	9,64E-05	9,62E-08	8,33E-01	4,21E-02
20	2,48E-03	4,98E-02	2,16E-05	2,97E-07	9,72E-01	2,86E-03	5,35E-02	2,37E-05	3,87E-07	8,31E-01	2,67E-03
21	3,36E-03	5,79E-02	2,24E-05	6,20E-07	9,75E-01	1,01E-03	3,18E-02	2,19E-05	4,26E-07	1,90E-01	2,18E-03
22	2,13E-03	4,62E-02	2,12E-05	8,49E-08	9,78E-01	1,04E-02	1,02E-01	3,13E-05	1,64E-07	5,30E-01	6,29E-03
23	4,58E-03	6,77E-02	2,37E-05	2,33E-07	9,76E-01	5,37E-03	7,32E-02	2,62E-05	2,01E-07	8,81E-01	4,98E-03
24	5,82E-03	7,63E-02	2,49E-05	2,32E-07	9,73E-01	6,58E-02	2,56E-01	8,66E-05	4,84E-07	5,49E-01	3,58E-02
25	7,32E-02	2,71E-01	9,23E-05	5,79E-07	9,70E-01	9,79E-03	9,90E-02	3,07E-05	2,00E-07	6,54E-01	4,15E-02
26	2,47E-03	4,97E-02	2,16E-05	2,23E-07	7,15E-01	7,33E-03	8,56E-02	2,82E-05	6,36E-07	7,76E-01	4,90E-03
27	1,06E-02	1,03E-01	2,97E-05	4,56E-07	9,71E-01	4,82E-03	6,94E-02	2,57E-05	2,67E-07	7,85E-01	7,70E-03
28	8,81E-03	9,38E-02	2,79E-05	4,02E-07	9,71E-01	6,35E-03	7,97E-02	2,72E-05	2,29E-07	8,63E-01	7,58E-03
29	1,15E-02	1,07E-01	3,06E-05	1,61E-07	9,46E-01	4,65E-03	6,82E-02	2,55E-05	3,04E-07	4,30E-01	8,07E-03
30	9,99E-03	1,00E-01	2,91E-05	1,18E-07	9,77E-01	1,01E-03	3,17E-02	2,19E-05	1,76E-07	2,37E-01	5,50E-03
31	9,35E-03	9,67E-02	2,84E-05	5,21E-08	9,74E-01	4,47E-03	6,69E-02	2,53E-05	1,94E-07	8,36E-01	6,91E-03
32	5,13E-02	2,26E-01	7,04E-05	5,48E-07	7,69E-01	5,87E-02	2,42E-01	7,96E-05	1,62E-07	9,68E-01	5,50E-02
33	2,45E-03	4,95E-02	2,15E-05	3,47E-07	9,71E-01	5,98E-02	2,45E-01	8,07E-05	5,80E-07	3,66E-01	3,11E-02
34	7,17E-03	8,47E-02	2,63E-05	5,99E-08	9,76E-01	1,07E-02	1,03E-01	3,16E-05	4,42E-07	5,53E-01	8,93E-03
35	6,50E-03	8,06E-02	2,56E-05	3,20E-07	9,72E-01	8,35E-03	9,14E-02	2,92E-05	1,99E-07	6,20E-01	7,43E-03
36	4,48E-02	2,12E-01	6,39E-05	4,01E-07	9,73E-01	6,18E-02	2,49E-01	8,27E-05	3,50E-07	7,24E-01	5,33E-02
37	4,96E-03	7,04E-02	2,40E-05	2,04E-07	9,70E-01	5,91E-03	7,69E-02	2,68E-05	6,57E-07	2,66E-01	5,43E-03
38	9,97E-03	9,98E-02	2,91E-05	3,17E-07	9,70E-01	2,87E-03	5,36E-02	2,38E-05	6,35E-07	1,58E-01	6,42E-03
39	4,06E-02	2,01E-01	5,97E-05	2,11E-07	9,74E-01	4,78E-02	2,19E-01	6,87E-05	2,16E-07	6,05E-01	4,42E-02
40	3,55E-03	5,96E-02	2,26E-05	4,28E-07	8,38E-01	2,03E-03	4,50E-02	2,29E-05	5,92E-07	7,36E-01	2,79E-03
41	9,05E-03	9,51E-02	2,81E-05	2,01E-07	9,72E-01	1,52E-03	3,89E-02	2,24E-05	1,11E-07	4,48E-01	5,28E-03
42	5,46E-02	2,34E-01	7,37E-05	4,56E-07	9,70E-01	6,34E-02	2,52E-01	8,42E-05	5,79E-07	9,07E-01	5,90E-02

43	1,13E-02	1,06E-01	3,04E-05	2,67E-07	9,71E-01	5,79E-03	7,61E-02	2,67E-05	5,00E-07	3,93E-01	8,55E-03
44	1,14E-02	1,07E-01	3,05E-05	5,32E-07	9,74E-01	7,65E-03	8,74E-02	2,85E-05	4,24E-07	3,40E-01	9,53E-03
45	5,26E-03	7,25E-02	2,44E-05	2,96E-07	9,77E-01	5,15E-03	7,17E-02	2,60E-05	2,34E-07	2,89E-01	5,21E-03
46	4,99E-03	7,07E-02	2,41E-05	1,46E-07	9,71E-01	4,24E-03	6,51E-02	2,51E-05	6,38E-07	7,82E-02	4,62E-03
47	8,79E-03	9,38E-02	2,79E-05	5,06E-07	9,74E-01	8,41E-03	9,17E-02	2,93E-05	1,69E-07	8,47E-01	8,60E-03
48	2,89E-03	5,38E-02	2,20E-05	1,04E-07	9,75E-01	9,79E-02	3,13E-01	1,19E-04	1,06E-07	1,07E+00	5,04E-02
49	4,15E-02	2,04E-01	6,05E-05	8,39E-08	9,77E-01	5,86E-03	7,65E-02	2,67E-05	4,79E-07	9,97E-01	2,37E-02
50	3,41E-03	5,84E-02	2,25E-05	1,35E-07	8,87E-01	6,26E-02	2,50E-01	8,35E-05	3,86E-07	1,63E-01	3,30E-02

Tabla 7.10. Valores de error del primer sistema neurodifuso sin clustering para la detección de *Xanthomonas Campestris*
Fuente: Elaboración propia

Modelo ANFIS sin Clustering											
Ejecución	Entrenamiento					Validación					\overline{MSE}
	MSE	RMSE	MAE	MIN	MAX	MSE	RMSE	MAE	MIN	MAX	
1	6,48E-02	2,55E-01	2,95E-04	4,59E-07	1,06E+00	2,94E-03	5,42E-02	2,09E-03	6,04E-06	3,54E-01	3,39E-02
2	6,54E-02	2,56E-01	2,96E-04	3,65E-07	9,25E-01	4,13E-02	2,03E-01	2,13E-03	6,14E-06	1,83E-01	5,33E-02
3	6,21E-02	2,49E-01	2,92E-04	2,13E-07	1,02E+00	9,32E-03	9,65E-02	2,10E-03	6,21E-06	8,71E-01	3,57E-02
4	1,03E-01	3,20E-01	3,33E-04	3,93E-07	1,03E+00	6,38E-03	7,99E-02	2,09E-03	6,37E-06	2,03E-01	5,45E-02
5	6,20E-02	2,49E-01	2,92E-04	7,94E-07	1,05E+00	4,77E-03	6,91E-02	2,09E-03	6,10E-06	3,26E-01	3,34E-02
6	6,12E-02	2,47E-01	2,91E-04	6,92E-07	2,33E-01	1,60E-03	4,00E-02	2,09E-03	6,53E-06	8,96E-01	3,14E-02
7	1,24E-01	3,53E-01	3,55E-04	5,10E-07	9,41E-01	6,10E-03	7,81E-02	2,09E-03	5,99E-06	5,36E-01	6,53E-02
8	1,42E-01	3,77E-01	3,72E-04	7,97E-07	1,02E+00	7,44E-03	8,63E-02	2,10E-03	6,32E-06	1,88E-01	7,47E-02
9	6,23E-02	2,50E-01	2,92E-04	4,89E-07	9,93E-01	2,55E-03	5,05E-02	2,09E-03	6,52E-06	6,65E-01	3,24E-02
10	6,06E-02	2,46E-01	2,91E-04	6,60E-07	9,56E-01	7,98E-03	8,93E-02	2,10E-03	6,23E-06	7,62E-01	3,43E-02
11	5,97E-02	2,44E-01	2,90E-04	5,85E-07	1,05E+00	9,83E-03	9,92E-02	2,10E-03	6,41E-06	1,05E+00	3,48E-02
12	6,38E-02	2,53E-01	2,94E-04	5,79E-07	1,06E+00	6,67E-03	8,17E-02	2,09E-03	6,28E-06	2,28E-01	3,52E-02
13	1,10E-01	3,31E-01	3,40E-04	5,70E-07	9,35E-01	3,49E-03	5,91E-02	2,09E-03	6,21E-06	8,79E-01	5,65E-02
14	6,38E-02	2,53E-01	2,94E-04	2,92E-07	1,05E+00	4,71E-03	6,86E-02	2,09E-03	5,99E-06	1,03E+00	3,43E-02

15	9,82E-02	3,13E-01	3,28E-04	3,55E-07	1,02E+00	1,15E-03	3,38E-02	2,09E-03	6,04E-06	1,86E-01	4,96E-02
16	5,87E-02	2,42E-01	2,30E-04	2,11E-07	9,90E-01	9,67E-04	3,11E-02	2,09E-03	5,94E-06	2,77E-01	2,98E-02
17	6,10E-02	2,47E-01	2,91E-04	3,69E-07	9,96E-01	1,09E-02	1,04E-01	2,10E-03	6,16E-06	8,85E-01	3,60E-02
18	6,85E-02	2,62E-01	2,99E-04	4,56E-07	9,92E-01	6,64E-03	8,15E-02	2,09E-03	6,32E-06	7,22E-01	3,75E-02
19	6,01E-02	2,45E-01	2,90E-04	7,54E-07	7,19E-01	1,03E-02	1,01E-01	2,10E-03	6,10E-06	7,06E-01	3,52E-02
20	6,34E-02	2,52E-01	2,94E-04	4,95E-07	9,93E-01	1,32E-02	1,15E-01	2,10E-03	6,17E-06	5,15E-01	3,83E-02
21	6,75E-02	2,60E-01	2,98E-04	6,76E-07	9,97E-01	7,48E-03	8,65E-02	2,10E-03	6,44E-06	4,65E-01	3,75E-02
22	6,60E-02	2,57E-01	2,96E-04	6,95E-07	9,91E-01	7,49E-03	8,65E-02	2,10E-03	5,99E-06	9,08E-01	3,68E-02
23	6,16E-02	2,48E-01	2,92E-04	3,17E-07	9,93E-01	1,09E-02	1,04E-01	2,10E-03	5,99E-06	1,05E+00	3,62E-02
24	6,53E-02	2,56E-01	2,95E-04	7,94E-07	9,97E-01	8,15E-02	2,85E-01	2,17E-03	6,10E-06	2,03E-01	7,34E-02
25	6,28E-02	2,51E-01	2,93E-04	3,75E-07	5,59E-01	7,33E-03	8,56E-02	2,10E-03	6,23E-06	5,44E-01	3,51E-02
26	6,30E-02	2,51E-01	2,93E-04	3,18E-07	9,95E-01	4,51E-03	6,71E-02	2,09E-03	6,35E-06	9,33E-01	3,37E-02
27	1,46E-01	3,82E-01	3,76E-04	3,77E-07	9,93E-01	6,76E-02	2,60E-01	2,16E-03	6,11E-06	6,29E-01	1,07E-01
28	6,40E-02	2,53E-01	2,94E-04	4,15E-07	9,91E-01	9,23E-03	9,61E-02	2,10E-03	5,98E-06	4,94E-01	3,66E-02
29	6,60E-02	2,57E-01	2,96E-04	6,98E-07	9,91E-01	8,67E-03	9,31E-02	2,10E-03	6,49E-06	9,07E-01	3,73E-02
30	6,43E-02	2,54E-01	2,95E-04	2,54E-07	9,96E-01	3,26E-02	1,81E-01	2,12E-03	6,06E-06	3,78E-01	4,85E-02
31	6,55E-02	2,56E-01	2,96E-04	7,77E-07	2,18E-01	5,67E-03	7,53E-02	2,09E-03	6,10E-06	4,35E-01	3,56E-02
32	1,34E-01	3,66E-01	3,64E-04	5,67E-07	9,92E-01	2,62E-03	5,12E-02	2,09E-03	6,37E-06	8,44E-01	6,83E-02
33	7,98E-02	2,82E-01	3,10E-04	7,35E-07	1,65E-01	9,78E-03	9,89E-02	2,10E-03	6,02E-06	7,97E-01	4,48E-02
34	6,41E-02	2,53E-01	2,94E-04	5,23E-07	9,90E-01	3,57E-03	5,97E-02	2,09E-03	6,22E-06	1,03E+00	3,38E-02
35	6,45E-02	2,54E-01	2,95E-04	4,89E-07	9,90E-01	8,39E-03	9,16E-02	2,10E-03	5,98E-06	7,72E-01	3,65E-02
36	6,58E-02	2,57E-01	2,96E-04	7,66E-07	9,92E-01	2,21E-03	4,70E-02	2,09E-03	6,27E-06	2,30E-01	3,40E-02
37	1,25E-01	3,53E-01	3,55E-04	2,85E-07	9,97E-01	5,90E-03	7,68E-02	2,09E-03	6,10E-06	4,31E-01	6,53E-02
38	6,57E-02	2,56E-01	2,96E-04	7,23E-07	9,73E-01	8,41E-03	9,17E-02	2,10E-03	5,95E-06	4,42E-01	3,71E-02
39	6,30E-02	2,51E-01	2,93E-04	7,88E-07	9,94E-01	2,15E-03	4,64E-02	2,09E-03	6,14E-06	5,21E-01	3,26E-02
40	6,49E-02	2,55E-01	2,95E-04	3,23E-07	9,95E-01	8,33E-03	9,13E-02	2,10E-03	5,97E-06	1,03E+00	3,66E-02
41	1,11E-01	3,34E-01	3,41E-04	4,89E-07	9,97E-01	8,51E-02	2,92E-01	2,17E-03	6,12E-06	1,68E-01	9,81E-02
42	6,51E-02	2,55E-01	2,95E-04	2,33E-07	9,47E-01	4,00E-03	6,32E-02	2,09E-03	6,14E-06	5,08E-01	3,45E-02

43	6,45E-02	2,54E-01	2,95E-04	4,93E-07	9,96E-01	3,88E-02	1,97E-01	2,13E-03	6,37E-06	9,79E-01	5,17E-02
44	9,45E-02	3,07E-01	3,25E-04	7,11E-07	9,90E-01	7,44E-02	2,73E-01	2,16E-03	6,37E-06	6,13E-01	8,45E-02
45	6,84E-02	2,62E-01	2,99E-04	4,95E-07	9,97E-01	2,75E-03	5,24E-02	2,09E-03	6,40E-06	5,63E-01	3,56E-02
46	5,97E-02	2,44E-01	2,90E-04	6,23E-07	9,76E-01	5,89E-03	7,67E-02	2,09E-03	5,95E-06	1,05E+00	3,28E-02
47	1,24E-01	3,52E-01	3,54E-04	6,82E-07	9,96E-01	4,54E-03	6,74E-02	2,09E-03	6,12E-06	9,09E-01	6,44E-02
48	1,25E-01	3,54E-01	3,55E-04	7,17E-07	9,95E-01	1,55E-03	3,94E-02	2,09E-03	5,98E-06	7,63E-01	6,34E-02
49	1,26E-01	3,54E-01	3,56E-04	6,62E-07	9,84E-01	8,33E-03	9,13E-02	2,10E-03	6,32E-06	1,60E-01	6,69E-02
50	8,56E-02	2,93E-01	3,16E-04	3,44E-07	9,10E-01	9,34E-03	9,66E-02	2,10E-03	6,14E-06	7,94E-01	4,75E-02

Tabla 7.11. Valores de error del segundo sistema neurodifuso con clustering para la detección de *Xanthomonas Campestris*
Fuente: Elaboración propia

Sistema 2 ANFIS con Clustering											
Ejecución	Entrenamiento					Validación					\overline{MSE}
	MSE	RMSE	MAE	MIN	MAX	MSE	RMSE	MAE	MIN	MAX	
1	1,11E-02	1,05E-01	3,20E-03	8,14E-05	2,58E-01	1,25E-02	1,12E-01	4,52E-03	4,51E-03	5,80E-01	1,18E-02
2	1,16E-02	1,08E-01	3,20E-03	1,24E-04	2,31E-01	6,73E-03	8,20E-02	4,52E-03	4,51E-03	6,26E-01	9,16E-03
3	9,47E-03	9,73E-02	3,20E-03	6,99E-05	2,45E-01	8,13E-03	9,02E-02	4,52E-03	4,51E-03	7,77E-01	8,80E-03
4	1,62E-02	1,27E-01	3,20E-03	7,06E-05	2,38E-01	1,30E-02	1,14E-01	4,52E-03	4,52E-03	2,99E-01	1,46E-02
5	1,32E-02	1,15E-01	3,20E-03	1,11E-04	2,84E-01	5,42E-03	7,36E-02	4,52E-03	4,51E-03	4,46E-01	9,29E-03
6	1,22E-02	1,11E-01	3,20E-03	1,24E-04	2,70E-01	1,33E-02	1,15E-01	4,52E-03	4,51E-03	3,82E-01	1,27E-02
7	1,07E-02	1,03E-01	3,20E-03	9,24E-05	2,47E-01	9,93E-03	9,97E-02	4,52E-03	4,51E-03	9,83E-01	1,03E-02
8	8,71E-03	9,33E-02	3,20E-03	8,03E-05	3,00E-01	6,24E-03	7,90E-02	4,52E-03	4,51E-03	3,96E-01	7,48E-03
9	8,62E-03	9,28E-02	3,20E-03	1,03E-04	2,81E-01	9,19E-03	9,59E-02	4,52E-03	4,52E-03	9,40E-01	8,91E-03
10	1,43E-02	1,20E-01	3,20E-03	1,02E-04	2,92E-01	1,07E-02	1,03E-01	4,52E-03	4,51E-03	3,74E-01	1,25E-02
11	1,02E-02	1,01E-01	3,20E-03	1,04E-04	2,95E-01	5,51E-03	7,42E-02	4,52E-03	4,51E-03	8,19E-01	7,87E-03
12	1,51E-02	1,23E-01	3,20E-03	1,01E-04	2,92E-01	1,59E-02	1,26E-01	4,53E-03	4,51E-03	7,29E-01	1,55E-02
13	1,65E-02	1,29E-01	3,20E-03	1,13E-04	2,55E-01	9,34E-03	9,67E-02	4,52E-03	4,51E-03	2,26E-01	1,29E-02

14	8,77E-03	9,37E-02	3,20E-03	8,49E-05	2,59E-01	1,48E-02	1,22E-01	4,52E-03	4,52E-03	2,36E-01	1,18E-02
15	1,58E-02	1,26E-01	3,20E-03	8,41E-05	2,94E-01	1,42E-02	1,19E-01	4,52E-03	4,51E-03	6,18E-01	1,50E-02
16	8,65E-03	9,30E-02	3,20E-03	8,95E-05	2,64E-01	6,67E-03	8,17E-02	4,52E-03	4,51E-03	6,64E-01	7,66E-03
17	1,52E-02	1,23E-01	3,20E-03	7,26E-05	2,90E-01	1,13E-02	1,06E-01	4,52E-03	4,51E-03	3,42E-01	1,33E-02
18	1,36E-02	1,17E-01	3,20E-03	1,16E-04	2,75E-01	8,72E-03	9,34E-02	4,52E-03	4,52E-03	5,96E-01	1,12E-02
19	7,52E-03	8,67E-02	3,20E-03	8,92E-05	2,94E-01	1,30E-02	1,14E-01	4,52E-03	4,51E-03	4,93E-01	1,02E-02
20	1,48E-02	1,22E-01	3,20E-03	1,17E-04	2,88E-01	1,02E-02	1,01E-01	4,52E-03	4,51E-03	7,54E-01	1,25E-02
21	6,83E-03	8,27E-02	3,19E-03	1,07E-04	2,92E-01	8,91E-03	9,44E-02	4,52E-03	4,51E-03	3,84E-01	7,87E-03
22	1,52E-02	1,23E-01	3,20E-03	7,77E-05	2,56E-01	7,02E-03	8,38E-02	4,52E-03	4,51E-03	5,38E-01	1,11E-02
23	7,09E-03	8,42E-02	3,19E-03	7,49E-05	2,40E-01	1,21E-02	1,10E-01	4,52E-03	4,51E-03	3,65E-01	9,57E-03
24	8,67E-03	9,31E-02	3,20E-03	1,07E-04	1,51E-01	1,18E-02	1,09E-01	4,52E-03	4,51E-03	2,29E-01	1,02E-02
25	1,29E-02	1,14E-01	3,20E-03	8,64E-05	3,02E-01	1,40E-02	1,18E-01	4,52E-03	4,51E-03	6,91E-01	1,34E-02
26	1,27E-02	1,13E-01	3,20E-03	1,02E-04	2,27E-01	5,57E-03	7,46E-02	4,52E-03	4,51E-03	3,34E-01	9,14E-03
27	1,64E-02	1,28E-01	3,20E-03	1,18E-04	1,02E-01	1,23E-02	1,11E-01	4,52E-03	4,51E-03	9,49E-01	1,43E-02
28	1,46E-02	1,21E-01	3,20E-03	8,94E-05	1,15E-01	1,14E-02	1,07E-01	4,52E-03	4,51E-03	2,96E-01	1,30E-02
29	7,21E-03	8,49E-02	3,19E-03	6,51E-05	2,24E-01	6,18E-03	7,86E-02	4,51E-03	4,51E-03	2,18E-01	5,38E-03
30	1,37E-02	1,17E-01	3,20E-03	8,67E-05	2,28E-01	1,32E-02	1,15E-01	4,52E-03	4,51E-03	6,88E-01	1,35E-02
31	1,61E-02	1,27E-01	3,20E-03	1,14E-04	2,29E-01	1,37E-02	1,17E-01	4,52E-03	4,51E-03	2,86E-01	1,49E-02
32	6,91E-03	8,31E-02	3,19E-03	1,25E-04	2,28E-01	1,56E-02	1,25E-01	4,53E-03	4,51E-03	8,12E-01	1,13E-02
33	1,25E-02	1,12E-01	3,20E-03	8,01E-05	1,01E-01	1,35E-02	1,16E-01	4,52E-03	4,52E-03	3,86E-01	1,30E-02
34	7,92E-03	8,90E-02	3,20E-03	1,03E-04	2,28E-01	1,52E-02	1,23E-01	4,53E-03	4,51E-03	7,43E-01	1,15E-02
35	1,42E-02	1,19E-01	3,20E-03	1,18E-04	2,28E-01	6,14E-03	7,84E-02	4,52E-03	4,51E-03	3,87E-01	1,02E-02
36	1,61E-02	1,27E-01	3,20E-03	1,10E-04	2,25E-01	1,18E-02	1,09E-01	4,52E-03	4,52E-03	3,93E-01	1,39E-02
37	1,06E-02	1,03E-01	3,20E-03	1,03E-04	8,89E-02	1,50E-02	1,22E-01	4,53E-03	4,51E-03	8,98E-01	1,28E-02
38	1,03E-02	1,01E-01	3,20E-03	7,02E-05	2,26E-01	1,04E-02	1,02E-01	4,52E-03	4,52E-03	4,82E-01	1,03E-02
39	1,72E-02	1,31E-01	3,20E-03	6,64E-05	2,25E-01	1,55E-02	1,25E-01	4,53E-03	4,51E-03	6,54E-01	1,64E-02
40	1,01E-02	1,01E-01	3,20E-03	9,90E-05	2,31E-01	1,06E-02	1,03E-01	4,52E-03	4,52E-03	1,02E+00	1,04E-02
41	1,58E-02	1,26E-01	3,20E-03	1,17E-04	1,62E-01	1,31E-02	1,14E-01	4,52E-03	4,51E-03	3,72E-01	1,44E-02

42	7,18E-03	8,47E-02	3,19E-03	1,01E-04	2,25E-01	1,26E-02	1,12E-01	4,52E-03	4,51E-03	8,24E-01	9,88E-03
43	8,66E-03	9,31E-02	3,20E-03	9,78E-05	2,25E-01	1,43E-02	1,20E-01	4,52E-03	4,51E-03	8,83E-01	1,15E-02
44	1,30E-02	1,14E-01	3,20E-03	1,11E-04	2,29E-01	1,57E-02	1,25E-01	4,53E-03	4,52E-03	3,87E-02	1,44E-02
45	1,24E-02	1,11E-01	3,20E-03	8,39E-05	6,65E-02	1,49E-02	1,22E-01	4,52E-03	4,51E-03	3,28E-01	1,36E-02
46	1,45E-02	1,20E-01	3,20E-03	7,29E-05	2,30E-01	5,86E-03	7,65E-02	4,52E-03	4,51E-03	4,48E-01	1,02E-02
47	1,13E-02	1,06E-01	3,20E-03	7,05E-05	2,26E-01	6,24E-03	7,90E-02	4,52E-03	4,51E-03	9,52E-01	8,76E-03
48	1,49E-02	1,22E-01	3,20E-03	1,14E-04	1,54E-01	6,64E-03	8,15E-02	4,52E-03	4,51E-03	5,85E-01	1,08E-02
49	6,93E-03	8,32E-02	3,19E-03	8,68E-05	2,31E-01	7,03E-03	8,39E-02	4,52E-03	4,51E-03	7,69E-01	6,98E-03
50	6,44E-03	8,03E-02	3,19E-03	7,18E-05	2,27E-01	7,12E-03	8,44E-02	4,52E-03	4,52E-03	2,78E-01	6,78E-03

Tabla 7.12. Valores de error del segundo sistema neurodifuso sin clustering para la detección de *Xanthomonas Campestris*
Fuente: Elaboración propia

Sistema 2 ANFIS sin Clustering											
Ejecución	Entrenamiento					Validación					\overline{MSE}
	MSE	RMSE	MAE	MIN	MAX	MSE	RMSE	MAE	MIN	MAX	
1	1,03E-02	1,01E-01	3,21E-03	3,06E-05	2,31E-01	1,27E-02	1,13E-01	3,41E-03	1,65E-04	9,02E-01	1,15E-02
2	1,20E-02	1,09E-01	3,21E-03	3,98E-05	2,47E-01	6,48E-03	8,05E-02	3,41E-03	1,61E-04	8,15E-01	9,22E-03
3	1,02E-02	1,01E-01	3,21E-03	2,57E-05	2,42E-01	1,45E-02	1,20E-01	3,42E-03	1,59E-04	3,04E-01	1,23E-02
4	8,99E-03	9,48E-02	3,20E-03	6,00E-05	2,67E-01	9,73E-03	9,87E-02	3,41E-03	1,62E-04	4,64E-01	9,36E-03
5	1,03E-02	1,02E-01	3,21E-03	2,98E-05	2,50E-01	1,24E-02	1,11E-01	3,41E-03	1,62E-04	7,60E-01	1,13E-02
6	1,20E-02	1,09E-01	3,21E-03	6,10E-05	2,30E-01	8,83E-03	9,40E-02	3,41E-03	1,64E-04	9,86E-01	1,04E-02
7	1,30E-02	1,14E-01	3,21E-03	1,35E-05	3,00E-01	9,15E-03	9,57E-02	3,41E-03	1,64E-04	7,36E-01	1,11E-02
8	8,25E-03	9,08E-02	3,20E-03	2,29E-06	2,59E-01	1,16E-02	1,08E-01	3,41E-03	1,59E-04	7,13E-01	9,90E-03
9	7,30E-03	8,55E-02	3,20E-03	3,53E-05	2,69E-01	9,16E-03	9,57E-02	3,41E-03	1,64E-04	7,98E-01	8,23E-03
10	6,56E-03	8,10E-02	3,20E-03	5,56E-05	2,43E-01	8,03E-03	8,96E-02	3,41E-03	1,62E-04	7,73E-01	7,30E-03
11	8,10E-03	9,00E-02	3,20E-03	2,19E-05	2,47E-01	1,29E-02	1,14E-01	3,41E-03	1,61E-04	9,22E-01	1,05E-02
12	6,54E-03	8,09E-02	3,20E-03	5,45E-05	2,79E-01	1,30E-02	1,14E-01	3,41E-03	1,64E-04	3,47E-01	9,77E-03

13	1,65E-02	1,29E-01	3,21E-03	1,02E-05	5,99E-02	5,94E-03	7,71E-02	3,41E-03	1,64E-04	8,51E-01	1,12E-02
14	8,64E-03	9,29E-02	3,20E-03	3,07E-05	2,67E-01	1,24E-02	1,12E-01	3,41E-03	1,62E-04	3,49E-01	1,05E-02
15	7,38E-03	8,59E-02	3,20E-03	2,07E-05	2,94E-01	5,06E-03	7,12E-02	3,41E-03	1,62E-04	3,66E-01	6,22E-03
16	9,28E-03	9,63E-02	3,21E-03	4,89E-05	1,13E-01	5,77E-03	7,60E-02	3,41E-03	1,60E-04	5,56E-01	7,53E-03
17	1,49E-02	1,22E-01	3,21E-03	6,31E-06	7,11E-02	9,54E-03	9,77E-02	3,41E-03	1,61E-04	4,17E-01	1,22E-02
18	7,25E-03	8,52E-02	3,20E-03	1,96E-06	2,24E-01	5,14E-03	7,17E-02	3,40E-03	1,59E-04	2,18E-01	6,19E-03
19	1,08E-02	1,04E-01	3,21E-03	4,33E-05	2,31E-01	1,08E-02	1,04E-01	3,41E-03	1,63E-04	4,84E-01	1,08E-02
20	9,62E-03	9,81E-02	3,21E-03	5,84E-05	2,28E-01	1,08E-02	1,04E-01	3,41E-03	1,61E-04	2,84E-01	1,02E-02
21	7,07E-03	8,41E-02	3,20E-03	1,90E-05	2,26E-01	8,62E-03	9,28E-02	3,41E-03	1,60E-04	6,41E-01	7,84E-03
22	1,51E-02	1,23E-01	3,21E-03	4,32E-05	2,01E-01	1,40E-02	1,18E-01	3,41E-03	1,61E-04	8,73E-01	1,45E-02
23	1,52E-02	1,23E-01	3,21E-03	8,86E-06	2,32E-01	6,63E-03	8,14E-02	3,41E-03	1,63E-04	6,77E-01	1,09E-02
24	1,46E-02	1,21E-01	3,21E-03	4,49E-05	2,25E-01	4,61E-03	6,79E-02	3,41E-03	1,61E-04	2,91E-01	9,58E-03
25	1,57E-02	1,25E-01	3,21E-03	5,12E-05	2,28E-01	1,41E-02	1,19E-01	3,41E-03	1,61E-04	2,25E-01	1,49E-02
26	1,56E-02	1,25E-01	3,21E-03	2,04E-05	6,63E-02	9,79E-03	9,89E-02	3,41E-03	1,59E-04	7,62E-01	1,27E-02
27	7,94E-03	8,91E-02	3,20E-03	1,97E-05	2,28E-01	8,91E-03	9,44E-02	3,41E-03	1,62E-04	9,97E-01	8,43E-03
28	1,61E-02	1,27E-01	3,21E-03	3,16E-05	2,25E-01	8,13E-03	9,02E-02	3,41E-03	1,59E-04	3,92E-01	1,21E-02
29	1,47E-02	1,21E-01	3,21E-03	5,05E-05	2,29E-01	8,62E-03	9,29E-02	3,41E-03	1,59E-04	3,45E-01	1,17E-02
30	1,21E-02	1,10E-01	3,21E-03	3,16E-05	2,16E-01	1,05E-02	1,02E-01	3,41E-03	1,60E-04	9,34E-01	1,13E-02
31	6,88E-03	8,30E-02	3,20E-03	3,47E-06	2,28E-01	9,52E-03	9,76E-02	3,41E-03	1,63E-04	1,01E+00	8,20E-03
32	1,69E-02	1,30E-01	3,21E-03	1,30E-05	2,30E-01	8,99E-03	9,48E-02	3,41E-03	1,62E-04	5,64E-01	1,30E-02
33	9,37E-03	9,68E-02	3,21E-03	3,92E-05	2,25E-01	1,36E-02	1,16E-01	3,41E-03	1,60E-04	9,62E-02	1,15E-02
34	1,38E-02	1,18E-01	3,21E-03	2,06E-05	3,21E-02	1,38E-02	1,18E-01	3,41E-03	1,62E-04	2,31E-01	1,38E-02
35	8,16E-03	9,03E-02	3,20E-03	5,38E-05	2,31E-01	4,28E-03	6,54E-02	3,41E-03	1,62E-04	7,59E-01	6,22E-03
36	1,45E-02	1,20E-01	3,21E-03	3,60E-05	2,29E-01	9,06E-03	9,52E-02	3,41E-03	1,61E-04	3,68E-01	1,18E-02
37	1,38E-02	1,18E-01	3,21E-03	3,76E-06	5,85E-02	8,69E-03	9,32E-02	3,41E-03	1,65E-04	7,25E-01	1,13E-02
38	7,16E-03	8,46E-02	3,20E-03	1,79E-05	2,28E-01	1,14E-02	1,07E-01	3,41E-03	1,62E-04	8,92E-01	9,28E-03
39	7,14E-03	8,45E-02	3,20E-03	2,05E-05	2,29E-01	8,35E-03	9,14E-02	3,41E-03	1,59E-04	3,55E-01	7,74E-03
40	1,34E-02	1,16E-01	3,21E-03	5,94E-06	2,25E-01	5,97E-03	7,73E-02	3,41E-03	1,61E-04	2,67E-01	9,71E-03

41	1,29E-02	1,14E-01	3,21E-03	4,41E-05	1,86E-01	4,37E-03	6,61E-02	3,41E-03	1,62E-04	2,34E-01	8,65E-03
42	6,30E-03	7,93E-02	3,20E-03	2,38E-05	2,25E-01	9,79E-03	9,89E-02	3,41E-03	1,64E-04	9,47E-01	8,04E-03
43	1,01E-02	1,01E-01	3,21E-03	2,31E-05	2,31E-01	1,21E-02	1,10E-01	3,41E-03	1,59E-04	2,46E-01	1,11E-02
44	7,50E-03	8,66E-02	3,20E-03	1,42E-05	2,30E-01	4,38E-03	6,62E-02	3,41E-03	1,65E-04	3,03E-01	5,94E-03
45	6,78E-03	8,23E-02	3,20E-03	3,73E-05	1,96E-01	8,15E-03	9,03E-02	3,41E-03	1,64E-04	3,16E-01	7,46E-03
46	1,08E-02	1,04E-01	3,21E-03	5,64E-05	2,26E-01	9,06E-03	9,52E-02	3,41E-03	1,59E-04	1,01E+00	9,95E-03
47	1,10E-02	1,05E-01	3,21E-03	3,97E-05	2,28E-01	4,61E-03	6,79E-02	3,41E-03	1,63E-04	9,72E-01	7,83E-03
48	1,60E-02	1,26E-01	3,21E-03	3,85E-05	1,49E-01	1,32E-02	1,15E-01	3,41E-03	1,63E-04	9,62E-01	1,46E-02
49	6,66E-03	8,16E-02	3,20E-03	3,58E-05	2,26E-01	4,85E-03	6,96E-02	3,41E-03	1,61E-04	5,15E-01	5,75E-03
50	1,68E-02	1,30E-01	3,21E-03	1,32E-05	2,24E-01	8,48E-03	9,21E-02	3,41E-03	1,63E-04	7,83E-01	1,27E-02

7.4 RESULTADOS

Como fase final del desarrollo de la investigación, en esta sección se describen los resultados obtenidos al ejecutar los sistemas neurodifusos con o sin clustering, para la detección de la enfermedad causada por *Xanthomonas Campestris*, a partir de las imágenes procesadas en modelo de color HSB de 15 cultivos diferentes entre sanas o infectadas en una equivalencia de 70% para entrenamiento y 30% de prueba.

Cada sistema es ejecutado en 50 oportunidades y se obtienen los valores estadísticos máximo, mínimo, error cuadrático medio y error absoluto medio, para las fases de entrenamiento y prueba como se evidencia en la sección 7.3.2.

Posteriormente, para determinar el mejor sistema entre el desarrollado con clustering o sin clustering se realiza una prueba t y una prueba de Kruskal-Wallis, para así establecer una comparación del modelo propuesto respecto a algunos modelos similares que se encuentran en literatura relacionada.

7.4.1 Entrenamiento

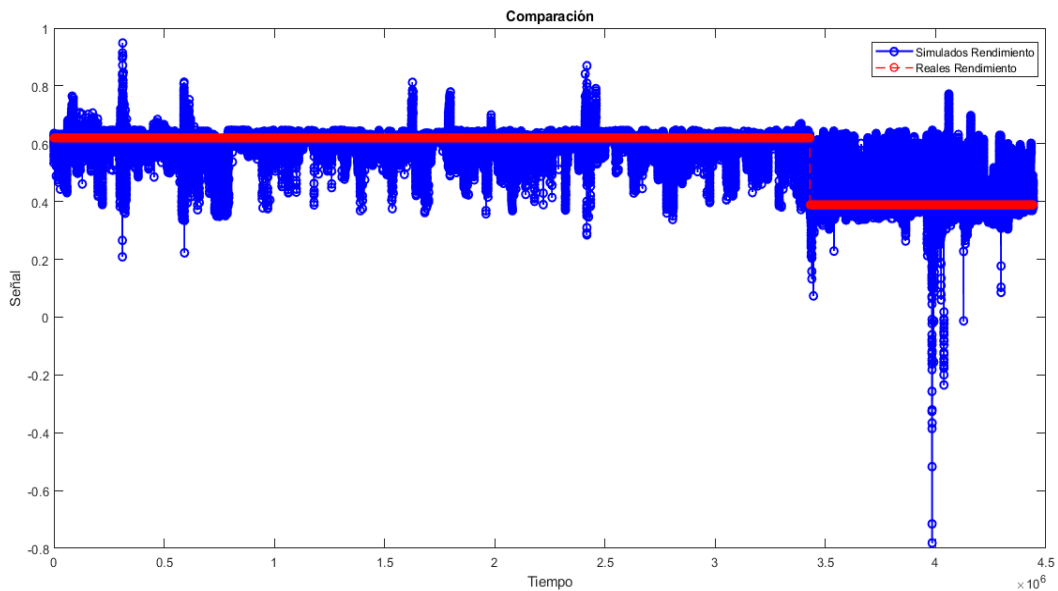
En el entrenamiento se seleccionaron las imágenes que presentaban patrones de la enfermedad mejor reconocibles para evitar redundancia de información o confusiones que pudiesen afectar el diagnóstico resultante. Posteriormente, se llevó a cabo la ejecución de los sistemas en 50 oportunidades obteniendo los resultados resumidos en la tabla 7.13, resumiendo los valores estadísticos y la ejecución en la que fue posible alcanzar el mejor resultado.

*Tabla 7.13. Resultados de fase de entrenamiento
Fuente: Elaboración propia*

Sistema de Inferencia Difusa	Ejecución	Entrenamiento					
		MSE	RMSE	MAE	MIN	MAX	Precisión
Sistema 1 ANFIS con Clustering	6	1,62E-03	4,02E-02	1,91E-05	2,11E-08	9,70E-01	95,98%
Sistema 1 ANFIS sin Clustering	16	5,87E-02	2,42E-01	2,30E-04	2,11E-07	9,90E-01	75,77%
Sistema 2 ANFIS con Clustering	29	7,21E-03	8,49E-02	3,19E-03	6,51E-05	2,24E-01	91,51%
Sistema 2 ANFIS sin Clustering	18	7,25E-03	8,52E-02	3,20E-03	1,96E-06	2,24E-01	91,48%

Seguidamente, en la figura 7.7 se presenta la comparación entre los resultados obtenidos y los resultados esperados del entrenamiento del Sistema 1, en ella se observa que los datos esperados para un pixel sano son cercanos al valor de 0.62, mientras que, el comportamiento

resultante de la comparación de los valores simulados y esperados, para pixeles que representen el estado de enfermo (0.3884), tiene un comportamiento similar.



*Figura 7.7. Comparación de valores reales y simulados de en fase de entrenamiento del primer sistema neurodifuso.
Fuente: Elaboración propia. (Entorno MATLAB®).*

Teniendo presente que como métrica establecida para determinar la mejor configuración de los diferentes sistemas es la relacionada con el error cuadrático medio, en la figura 7.8 se evidencia esta métrica para la mejor configuración del sistema 1, el cual presenta un valor de $1,62E-03$, producto de diferencias sustanciales entre los resultados de la simulación.

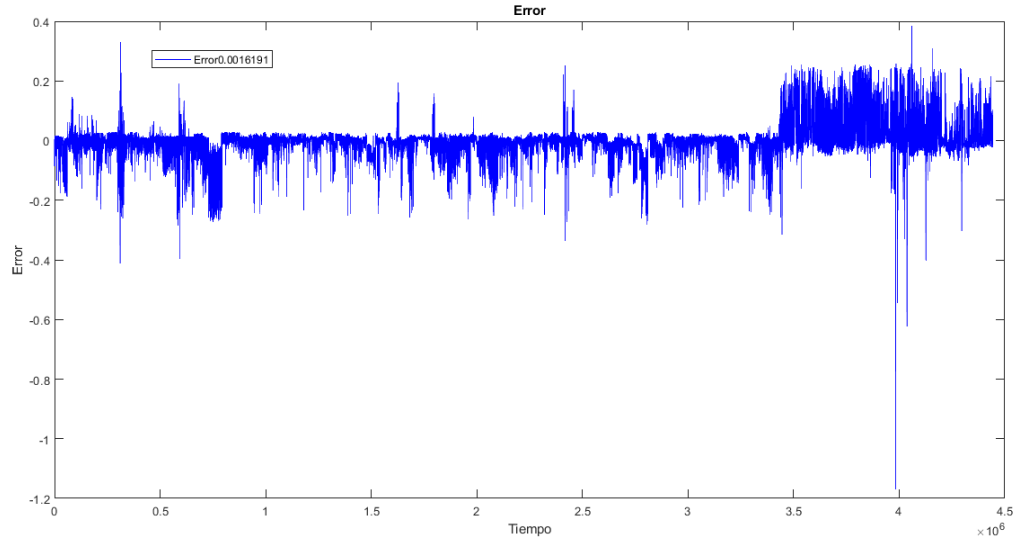


Figura 7.8. Error de valores reales del primer sistema neurodifuso en fase de entrenamiento
Fuente: Elaboración propia. (Entorno MATLAB®).

Para el sistema 2, el cual tiene como entrada el promedio de los valores obtenidos del sistema 1 y se obtiene como resultado valores sobre el estado de la planta, se muestra en la figura 7.9, la comparación entre los resultados obtenidos y los resultados esperados, teniendo en cuenta que los valores para un pixel sano son aproximados al valor de 0.62, mientras que, el comportamiento resultante de la comparación de los valores simulados y esperados, para pixeles que representen el estado de enfermo (0.3884).

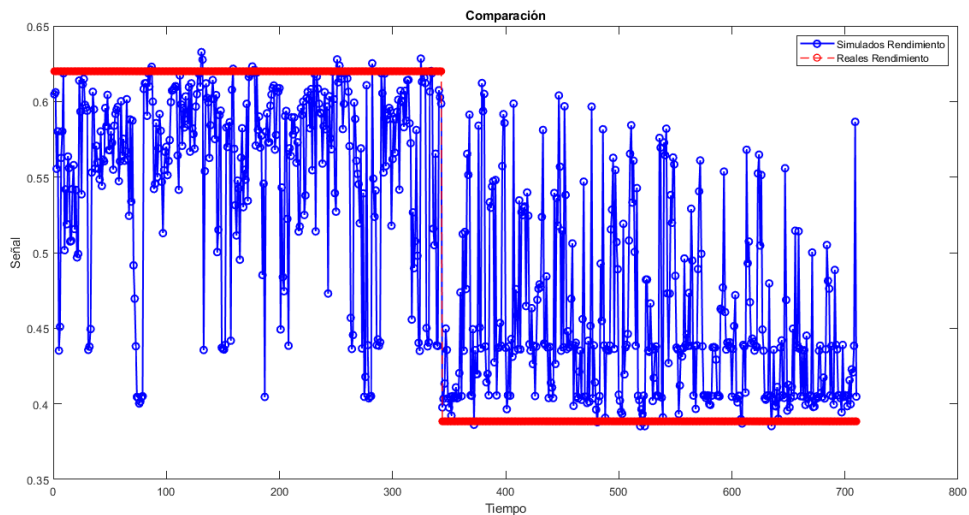


Figura 7.9. Comparación de valores reales y simulados de en fase de entrenamiento del sistema neurodifuso.

Fuente: Elaboración propia. (Entorno MATLAB®).

Para verificar visualmente la métrica establecida para determinar la mejor configuración, se presenta la figura 7.10, que contiene el error, en la cual se obtiene un MSE de 7,21E-03.

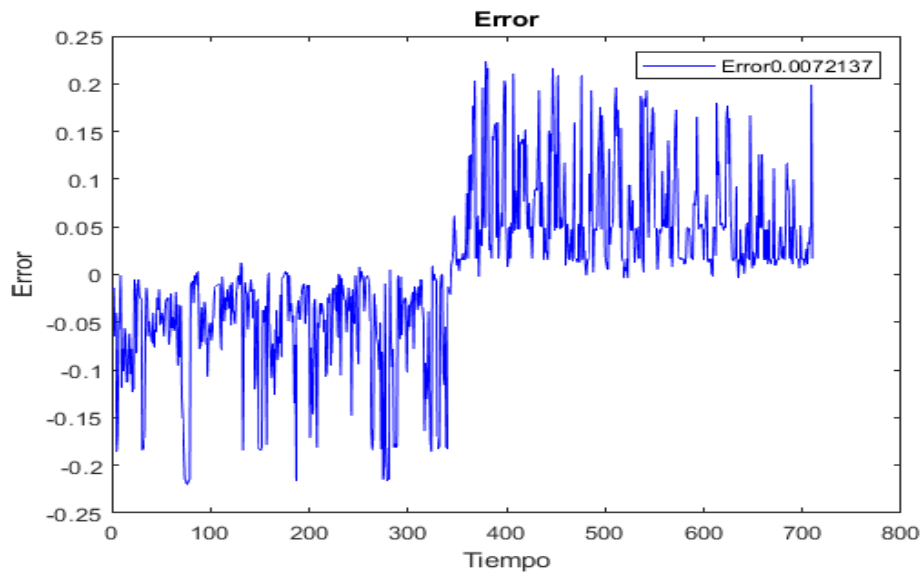


Figura 7.10. Error de valores reales del primer sistema neurodifuso en fase de entrenamiento
Fuente: Elaboración propia. (Entorno MATLAB®).

7.4.2 Prueba

Para la etapa de prueba se eligieron imágenes al azar entre sanas e infectadas por *Xanthomonas Campestris*, con las cuales el modelo no hubiese sido entrenado para así verificar con más precisión la capacidad del modelo para reconocer patrones que no hubiese identificado anteriormente. Al igual que en la fase de entrenamiento, se ejecuta el algoritmo en 50 oportunidades, obteniendo así el resumen de datos estadísticos de la tabla 7.14 y una aproximación a la precisión real de los sistemas.

Tabla 7.14. Resultados fase de prueba
Fuente: Elaboración propia

Sistema de Inferencia Difusa	Ejecución	Pruebas					
		MSE	RMSE	MAE	MIN	MAX	Precisión
Sistema 1 ANFIS con Clustering	6	8,30E-04	2,88E-02	2,09E-05	5,94E-08	2,77E-01	97,12%
Sistema 1 ANFIS sin Clustering	16	9,67E-04	3,11E-02	2,09E-03	5,94E-06	2,77E-01	96,89%
Sistema 2 ANFIS con Clustering	29	6,18E-03	7,86E-02	4,51E-03	4,51E-03	2,18E-01	92,14%
Sistema 2 ANFIS sin Clustering	18	7,25E-03	8,52E-02	3,40E-03	1,59E-04	2,18E-01	91,86%

Por otro lado, para el caso de pruebas, también se realiza la comparación entre los resultados obtenidos y los resultados esperados de los sistemas.

En primer lugar, de manera similar al entrenamiento, para el primer sistema se obtiene que los datos esperados para un pixel sano son cercanos al valor de 0.62, mientras que, el comportamiento resultante de la comparación de los valores simulados y esperados, para pixeles que representen el estado de enfermo (0.3884) como se muestra en la figura 7.11.

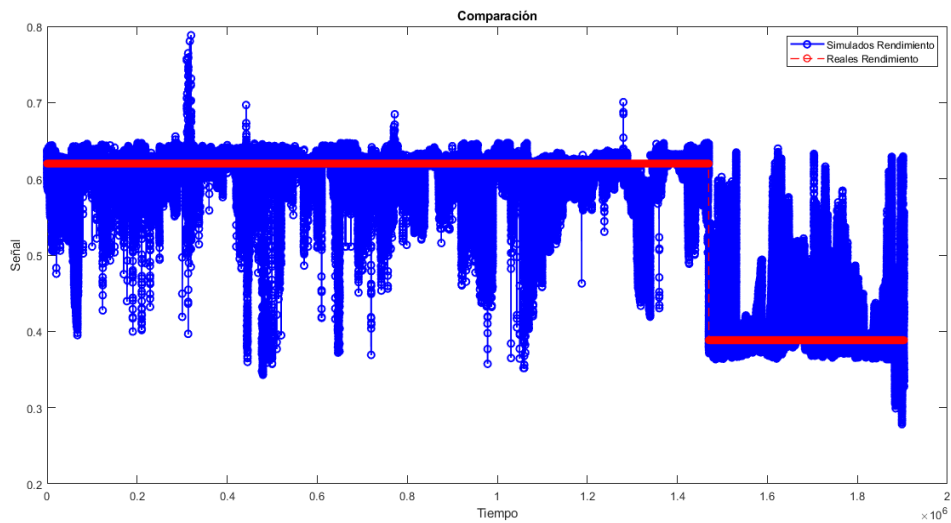


Figura 7.11. Comparación de valores reales y simulados de en fase de prueba del primer sistema neurodifuso. (Entorno MATLAB®).

Fuente: Elaboración propia

De igual manera, se presenta la gráfica del Error el cual presenta un valor de $8,30E-04$, como se evidencia en la figura 7.12, lo cual representa que, en la prueba se obtiene un mejor rendimiento que a nivel de entrenamiento.

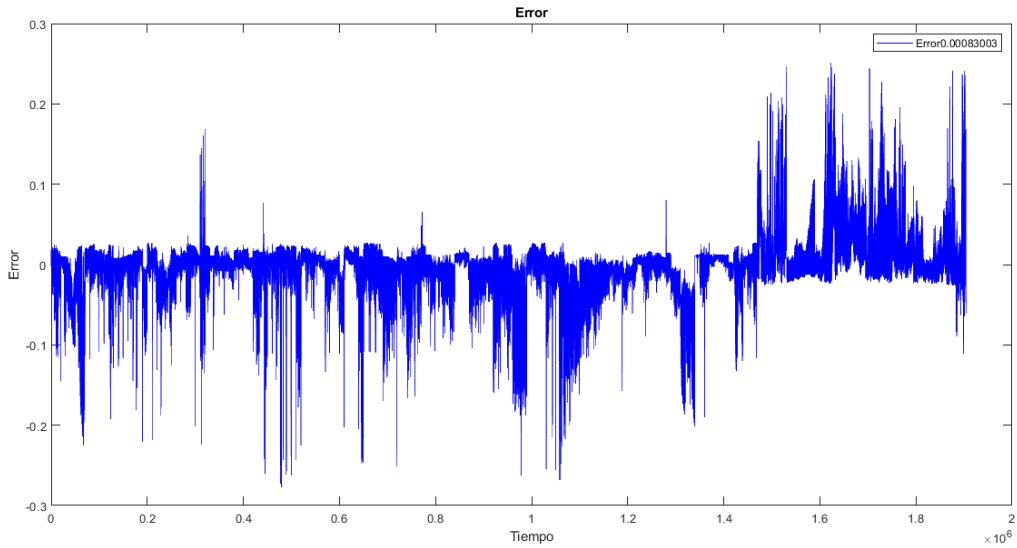


Figura 7.12. Error de valores reales del primer sistema neurodifuso en fase de prueba
Fuente: Elaboración propia. (Entorno MATLAB®).

Finalmente, para observar el comportamiento de la prueba en el sistema 2, en la figura 7.13. se encuentra nuevamente la comparación entre los resultados obtenidos y los resultados esperados, teniendo un comportamiento similar al obtenido en el entrenamiento, lo cual permite evidenciar buenos resultados en la detección de la enfermedad.

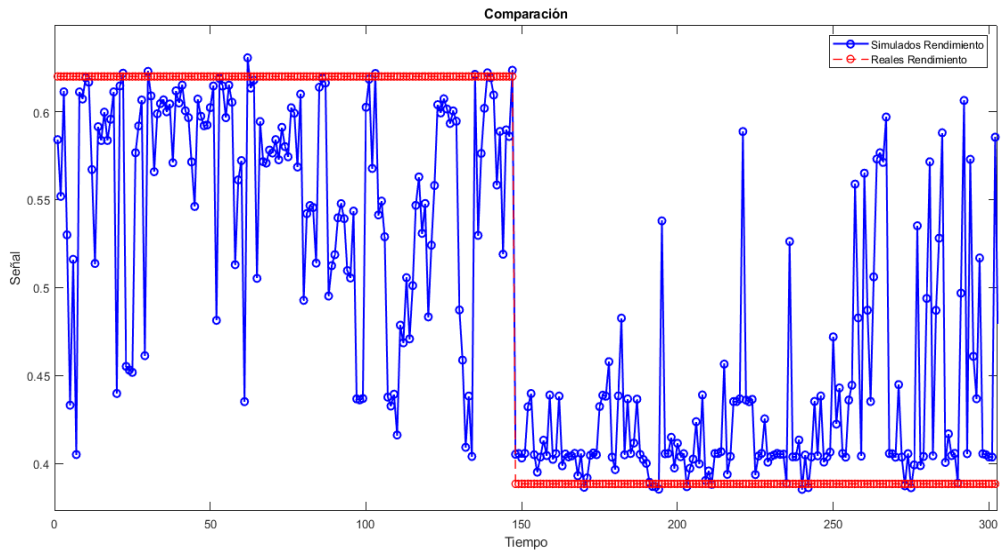


Figura 7.13. Comparación de valores reales y simulados de en fase de prueba del segundo sistema neurodifuso.
Fuente: Elaboración propia. (Entorno MATLAB®).

Para verificar visualmente la métrica establecida, y así determinar la mejor configuración, se presenta la gráfica del error que se muestra en la figura 7.14, en la cual se obtiene para las pruebas, un MSE de 6,18E-03.

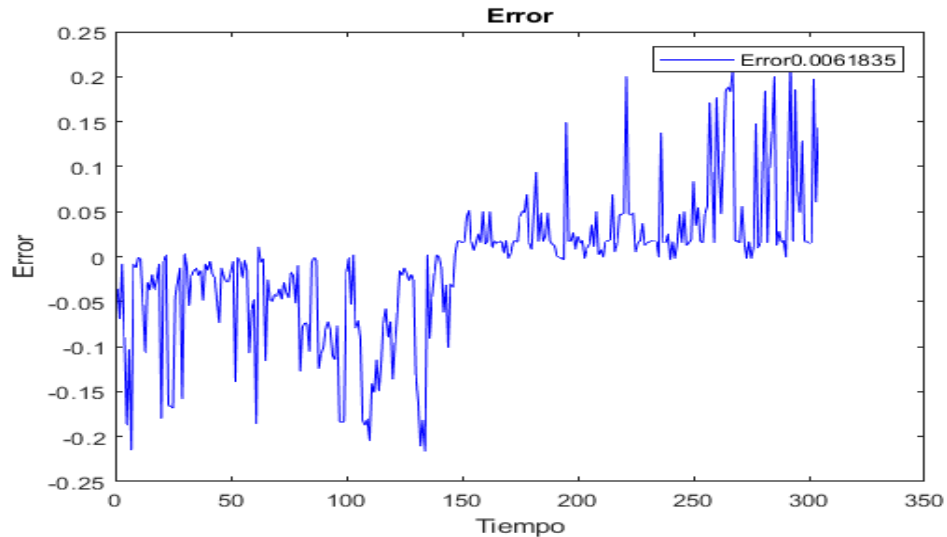


Figura 7.14. Error de valores reales del segundo sistema neurodifuso en fase de prueba.
Fuente: Elaboración propia. (Entorno MATLAB®).

7.4.3 Validación del mejor modelo

Con base en los resultados obtenidos en la sección 8.4.1 y 8.4.2, es posible definir como el mejor modelo al segundo sistema neurodifuso con clustering. Sin embargo, para verificar dicha afirmación se desarrollaron dos pruebas, la primera es la prueba de Kruskal-Wallis que compara las medianas de los grupos de datos para determinar si las muestras provienen de la misma población, comprobando si las diferentes muestras están equidistribuidas, es decir, una vez aplicados los algoritmos de optimización esta prueba permite comprobar si los resultados obtenidos por los modelos con respecto a los datos de entrenamiento y a su valor esperado, pertenecen a una misma población.

Con ayuda de la función *kruskalwallis* de Matlab se realizó dicha prueba con un intervalo de confianza de 0,05 para el segundo modelo neurodifuso con clustering, obteniendo los datos descritos en la figura 7.15 con un p-value de 0,646, el cual al ser mayor que el intervalo de confianza indica que la muestra está equidistribuida y por lo tanto si corresponde a la misma distribución, de modo que no es posible rechazar la hipótesis nula de la prueba.

Kruskal-Wallis ANOVA Table					
Source	SS	df	MS	Chi-sq	Prob>Chi-sq
Columns	6305.53	1	6305.5	0.21	0.646
Error	18136110.47	606	29927.6		
Total	18142416	607			

Figura 7.15. Resultados de prueba Kruskal-Wallis para segundo sistema con clustering

Fuente: Elaboración propia. (Entorno MATLAB®).

Por otra parte, para determinar el grado de significancia del resultado dado se realiza una prueba t a partir de la predicción de si los datos provienen de la misma distribución con la misma media y varianza apoyando lo resultante de la prueba Kruskal-Wallis, para ello se plantea la hipótesis nula de que la distribución de la muestra es normal con medias y varianzas iguales, se tiene en cuenta un intervalo de confianza de 0,05 y el p-value resultante es de 0.0668 cuyo valor es cercano a lo dado en la prueba Kruskal-Wallis y al ser mayor que el intervalo de confianza también indica que la muestra está equidistribuida y por lo tanto si corresponde a la misma distribución, de modo que no es posible rechazar la hipótesis nula de la prueba.

7.4.4 Composición del modelo

Para determinar la estructura interna del modelo compuesto por los sistemas que obtuvieron mejor resultado en la fase de pruebas, se procede en la presente subsección a presentar la conformación final de cada uno de los sistemas pertenecientes al modelo final.

Sistema 1 con clustering

Teniendo en cuenta la configuración establecida para los diferentes tipos de sistema a elaborar, en la figura 7.16 se observa el sistema de inferencia de lógica neurodifusa, el cual cuenta con 3 conjuntos de entrada correspondientes a los diferentes valores a nivel de Pixel en escala HSB, cada uno de los conjuntos con 3 funciones de pertenencia según la configuración establecida en los clúster, y un conjunto de salida que permite establecer el estado que representa el pixel (Sano o enfermo) y la estructura propia del sistema.

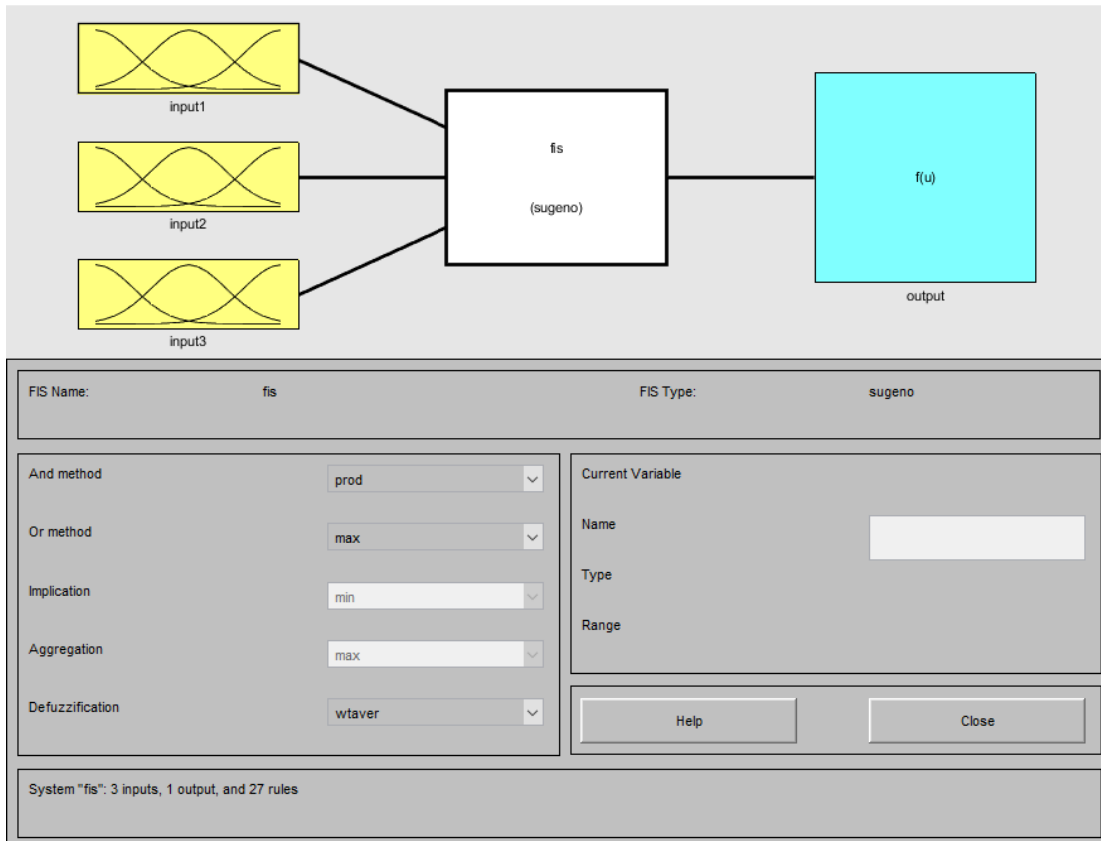


Figura 7.16. Configuración general primer sistema neurodifuso con clustering.
Fuente: Elaboración propia. (Entorno MATLAB®).

A su vez, el conjunto de reglas y su respectiva correspondencia a nivel de las funciones de pertenencia de la configuración del sistema obtenida en su mejor configuración se puede observar en la figura 7.17, en donde se observan los resultados a nivel de la función lineal (Salida) a partir de los diferentes valores que pueden tomar las funciones de pertenencia de las entradas (valores HSB).

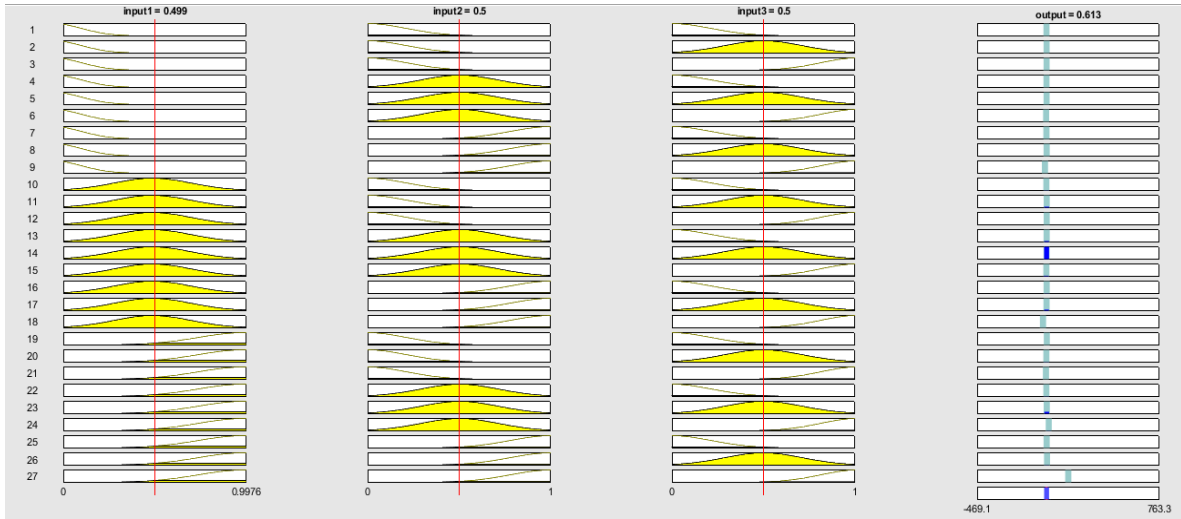


Figura 7.17. Visualización del conjunto de reglas del primer sistema de inferencia neurodifuso con clustering
Fuente: Elaboración propia. (Entorno MATLAB®).

Como producto de las diferentes correspondencias de las funciones de pertenencia se pueden obtener la superficie de reglas, el cual, como su nombre lo indica permite reflejar las reglas a través del espacio muestral para las variables de entrada y salida en 3 dimensiones como se muestra en la figura 7.18.

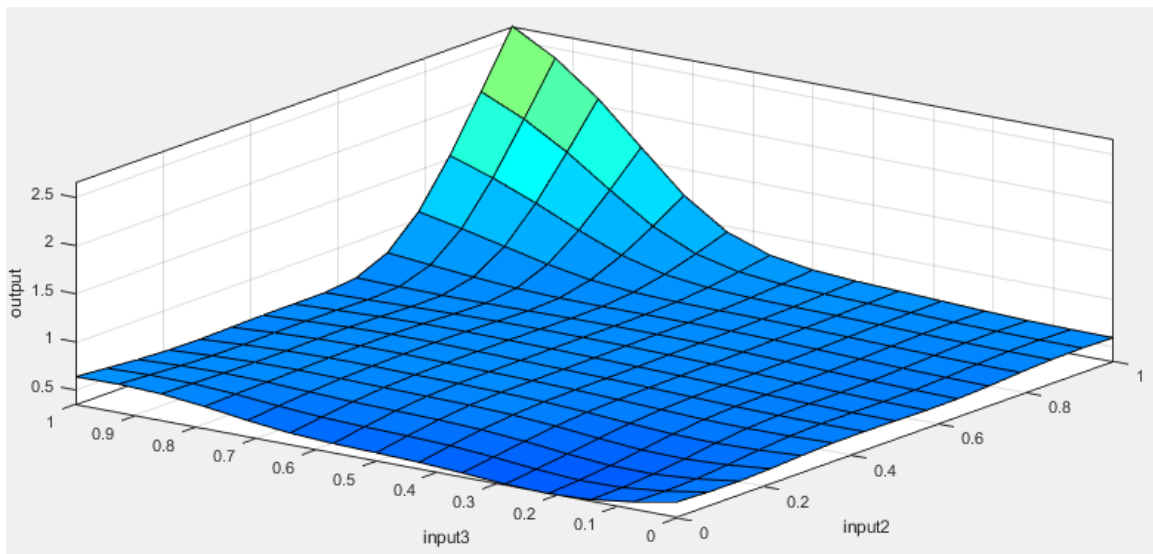
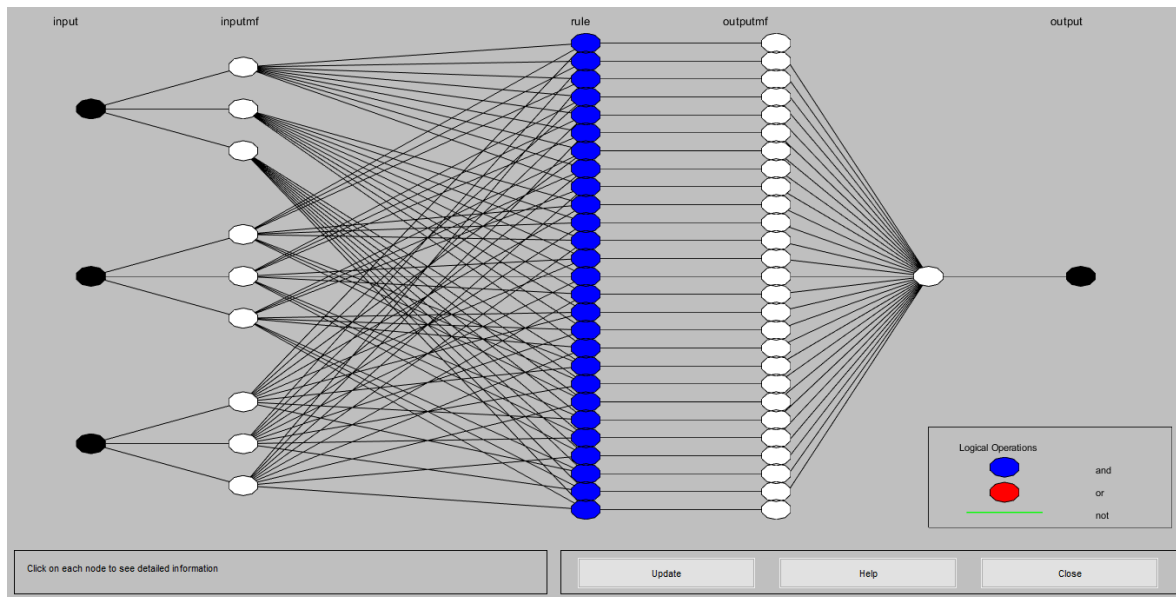


Figura 7.18. Superficie de salida del primer sistema neurodifuso con clustering.
Fuente: Elaboración propia. (Entorno MATLAB®).

La estructura final del primer sistema de inferencia neurodifuso se puede observar en la figura 7.19, en la cual se aprecian los 3 conjuntos de entrada, cada uno con 3 funciones de pertenencia,

junto con las 27 reglas establecidas mediante el clustering para el presente modelo, de manera tal que se pueda obtener una salida en términos de la validación del estado del píxel.



*Figura 7.19. Estructura ANFIS para primer sistema con clustering.
Fuente: Elaboración propia. (Entorno MATLAB®).*

Sistema 2 con clustering

Por otro lado, para el segundo sistema del modelo, se presenta el sistema de inferencia de lógica neurodifusa, que cuenta con 1 conjunto de entrada correspondiente al promedio de los valores obtenidos de una imagen evaluada por el sistema uno, y una única salida, como se muestra en la figura 7.20.

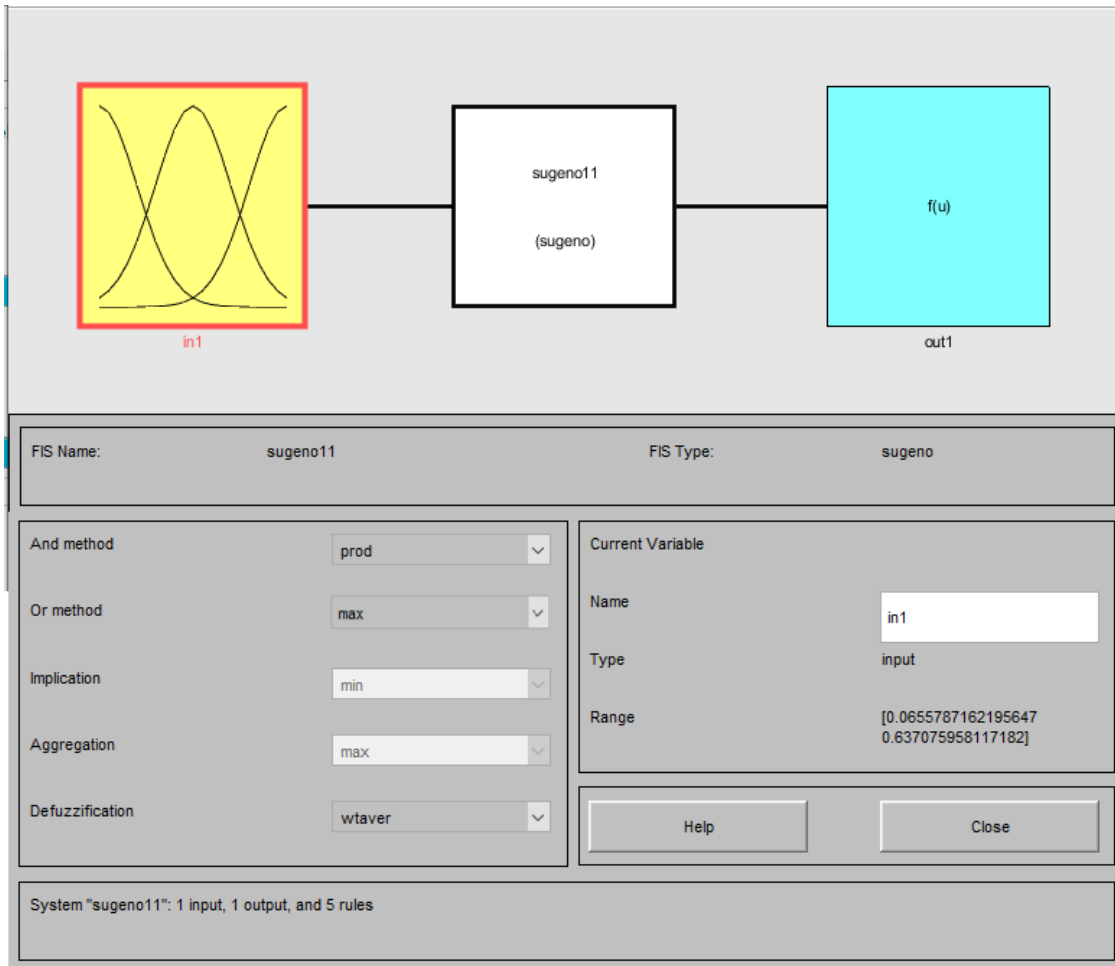


Figura 7.20. Configuración general del segundo sistema neurodifuso con clustering
Fuente: Elaboración propia. (Entorno MATLAB®).

Para este sistema, la correspondencia a nivel de las funciones de pertenencia con las diferentes reglas establecidas del sistema 2, se puede observar en la figura 7.21.

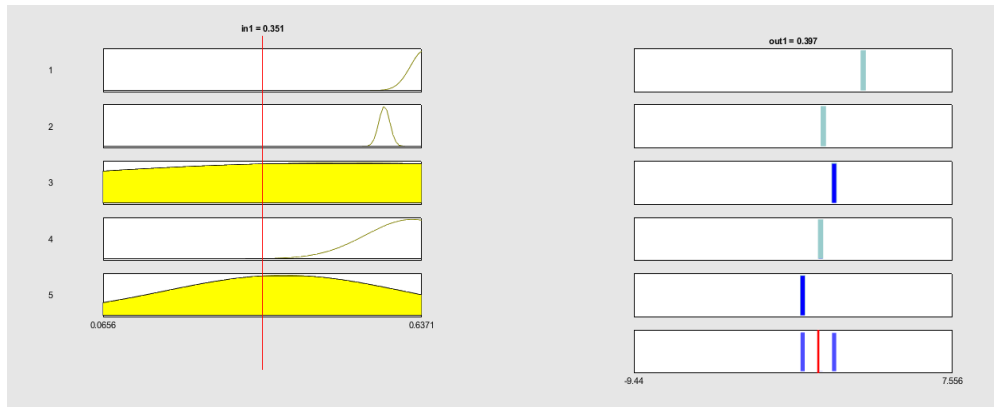


Figura 7.21. Visualización del conjunto de reglas del segundo sistema de inferencia neurodifuso con clustering.
Fuente: Elaboración propia. (Entorno MATLAB®).

Finalmente, se determina la estructura del segundo sistema de inferencia neurodifuso la cual se puede observar en la figura 7.22, en ella se aprecia el único conjunto de entrada, con 5 funciones de pertenencia, junto con las 5 reglas establecidas mediante el clustering para el presente modelo, de manera tal que se pueda obtener una salida en términos de la validación del estado de una imagen de forma rápida y precisa.

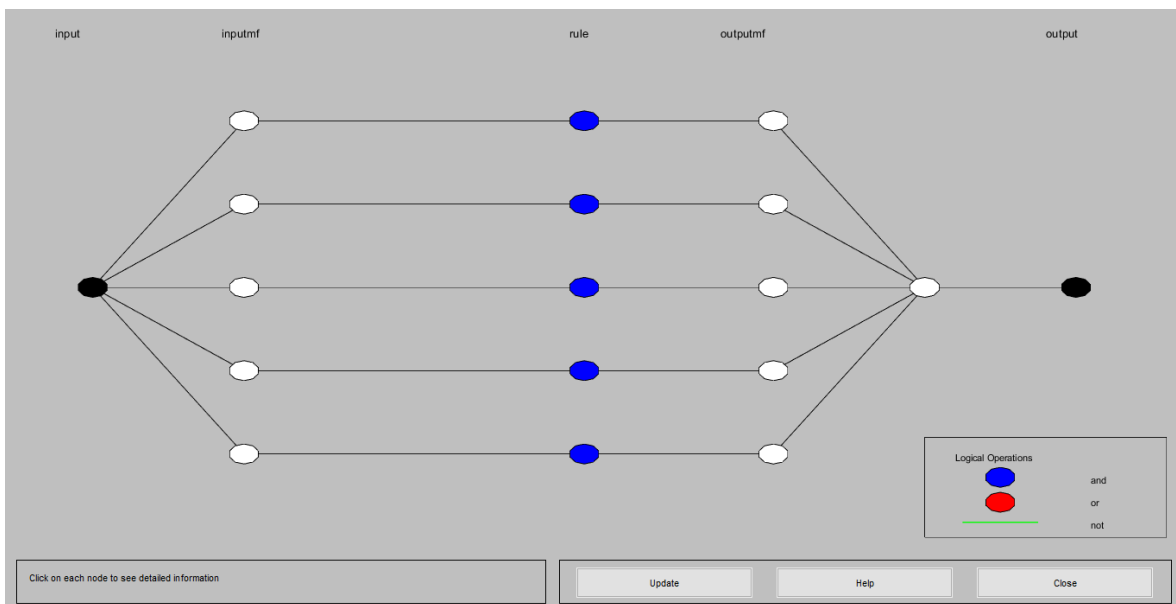


Figura 7.22. Estructura ANFIS para el segundo sistema con clustering
Fuente: Elaboración propia. (Entorno MATLAB®).

7.4.5 Desempeño general del modelo

Tras la consecución de resultados óptimos del segundo sistema de inferencia neurodifusa con clustering, se evalúa su correcta configuración con una matriz de confusión la cual es utilizada comúnmente para evaluar el desempeño de algoritmos de aprendizaje supervisado, para ello se utilizó el conjunto de 1471 imágenes que reúne el conjunto de datos total utilizado para el desarrollo del modelo.

En la figura 7.23 se muestran los resultados de las pruebas indicando que: el modelo con mayor índice de efectividad en la detección de la enfermedad corresponde al segundo sistema de inferencia neurodifusa con clustering optimizado con algoritmo inteligente híbrido, cuyos resultados son: 92,34% de precisión, 95,28% de sensibilidad, 92,40% de especificidad y 93,81% de exactitud.

Detección Xanthomonas Campestris

Salida	693	57	92,40%
	34	687	95,28%
	95,32%	92,34%	93,81%
	Clase objetivo		

*Figura 7.23. Matriz de confusión del modelo con mejor desempeño
Fuente: Elaboración propia*

De esta manera, a diferencia de los resultados obtenidos en modelos similares diseñados para la detección de enfermedades, se obtiene un modelo altamente interpretable y preciso respecto aquellos que se han enfocado principalmente en obtener buenos resultados en sus modelos como los que se encuentran en la tabla 7.15.

Por ejemplo, algunos autores proponen modelos con técnicas como Redes neuronales, Árboles de Decisión o Máquina de Vectores de Soporte, optimizados de acuerdo a sus necesidades y que alcanzan una precisión en un intervalo del 88% a 93%, pero sus resultados no proporcionan interpretabilidad, similar a lo que sucede en una caja negra, por lo tanto, son modelos de difícil

comprensión; de igual manera, si bien existe un modelo interpretable en comparación con el propuesto, su precisión es inferior representando una ventaja sobre el desarrollado en este proyecto.

Tabla 7.15. Comparación del modelo propuesto respecto a los encontrados en literatura relacionada
Fuente: Elaboración Propia

Autores	Conjunto de Datos			Modelo de identificación	Exactitud	Interpretabilidad	
	Tipo de Cultivo	Enfermedad	Fuente			Interpretable	No Interpretable
Propio	15 especies	Xanthomonas Campestris	PlantVillage, ImageNet, Google y Cámara digital	Sistemas de inferencia lógica neurodifusa tipo Sugeno optimizado con algoritmo inteligente híbrido	93,81%	X	
(Van De Vijver et al. 2020)	Solanum tuberosum L.	Alternaria solani	Sensor de imágenes	Detección a partir de árbol de decisión y SVM, optimizado con optimizador de gradiente	92%		X
(Arribas et al. 2011)	Helianthus annuus	Maleza	Cámara digital	Red neuronal optimizada con algoritmo genético	90%		X
(Omrani et al. 2014)	Manzano	Alternaria, la mancha negra de la manzana y la plaga del minador	Cámara digital y de computador	Sistema adaptativo neurodifuso desarrollado con Levenberg-Marquardt	89%	X	
(Yang et al. 2020)	14 especies	21 enfermedades	Base de datos PlantVillage e ImageNet	Redes neuronales convolucionales optimizadas con descenso de gradiente estocástico	93,05%		X
(Deng et al. 2016)	Plantas de cítricos	Citrus Huanglongbing	Cámara digital	Modelo de detección a partir de SVM, optimizado con Particle Swarm Optimization	88,20%		X
(Zhang and Meng 2011)	Plantas de cítricos	Xanthomonas citri	Cámara digital	Modelo de detección jerárquica a partir de redes neuronales, optimizado con algoritmo de unión de ventanas	88%		X

7.4.6 Prototipo Software

En esta subsección se expone el uso del aplicativo desarrollado que permite utilizar el modelo elaborado en la presente investigación. El aplicativo desarrollado permite ser usado en ambiente web, ambiente de desarrollo y en ambiente de escritorio, los siguientes son los pasos para su uso.

Ambiente web

El aplicativo está desplegado en un servidor en la Nube de Amazon, permitiendo de esta manera garantizar su disponibilidad y acceso. Usted puede mediante su navegador, dirigirse mediante la URL al <http://34.201.19.202/webapps/home/> como se muestra en la figura 7.24, y acceder al servicio sin necesidad de tener que instalar algún software adicional en su pc. En este caso, de clic en el software denominado DetXantCamp.



Figura 7.24. Aplicación en servidor web

Fuente: Elaboración propia. (Entorno MATLAB®).

Una vez llevado a cabo el anterior paso, se le abrirá una interfaz similar a la mostrada en la siguiente figura 7.25.

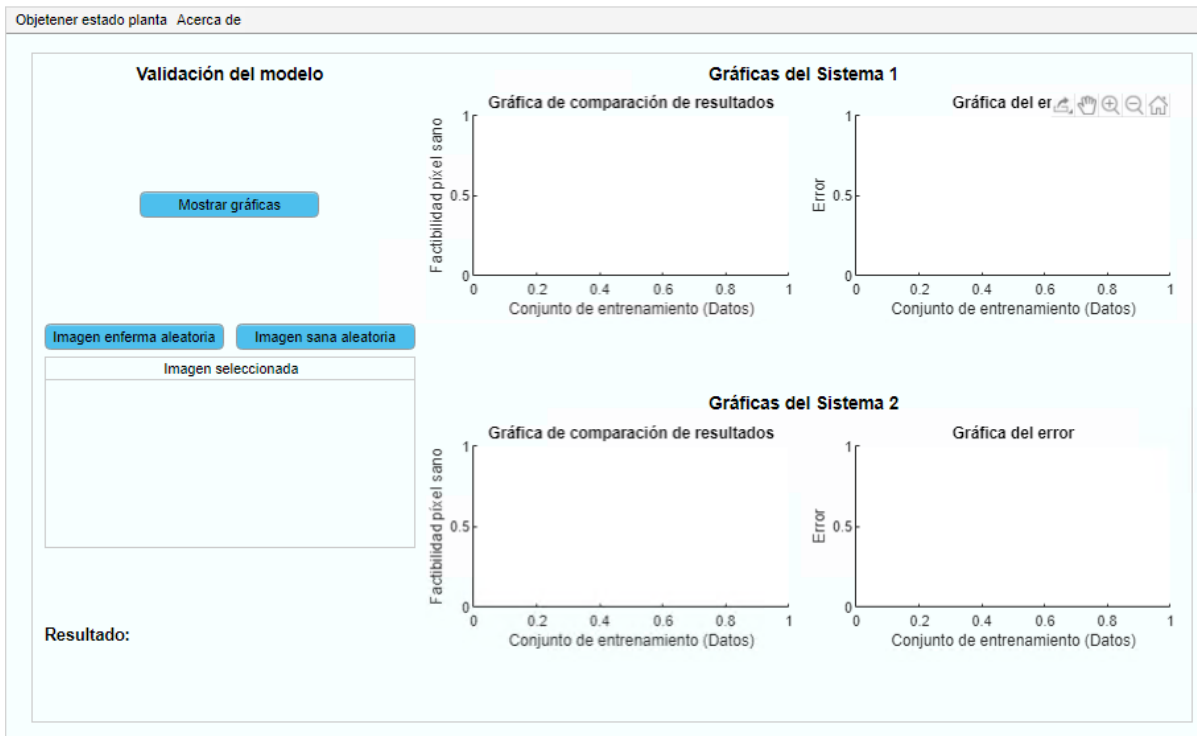


Figura 7.25. Interfaz de usuario DetXantCamp
Fuente: Elaboración propia. (Entorno MATLAB®).

Observe que, en la parte superior izquierda se muestra las diferentes opciones generales a nivel de ventanas de la aplicación, entre estas esta la página principal denominada obtener estado de planta, mostrada en la anterior figura, y la pestaña denominada Acerca de.

Con respecto a la página principal, se observa que hay inicialmente 3 opciones, la opción mostrar gráficas, que permite evidenciar el comportamiento del entrenamiento en cuanto a graficas de comparación de resultados y graficas del error de ambos sistemas pertenecientes al modelo como se muestra en figura 7.26.

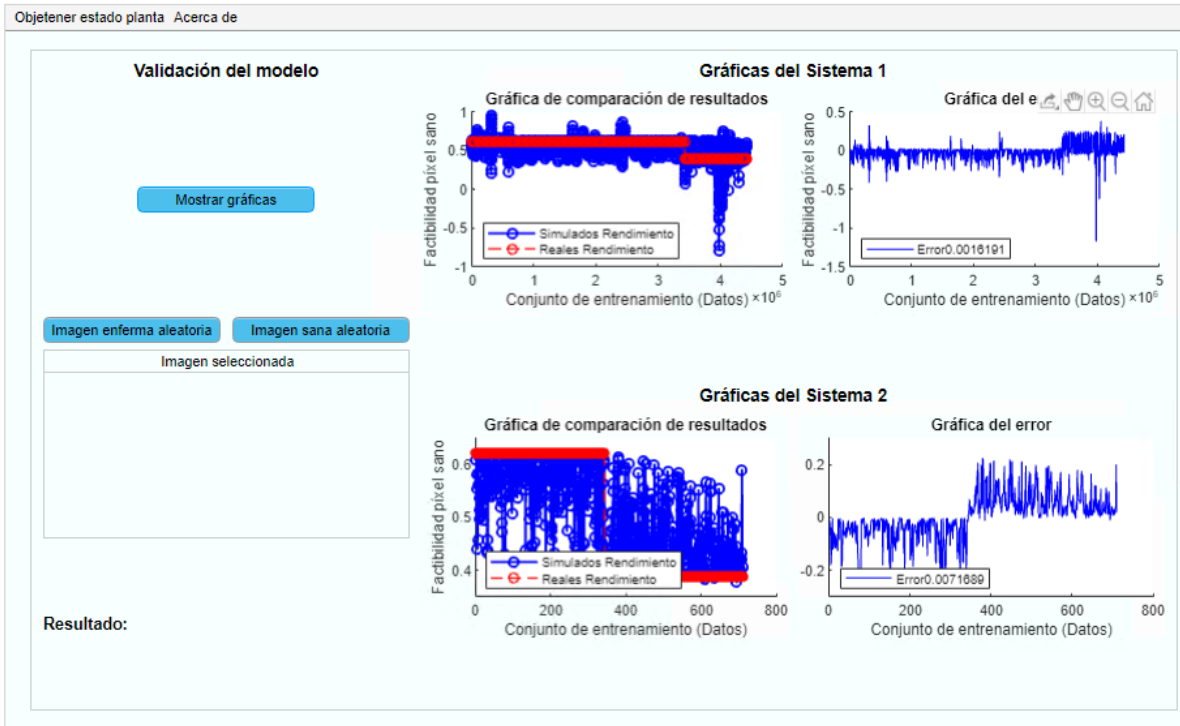


Figura 7.26. Página Principal
Fuente: Elaboración propia. (Entorno MATLAB®).

Adicionalmente, observe que hay otras dos opciones habilitadas a nivel de botón las cuales son imagen enferma aleatoria e imagen sana aleatoria, estas permiten seleccionar aleatoriamente del banco de imágenes del proyecto una imagen que cumpla con esas características para ser evaluada por el modelo. Una vez usted de clic en cualquiera de estas opciones le aparecerá la imagen elegida y posteriormente se le habilitará el botón analizar imagen como se muestra en la figura 7.27.

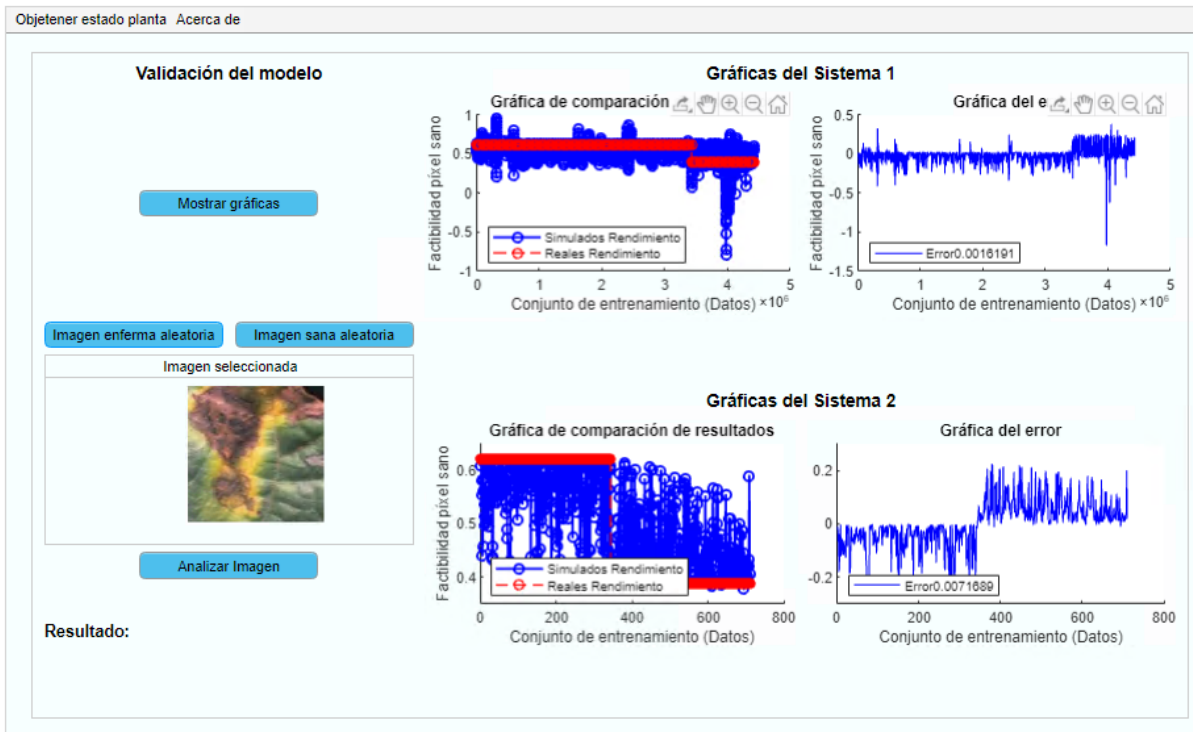


Figura 7.27. Selección de imagen a diagnosticar
Fuente: Elaboración propia. (Entorno MATLAB®).

Posteriormente, de clic en el botón analizar imagen, para que el software proceda a analizar la misma en el modelo desarrollado y dé un resultado sobre si en la imagen se encuentra evidenciada la enfermedad o no como se observa en la figura 7.28.

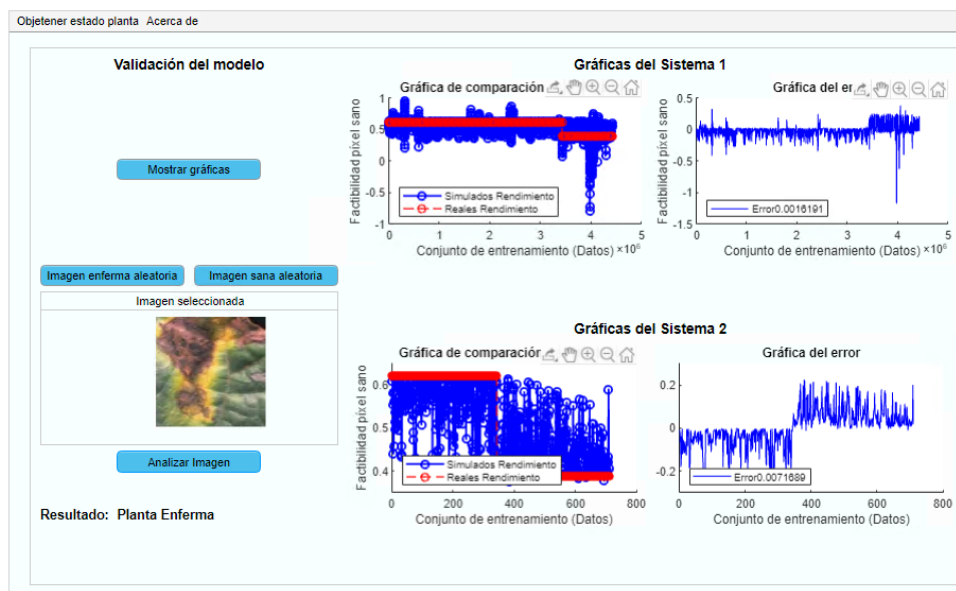


Figura 7.28. Diagnóstico del estado de la planta dado por la aplicación
Fuente: Elaboración propia. (Entorno MATLAB®).

En la figura 7.29, se observa la ventana denominada Acerca de, en esta se puede observar la documentación del software.

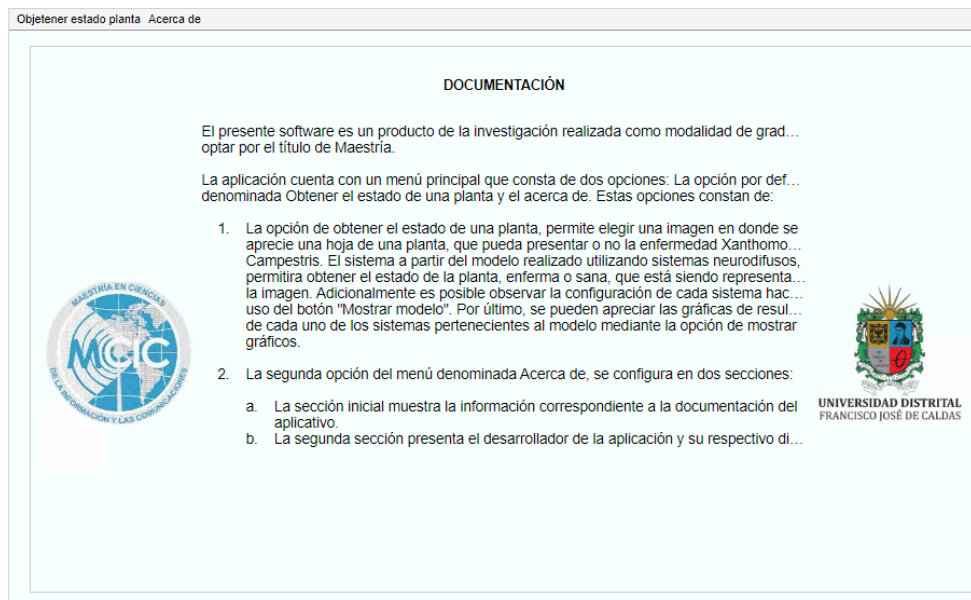
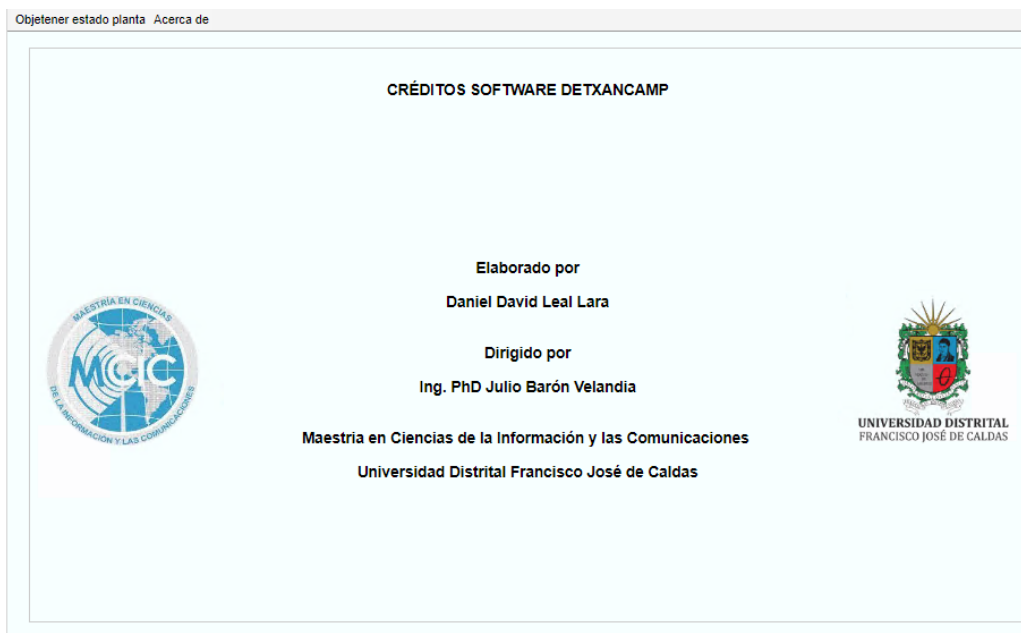


Figura 7.29. Documentación del software
Fuente: Elaboración propia. (Entorno MATLAB®).

Y por último, en la figura 7.30 se observan los créditos del software.



*Figura 7.30. Créditos Software DetXantCamp
Fuente: Elaboración propia. (Entorno MATLAB®).*

Ambiente de desarrollo

La interfaz realizada para el entorno de desarrollo sigue la misma estructura general que la interfaz realizada para el ambiente web, los cambios que se evidencian es con respecto a la posibilidad que en este ambiente se pueda mostrar la configuración de los sistemas con los botones mostrar sistema, y la posibilidad de seleccionar una imagen desde el propio PC, la interfaz del aplicativo se observa en la figura 7.31.

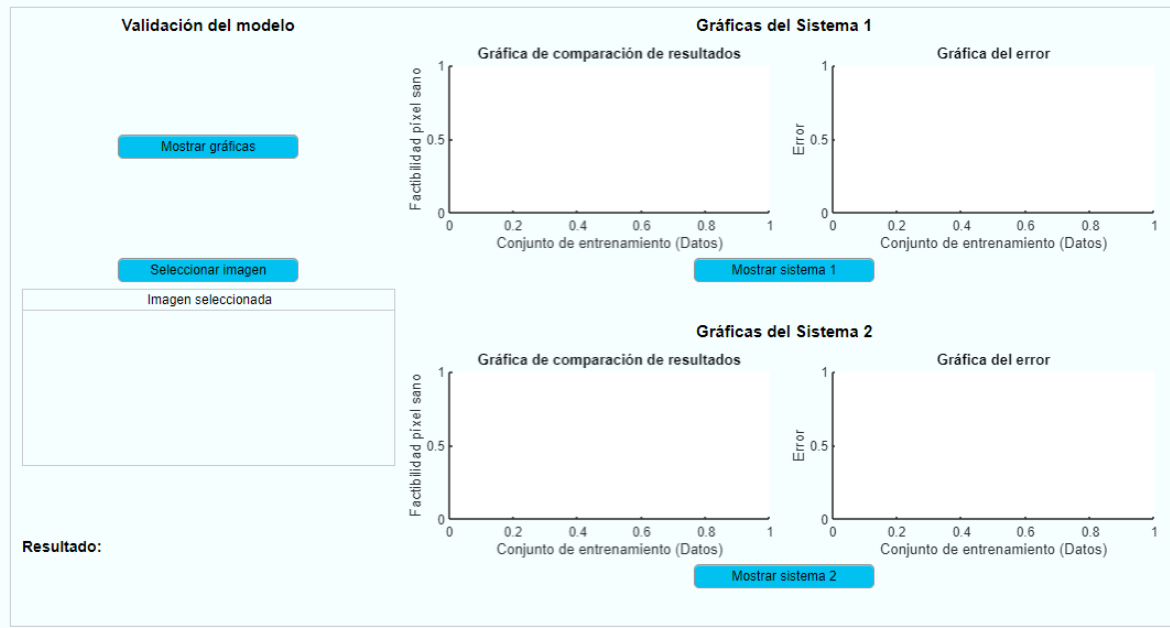


Figura 7.31. Interfaz de entorno
Fuente: Elaboración propia. (Entorno MATLAB®).

Ambiente de escritorio

Para realizar el despliegue de la aplicación de escritorio, como primer paso, debe remitirse al archivo denominado “MyAppInstaller_web.exe”, como se muestra en la figura 7.32 y ejecútelo.

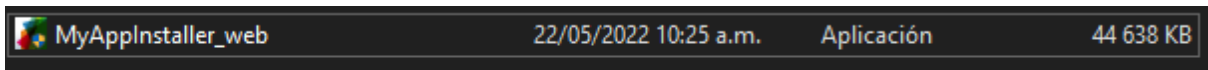
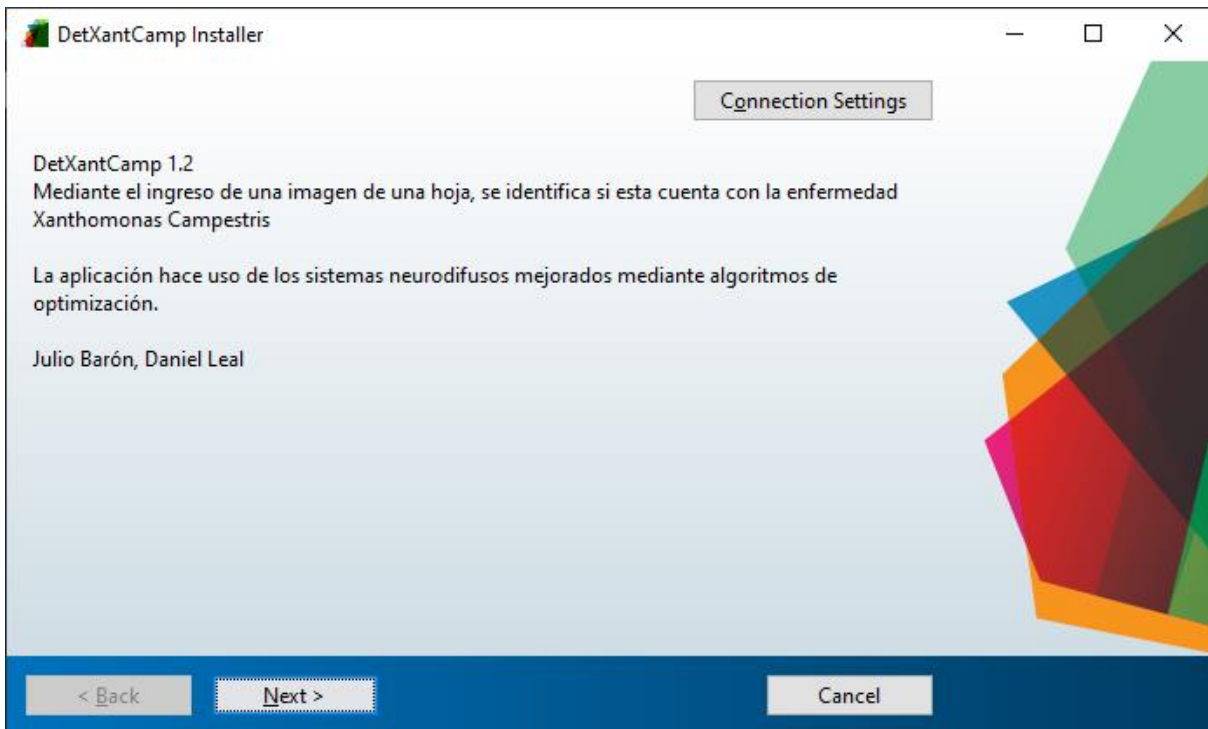


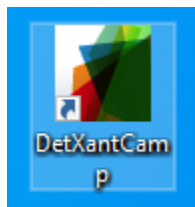
Figura 7.32. Instalación de la aplicación
Fuente: Elaboración propia.

Posteriormente, acepte los términos y condiciones del uso de aplicación, y continúe con el proceso de instalación como se observa en la figura 7.33, hasta culminarlo.



*Figura 7.33. Instalación y términos y condiciones de la aplicación
Fuente: Elaboración propia*

Una vez culminada la instalación, en el escritorio se le mostrara un icono de acceso directo al aplicativo recién instalado similar a la figura 7.34.



*Figura 7.34. Ícono en escritorio de aplicación
Fuente: Elaboración propia*

8. CONCLUSIONES

El modelo de predicción propuesto para la detección automatizada de la enfermedad causada por *Xanthomonas Campestris*, es un modelo basado en el reconocimiento de patrones y procesamiento digital de imágenes que tiene en cuenta las etapas de procesamiento, segmentación y clasificación. Se realiza la captura de información de distintas fuentes de imágenes desde especializadas a propias en modelo de color RGB en condiciones aleatorias de luz, sombra, brillo y contraste, para posteriormente en modelo de color HSB, el modelo de aprendizaje automático tuviese la capacidad de reconocer óptimamente las características de la enfermedad, sin que se vea afectado por características medioambientales como la luz, sombra, brillo y contraste, este proceso representó mejor definición de los objetos y mejorías notorias en el entrenamiento del modelo.

El modelo propuesto se compone de 2 configuraciones diferentes de sistemas neurodifusos optimizados por método híbrido combinado a partir de técnicas heurísticas y exactas como lo son el algoritmo de punto interior y los algoritmos genéticos, obteniendo de esta manera un alto grado de interpretabilidad al ser un modelo basado en lógica difusa, y cuya técnica de optimización permitió aumentar la precisión sin afectar significativamente la interpretabilidad, permitiendo a los usuarios entender la dinámica del modelo conceptual planteado con base en técnicas de Machine Learning alcanzando una exactitud de 93,81% el cual es superior a otros modelos propuestos enfocados únicamente en la precisión.

Se evidencia que utilizando el método de agrupamiento difuso FCMClustering, se obtienen mejores resultados para los diversos sistemas propuestos en la presente investigación. Obteniendo valores MSE de para el primer sistema de $1,62E-03$ y $8,30E-04$ en el entrenamiento y prueba respectivamente, y para el segundo sistema valores de $1,62E-03$ y $6,18E-03$, verificando de esta manera la efectividad de los sistemas que componen el modelo.

Por medio de pruebas estadísticas específicas en el campo de la inteligencia artificial, se presenta al usuario un modelo confiable capaz de optimizar el proceso de cuidado y la productividad de los cultivos de 15 especies distintas con un 92,34% de precisión, 95,28% de sensibilidad y 93,81% de exactitud en su diagnóstico.

A diferencia de los resultados alcanzados en diversas propuestas de investigación, en donde los modelos realizados se enfocan principalmente en la efectividad, dejando de lado la interpretabilidad, el modelo propuesto cuenta con una exactitud del 93,81% en la correcta clasificación del estado de las hojas con respecto a la enfermedad *Xanthomonas Campestris*, sin dejar de lado el nivel de interpretabilidad, ya que permite relacionar la escala cromática HSB con valores cualitativos, logrando dar significancia a la estructuración de las diversas reglas del modelo establecidas mediante Clustering difuso. Facilitando de esta manera que personas sin conocimientos especializados, entiendan la dinámica del modelo.

8.1 PRINCIPALES APORTES

Con base en los resultados obtenidos en el desarrollo del proyecto de investigación y alcanzando los objetivos propuestos se tienen los siguientes aportes:

- Se obtuvo un modelo interpretable y preciso a partir de un sistema de inferencia neurodifuso optimizado con algoritmo híbrido capacitado para dar un diagnóstico rápido y preciso de forma no destructible a partir de imágenes de hojas de las plantas, definiendo si estas se encuentran infectadas por la enfermedad causada por *Xanthomonas Campestris* representando un beneficio tecnológico para los agricultores empíricos.
- El modelo propuesto alcanza una efectividad por encima del 93% y llena un vacío bibliográfico en desarrollo de modelos interpretables para la detección de enfermedades en las plantas, teniendo en cuenta factores medioambientales como lo son la luz, sombra, brillo y contraste encontrados en las imágenes de las plantas, alcanzando buenos resultados y brindando al usuario una explicación de la dinámica del mismo.
- El prototipo software diseñado para la detección de la enfermedad reduce la dependencia de un profesional para el cuidado y seguimiento de los cultivos, al ser

una interfaz apoyada por componentes gráficos sencillos de entender con conocimientos básicos de computación contribuyendo así a la toma de mejores decisiones sobre los cultivos y a su vez mejorar la productividad de los mismos.

9. PROSPECTIVA DEL TRABAJO DE GRADO

9.1 LINEAS DE INVESTIGACIÓN FUTURAS

La agricultura es una de las actividades económicas más antiguas e importantes para la producción interna los países, ya que gracias a ella se obtienen los alimentos y algunos insumos vitales. Por ello, en búsqueda de mejorar su desarrollo en rendimiento y calidad, a lo largo de los años se han implementado junto con el desarrollo de la tecnología nuevas técnicas y metodologías que apoyan los procesos de siembra, cosecha y cuidado.

Sin embargo, esta actividad es realizada en la mayoría de los casos por pequeños agricultores quienes se han visto beneficiados por el interés creciente de investigar la aplicación de algoritmos de aprendizaje automático para la identificación de enfermedades de las plantas, un trabajo que ha revolucionado a partir de diversos esfuerzos y proyectos de investigación se vean enfocados en la agricultura, enfocándose principalmente en la posibilidad de automatizar los procesos del agricultor con el objetivo de volver cada vez más rentable las cosechas.

El desarrollo de métodos para la detección y mitigación de enfermedades de las plantas tiene el doble propósito de aumentar el rendimiento de los cultivos y reducir el uso de plaguicidas a medida que la población aumente, por ello es importante el trabajo en las distintas líneas de investigación que se encargan de esta tarea como lo son las granjas inteligentes, el control de la temperatura y humedad, entre otros.

9.2 TRABAJOS DE INVESTIGACIÓN FUTUROS

Dentro de los trabajos de investigación futuros, se proyecta integrar al modelo, un sistema que permita en tiempo real, detectar las hojas de una planta a partir de la imagen de un cultivo, mediante las técnicas de detección y segmentación de bordes, para poder luego ser analizadas por el presente modelo. Adicionalmente, se pretende elaborar un meta-modelo que permita identificar distintos tipos de enfermedad siguiendo el proceso metodológico llevado a cabo en el presente proyecto. Finalmente, se optará por desarrollar un prototipo portable y de fácil acceso para cualquier tipo de población, incluyendo a las

personas ciudadanas mediante una propuesta que incite a la agricultura urbana, lo anterior haciendo uso de IOT y de la Nube.

9.3 PRODUCTOS ACADÉMICOS

En esta sección se presenta los diversos productos resultado del proceso investigativo del proyecto.

9.3.1 Artículos

- Optimization techniques on fuzzy inference systems to detect Xanthomonas campestris disease. International Journal of Electrical and Computer Engineering ISSN: 2088-8708, Publicado.
- Interpretability in the Field of Plant Disease Detection: A Review. Revista Facultad de Ingeniería, ISSN: 0121-1129 Publicado.
- Optimization techniques on neurofuzzy systems to detect Xanthomonas Campestris disease, en desarrollo.

9.3.2 Registro de Software y propiedad intelectual

- “DetXantCamp” Software desarrollado para el presente proyecto, en fase de registro de derechos de autor ante la Dirección Nacional de Derecho de Autor.
- Prototipo de invernadero para el cultivo de vegetales gestionado por un microcontrolador, en desarrollo para posterior estudio con la Oficina de Transferencias de Resultados de Investigación (OTRI) como patente de modelo de utilidad.

9.3.3 Libros

- Modelo para la detección de la enfermedad Xanthomonas Campestris en hojas de judía aplicando algoritmos de optimización genéticos y de gradiente. Pink Petals Publication Pvt Ltd. ISBN: 9-788195-380428. Publicado.

9.3.4 Ponencias en Eventos Científicos

- Consideraciones para la detección de la enfermedad *Xanthomonas Campestris* aplicando sistemas neurodifusos, en desarrollo.

10. BIBLIOGRAFIA

- Ab Jabal, Mohamad Faizal, Suhardi Hamid, Salehuddin Shuib, and Illiasaak Ahmad. 2013. "Leaf Features Extraction and Recognition Approaches to Classify Plant." *Journal of Computer Science* 9(10):1295–1304.
- Adriyendi. 2018. "Fuzzy Logic Using Tsukamoto Model and Sugeno Model on Prediction Cost." *International Journal of Intelligent Systems and Applications* 10(6):13–21.
- Agrios, G. 1995. *INTRODUCCIÓN A LA FITOPATOLOGÍA*.
- Ali, H., M. I. Lali, M. Z. Nawaz, M. Sharif, and B. A. Saleem. 2017. "Symptom Based Automated Detection of Citrus Diseases Using Color Histogram and Textural Descriptors." *Computers and Electronics in Agriculture* 138:92–104.
- Alsuwaidi, Ali, Bruce Grieve, and Hujun Yin. 2018. "Feature-Ensemble-Based Novelty Detection for Analyzing Plant Hyperspectral Datasets." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11(4):1041–55.
- ARAUZ CAVALLINI, LUIS FELIPE. n.d. "FITOPATOLOGÍA. Un Enfoque Agroecológico."
- Ariani, Fenty, and Robby Yuli Endra. 2013. "Implementation of Fuzzy Inference System With Tsukamoto Method for Study Programme." *International Conference on Engineering and Technology Development (Icetd)*:189–200.
- Arribas, Juan Ignacio, Gonzalo V. Sánchez-Ferrero, Gonzalo Ruiz-Ruiz, and Jaime Gómez-Gil. 2011. "Leaf Classification in Sunflower Crops by Computer Vision and Neural Networks." *Computers and Electronics in Agriculture* 78(1):9–18.
- Baquero Hernández, Lionel Rodolfo, Osviel Rodríguez Valdés, and Febe Ángel Ciudad Ricardo. 2016. "Lógica Difusa Basada En La Experiencia Del Usuario Para Medir La Usabilidad." *Revista Latinoamericana de Ingeniería de Software* 4(1):48.
- Barbedo, Jayme Garcia Arnal, Luciano Vieira Koenigkan, and Thiago Teixeira Santos. 2016. "Identifying Multiple Plant Diseases Using Digital Image Processing." *Biosystems Engineering* 147:104–16.
- Behmann, Jan, Jörg Steinrücken, and Lutz Plümer. 2014. "Detection of Early Plant Stress Responses in Hyperspectral Images." *ISPRS Journal of Photogrammetry and Remote Sensing* 93:98–111.
- Calvo, Ingrid. 2016. "Proyectacolor. Modelos de Color." 15.
- Cao, Feng, Fei Liu, Han Guo, Wenwen Kong, Chu Zhang, and Yong He. 2018. "Fast Detection of Sclerotinia Sclerotiorum on Oilseed Rape Leaves Using Low-Altitude Remote Sensing Technology." *Sensors (Switzerland)* 18(12).
- Chouhan, Siddharth Singh, Uday Pratap Singh, Utkarsh Sharma, and Sanjeev Jain. 2021. "Leaf Disease Segmentation and Classification of Jatropha Curcas L. and Pongamia Pinnata L. Biofuel Plants Using Computer Vision Based Approaches." *Measurement: Journal of the International Measurement Confederation* 171:108796.
- Chuvieco, E. 1990. "Fundamentos de Teledetección Espacial." (*Fundamentals of Remote Sensing from Space*).
- Conselleria d'Agricultura, Desenvolupament Rural, and Emergència Climàtica i Transició Ecològica. 2007. "Xanthomona Campestris Pv. Pruni."
- Correa-Henao, Gabriel Jaime, and Lina María Montoya-Suárez. 2013. "Aplicación Del Modelo ANFIS Para Predicción de Series de Tiempo." *Lámpsakos* (9):12.
- Cruz, Albert, Yiannis Ampatzidis, Roberto Pierro, Alberto Materazzi, Alessandra Panattoni, Luigi De Bellis, and Andrea Luvisi. 2019. "Detection of Grapevine Yellows Symptoms in Vitis Vinifera L. with Artificial Intelligence." *Computers and Electronics in Agriculture*

- 157(November 2018):63–76.
- D. Guzmán, V. M. Castaño. 2006. “La Lógica Difusa En Ingeniería: Principios, Aplicaciones y Futuro.” *Ciencia y Tecnología: Revista de La Universidad de Costa Rica* 24(2):87–107.
- Deng, Xiaoling, Yubin Lan, Tiansheng Hong, and Junxi Chen. 2016. “Citrus Greening Detection Using Visible Spectrum Imaging and C-SVC.” *Computers and Electronics in Agriculture* 130:177–83.
- Desktop, ArcGIS for. 2021. “Función Conversión de Modelo de Color.” Retrieved (<https://pro.arcgis.com/es/pro-app/latest/help/analysis/raster-functions/color-model-conversion-function.htm>).
- Elhariri, Esraa, Nashwa El-Bendary, and Aboul Ella Hassanien. 2014. “Plant Classification System Based on Leaf Features.” *Proceedings of 2014 9th IEEE International Conference on Computer Engineering and Systems, ICCES 2014* 271–76.
- Francis, Jobin, Anto Sahaya Dhas D, and Anoop B K. 2016. “Identification of Leaf Diseases in Pepper Plants Using Soft Computing Techniques.” Pp. 168–73 in. Institute of Electrical and Electronics Engineers (IEEE).
- Garcés Ruiz, Alejandro, and Oscar Gómez Carmona. 2008. “Solución Al Problema Del Despacho Hidrotérmico Mediante Simulación de Monte Carlo y Punto Interior.” *Revista Facultad de Ingeniería* (45):132–47.
- Genetic Algorithm Options - MATLAB & Simulink. (n.d.). Retrieved April 18, 2022, from <https://www.mathworks.com/help/gads/genetic-algorithm-options.html>
- Gu, Qing, Li Sheng, Tianhao Zhang, Yuwen Lu, Zhijun Zhang, Kefeng Zheng, Hao Hu, and Hongkui Zhou. 2019. “Early Detection of Tomato Spotted Wilt Virus Infection in Tobacco Using the Hyperspectral Imaging Technique and Machine Learning Algorithms.” *Computers and Electronics in Agriculture* 167(March):105066.
- Guo, Pengfei, Xuezhi Wang, and Yingshi Han. 2010. “The Enhanced Genetic Algorithms for the Optimization Design.” *Proceedings - 2010 3rd International Conference on Biomedical Engineering and Informatics, BMEI 2010* 7(Bmei):2990–94.
- Hernández, S., and Juan L. López. 2020. “Uncertainty Quantification for Plant Disease Detection Using Bayesian Deep Learning.” *Applied Soft Computing Journal* 96:106597.
- IICCA. 1997. “Técnicas Para El Diagnóstico de Enfermedades En Plantas.”
- Isanta Navarro, Roger. 2013. “Study of a Neural Network-Based System for Stability Augmentation of an Airplane.” (September).
- Jang, Jyh Shing Roger. 1993. “ANFIS: Adaptive-Network-Based Fuzzy Inference System.” *IEEE Transactions on Systems, Man and Cybernetics* 23(3):665–85.
- Kasinathan, Thenmozhi, Dakshayani Singaraju, and Srinivasulu Reddy Uyyala. 2021. “Insect Classification and Detection in Field Crops Using Modern Machine Learning Techniques.” *Information Processing in Agriculture* 8(3):446–57.
- Kaur, Sukhvir, Shreelekha Pandey, and Shivani Goel. 2018. “Semi-Automatic Leaf Disease Detection and Classification System for Soybean Culture.” *IET Image Processing* 12(6):1038–48.
- Kerkech, Mohamed, Adel Hafiane, and Raphael Canals. 2020a. “VddNet: Vine Disease Detection Network Based on Multispectral Images and Depth Map.” *Remote Sensing* 12(20):1–18.
- Kerkech, Mohamed, Adel Hafiane, and Raphael Canals. 2020b. “Vine Disease Detection in UAV Multispectral Images Using Optimized Image Registration and Deep Learning Segmentation Approach.” *Computers and Electronics in Agriculture* 174(April).
- Khan, Saiqa, and Meera Narvekar. 2020. “Novel Fusion of Color Balancing and Superpixel Based Approach for Detection of Tomato Plant Diseases in Natural Complex Environment.” *Journal of King Saud University - Computer and Information Sciences*

- (xxxx).
- Kim, Wan Soo, Dae Hyun Lee, and Yong Joo Kim. 2020. "Machine Vision-Based Automatic Disease Symptom Detection of Onion Downy Mildew." *Computers and Electronics in Agriculture* 168(November):105099.
- Knapp, R. B. 2007a. "Fuzzy Sets and Pattern Recognition - Adaptive Fuzzy."
- Knapp, R. B. 2007b. "Fuzzy Sets and Pattern Recognition - Fuzzy Inference Systems."
- Leal, Lara Daniel David. 2021. "Interpretability in the Field of Plant Disease Detection : A Review." *Revista Facultad de Ingeniería* 30(December):22.
- Luis Espinoza Pineda, Francisco, and Martin Urbina Chavarría. 2011. *L, Espinoza: Introducción a La Fitopatología*.
- Marín-Ortiz, Juan Carlos, Lilliana María Hoyos-Carvajal, and Veronica Botero Fernandez. 2018. "Detección de Plantas Asintomáticas de Solanum Lycopersicum L. Infectadas Con Fusarium Oxysporum Usando Espectroscopia de Reflectancia VIS." *Revista Colombiana de Ciencias Hortícolas* 12(2):436–46.
- Matsunaga, Tomoko M., Daisuke Ogawa, Fumio Taguchi-Shiobara, Masao Ishimoto, Sachihiko Matsunaga, and Yoshiki Habu. 2017. "Direct Quantitative Evaluation of Disease Symptoms on Living Plant Leaves Growing under Natural Light." *Breeding Science* 67(3):316–19.
- Mattio, Mónica C., María V. Peretti Canale, María C. Lurá, and María G. Latorre Rapela. 2018a. "Validation of the Technique Based on Dot Blot for the Detection of Cercospora Kikuchii in Soybean Plants." *Revista Argentina de Microbiología* 50(2):211–15.
- Mattio, Mónica C., María V. Peretti Canale, María C. Lurá, and María G. Latorre Rapela. 2018b. "Validation of the Technique Based on Dot Blot for the Detection of Cercospora Kikuchii in Soybean Plants." *Revista Argentina de Microbiología* 50(2):211–15.
- Mustafa, M. S., Z. Husin, W. K. Tan, M. F. Mavi, and R. S. M. Farook. 2020. "Development of Automated Hybrid Intelligent System for Herbs Plant Classification and Early Herbs Plant Disease Detection." *Neural Computing and Applications* 32(15):11419–41.
- N. H. Reddy, E. R. Kumar, M. V. Reddy, K. L. Reddy y G. S. Valli. 2019. *Bioinformatics and Image Processing—Detection of Plant Diseases*. Springer Singapore.
- Nagano, Shogo, Shogo Moriyuki, Kazumasa Wakamori, Hiroshi Mineno, and Hirokazu Fukuda. 2019. "Leaf-Movement-Based Growth Prediction Model Using Optical Flow Analysis and Machine Learning in Plant Factory." *Frontiers in Plant Science* 10(March):1–10.
- Nturambirwe, Jean Frederic Isingizwe, and Umezuruike Linus Opara. 2020. "Machine Learning Applications to Non-Destructive Defect Detection in Horticultural Products." *Biosystems Engineering* 189:60–83.
- Omrani, Elham, Benyamin Khoshnevisan, Shahaboddin Shamsirband, Hadi Saboohi, Nor Badrul Anuar, and Mohd Hairul Nizam Md Nasir. 2014. "Potential of Radial Basis Function-Based Support Vector Regression for Apple Disease Detection." *Measurement: Journal of the International Measurement Confederation* 55:512–19.
- Ozguven, Mehmet Metin, and Kemal Adem. 2019. "Automatic Detection and Classification of Leaf Spot Disease in Sugar Beet Using Deep Learning Algorithms." *Physica A: Statistical Mechanics and Its Applications* 535:122537.
- Perangin-angin, Mochammad Iswan, Wirda Fitriani, Nova Mayasari, Andysah Putera, and Utama Siahaan. 2016. "Tuition Reduction Determination Using Fuzzy Tsukamoto." *International Journal of Engineering Science Invention* 5(9):68–72.
- Petrellis, Nikos. 2019. "Plant Disease Diagnosis with Color Normalization." *2019 8th International Conference on Modern Circuits and Systems Technologies, MOCASST 2019* (May 2019).
- Poblete, T., C. Camino, P. S. A. Beck, A. Hornero, T. Kattenborn, M. Saponari, D. Boscia,

- J. A. Navas-Cortes, and P. J. Zarco-Tejada. 2020. "Detection of Xylella Fastidiosa Infection Symptoms with Airborne Multispectral and Thermal Imagery: Assessing Bandset Reduction Performance from Hyperspectral Analysis." *ISPRS Journal of Photogrammetry and Remote Sensing* 162(February):27–40.
- Queizan, Angel. 2006. "Algoritmos de Punto Interior Para Optimización Estructural Frente a Acciones Sísmicas." *Mecánica Computacional XXV*:1553–63.
- Quintero-Domínguez, Luis A., Manuel Fajardo, and Sancti Spíritus. 2019. "Sistema Experto Para El Diagnóstico Presuntivo de Enfermedades Fúngicas En Los Cultivos Expert System for the Presumptive Diagnosis of Fungal Diseases in Crops." *Revista Cubana de Ciencias Informáticas* 13(1):61–75.
- Rao, Singiresu S. 2009. *Engineering Optimization: Theory and Practice*. Vol. 77.
- Robert Czabanski, Michal Jezewski and Jacek Leski. 2017. "Introduction to Fuzzy Systems Robert." 23–43.
- Ross, Timothy J. 2010. *Fuzzy Logic with Engineering Applications*. John Wiley.
- S, Lakshmi, and Sivakumar R. 2019. "Plant Phenotyping Through Image Analysis Using Nature Inspired Optimization Techniques." *Springer International Publishing* 150:165–87.
- Sambasivam, G., and Geoffrey Duncan Opiyo. 2021. "A Predictive Machine Learning Application in Agriculture: Cassava Disease Detection and Classification with Imbalanced Dataset Using Convolutional Neural Networks." *Egyptian Informatics Journal* 22(1):27–34.
- Santoso, Heri, Hiroshi Tani, Xuifeng Wang, Agus Eko Prasetyo, and Rei Sonobe. 2019. "Classifying the Severity of Basal Stem Rot Disease in Oil Palm Plantations Using WorldView-3 Imagery and Machine Learning Algorithms." *International Journal of Remote Sensing* 40(19):7624–46.
- Sharif, Muhammad, Muhammad Attique Khan, Zahid Iqbal, Muhammad Faisal Azam, M. Ikram Ullah Lali, and Muhammad Younus Javed. 2018. "Detection and Classification of Citrus Diseases in Agriculture Based on Optimized Weighted Segmentation and Feature Selection." *Computers and Electronics in Agriculture* 150(April):220–34.
- Sharma, Parul, Yash Paul Singh Berwal, and Wiqas Ghai. 2020. "Performance Analysis of Deep Learning CNN Models for Disease Detection in Plants Using Image Segmentation." *Information Processing in Agriculture* 7(4):566–74.
- Shin, Jaemyung, Young K. Chang, Brandon Heung, Tri Nguyen-Quang, Gordon W. Price, and Ahmad Al-Mallahi. 2020. "Effect of Directional Augmentation Using Supervised Machine Learning Technologies: A Case Study of Strawberry Powdery Mildew Detection." *Biosystems Engineering* 194:49–60.
- Singh, Vijai, and A. K. Misra. 2017. "Detection of Plant Leaf Diseases Using Image Segmentation and Soft Computing Techniques." *Information Processing in Agriculture* 4(1):41–49.
- Sladojevic, Srdjan, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic. 2016. "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification." *Computational Intelligence and Neuroscience* 2016.
- Tan, Daniel Stanley, Robert Neil Leong, Ann Franchesca Laguna, Courtney Anne Ngo, Angelyn Lao, Divina M. Amalin, and Dionisio G. Alvindia. 2018. "AuToDiDAC: Automated Tool for Disease Detection and Assessment for Cacao Black Pod Rot." *Crop Protection* 103(January 2019):98–102.
- Tolerances and Stopping Criteria - MATLAB & Simulink. (n.d.). Retrieved April 21, 2022, from <https://www.mathworks.com/help/optim/ug/tolerances-and-stopping-criteria.html>
- Trivedi, Jay, Yash Shamnani, and Ruchi Gajjar. 2020. "Plant Leaf Disease Detection Using Machine Learning." *Communications in Computer and Information Science* 1214

CCIS:267–76.

- Verma, Brijesh, and Ping Zhang. 2007. "A Novel Neural-Genetic Algorithm to Find the Most Significant Combination of Features in Digital Mammograms." *Applied Soft Computing Journal* 7(2):612–25.
- Vicente, J. G., J. Conway, S. J. Roberts, and J. D. Taylor. 2001. "Identification and Origin of *Xanthomonas Campestris* Pv. *Campestris* Races and Related Pathovars." *Phytopathology* 91(5):492–99.
- VijayaLakshmi, Balasubramanian, and Vasudev Mohan. 2016. "Kernel-Based PSO and FRVM: An Automatic Plant Leaf Type Detection Using Texture, Shape, and Color Features." *Computers and Electronics in Agriculture* 125:99–112.
- Van De Vijver, Ruben, Koen Mertens, Kurt Heungens, Ben Somers, David Nuyttens, Irene Borra-Serrano, Peter Lootens, Isabel Roldán-Ruiz, Jürgen Vangeyte, and Wouter Saey. 2020. "In-Field Detection of *Alternaria Solani* in Potato Crops Using Hyperspectral Imaging." *Computers and Electronics in Agriculture* 168(March):105106.
- Witten, I. H. (Ian H. ..., Eibe Frank, Mark A. (Mark Andrew) Hall, and Christopher J. Pal. n.d. *Data Mining : Practical Machine Learning Tools and Techniques*.
- Wspanialy, Patrick, and Medhat Moussa. 2020. "A Detection and Severity Estimation System for Generic Diseases of Tomato Greenhouse Plants." *Computers and Electronics in Agriculture* 178(January):105701.
- Xia, Ji'An A., Hong Xin Cao, Yu Wang Yang, Wei Xin Zhang, Qian Wan, Lei Xu, Dao Kuo Ge, Wen Yu Zhang, Ya Qi Ke, and Bo Huang. 2019. "Detection of Waterlogging Stress Based on Hyperspectral Images of Oilseed Rape Leaves (*Brassica Napus* L.)." *Computers and Electronics in Agriculture* 159(December 2018):59–68.
- Xiong, Yonghua, Longfei Liang, Lin Wang, Jinhua She, and Min Wu. 2020. "Identification of Cash Crop Diseases Using Automatic Image Segmentation Algorithm and Deep Learning with Expanded Dataset." *Computers and Electronics in Agriculture* 177(January):105712.
- Yadav, Rishabh, Yogesh Kumar Rana, and Sushama Nagpal. 2019. *Plant Leaf Disease Detection and Classification Using Particle Swarm Optimization*. Vol. 11407 LNCS. Springer International Publishing.
- Yadav, Rishabh, Yogesh Kumar, and Sushama Nagpal. 2019. "Plant Leaf Disease Detection and Classification Using Particle Swarm Optimization." *Springer Nature Switzerland* 294–306.
- Yamamoto, Kyosuke, Wei Guo, and Seishi Ninomiya. 2016. "Node Detection and Internode Length Estimation of Tomato Seedlings Based on Image Analysis and Machine Learning." *Sensors (Switzerland)* 16(7).
- Yang, Guofeng, Yong He, Yong Yang, and Beibei Xu. 2020. "Fine-Grained Image Classification for Crop Disease Based on Attention Mechanism." *Frontiers in Plant Science* 11(December):1–15.
- Zapata, Mildred; Gaud, Rosario. 2001. "ESTRATEGIAS PARA DIFERENCIAR *Xanthomonas Campestris* Pv. *Phaseoli* CON SALES INORGÁNICAS." *Agronomía Mesoamericana* 12:1–8.
- Zhang, Min, and Qinggang Meng. 2011. "Automatic Citrus Canker Detection from Leaf Images Captured in Field." *Pattern Recognition Letters* 32(15):2036–46.
- Zhang, Shanwen, Zhuhong You, and Xiaowei Wu. 2019. "Plant Disease Leaf Image Segmentation Based on Superpixel Clustering and EM Algorithm." *Neural Computing and Applications* 31:1225–32.
- Zhu, Hongyan, Bingquan Chu, Chu Zhang, Fei Liu, Linjun Jiang, and Yong He. 2017. "Hyperspectral Imaging for Presymptomatic Detection of Tobacco Disease with Successive Projections Algorithm and Machine-Learning Classifiers." *Scientific*

Reports 7(1):1–12.

Zhuang, Jiayu, Shiwei Xu, Zhemin Li, Wei Chen, and Dongjie Wang. 2015. “Application of Intelligence Information Fusion Technology in Agriculture Monitoring and Early-Warning Research.” Pp. 114–17 in *Proceedings - 2015 International Conference on Control, Automation and Robotics, ICCAR 2015*. Institute of Electrical and Electronics Engineers Inc.
