



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO

CENTRO UNIVERSITARIO UAEM TEXCOCO



**SISTEMA EXPERTO CON VISIÓN ARTIFICIAL PARA LA IDENTIFICACIÓN DE
INSECTOS DESCORTEZADORES EN ESPECIES DE PINO DE LOS
ESTADOS FORESTALES MÁS IMPORTANTES DE MÉXICO.**

T E S I S

**QUE PARA OBTENER EL GRADO DE:
MAESTRO EN CIENCIAS DE LA COMPUTACIÓN.**

**PRESENTA:
JORGE RODRIGO DÁVILA GARCÍA.**

**TUTORA ACADÉMICA:
M. EN C. YEDID ERANDINI NIÑO MEMBRILLO.**

**TUTORES ADJUNTOS:
DR. EN C. DE LA COMP. FARID GARCÍA LAMONT.
M. EN C. RODOLFO CAMPOS BOLAÑOS.**

TEXCOCO, ESTADO DE MÉXICO, ENERO DE 2019 .



DICTÁMEN DE AUTORIZACIÓN Y OBTENCIÓN DE GRADO DE MAESTRÍA

Título del proyecto:

Sistema Experto con Visión Artificial para la Identificación de Insectos Descortezadores en Especies de Pino de los Estados Forestales más importantes de México

Tesista:

Jorge Rodrigo Dávila García

Dictamen:

No. de revisión: _____

- Rechazado
- Sujeto a modificaciones
- Aceptado, condicionado
- Aceptado

Observaciones generales:

Acceptado el Sistema de Información
Acceptado en Documento para Impresión
Acceptado para la Defensa de Grado

Tutor Adjunto	Tutor Académico	Tutor Adjunto
Dr. en C. de la Comp. Farid García Lamont	M. en C. Yedid Erandini Niño Membrillo	M. en C. Rodolfo Campos Bolaños

Agradecimientos

...Expreso mi gratitud a todas y cada una de las personas que con sus aportaciones hicieron posible la culminación de este trabajo...

Agradezco a la M. en C. Yedid Erandini Niño Membrillo, al Dr. Farid García Lamont y al M. en C. Rodolfo Campos Bolaños por su invaluable apoyo durante todos mis estudios de maestría.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por la beca otorgada, con el número de registro 783566, y a la Universidad Autónoma del Estado de México Centro Universitario Texcoco.

Al laboratorio de Entomología Forestal de la Universidad Autónoma Chapingo, al laboratorista Rosalio Mejía Vázquez por el apoyo y préstamo de los especímenes de insectos descortezadores utilizados en los experimentos.

Dedicatoria

A mis padres, Alma Rosa García Bustamante y Jorge Dávila Olivares, por haberme apoyado en todo momento, por sus consejos, sus valores, por la motivación constante que me ha permitido ser una persona de bien, pero más que nada, por su amor.

A Liliana y Alondra, por estar conmigo y apoyarme siempre, las quiero mucho.

A todos aquellos familiares que no recordé al momento de escribir esto. Ustedes saben quiénes son.

Contenido

Índice de tablas	VII
Índice de figuras	VIII
1. Introducción	1
2. Planteamiento del Problema	2
3. Justificación	6
4. Objetivo General	6
5. Objetivos Específicos	6
6. Hipótesis	7
7. Metodología	7
8. Marco Teórico	9
8.1. Economía del Sector Forestal	9
8.2. Insectos Descortezadores	10
8.2.1. Género Dendroctonus	11
8.2.2. Género Ips	12
8.3. Materiales y métodos	13
8.4. Herramientas Computacionales	16
8.4.1. Sistema Experto	16
8.4.1.1. Metodología del SE	18
8.4.1.2. Representación de conocimiento con reglas	19
8.4.1.3. PROLOG	20
8.4.2. Técnicas de IA	20
8.4.2.1. Red Neuronal Feed-Forward Backpropagation (RNFF)	21
8.4.2.2. Clasificador Bayesiano (CB)	24
8.4.2.3. Máquina de Soporte Vectorial (SVM)	24
8.4.3. Evaluación de los modelos	26
8.4.3.1. Índice kappa	26
8.4.4. Técnicas de procesamiento de imágenes	27
8.4.4.1. Redimensión de imágenes	27
8.4.4.2. Filtro Gaussiano	27
8.4.4.3. Transformación a escala de grises	27
8.4.4.4. Ecuación de histograma	28
8.4.4.5. Erosión	29
8.4.4.6. Binarización de la imagen a escala de grises	30
8.4.4.6.1. Otsu	30
8.4.5. Extracción de Características	32
8.4.5.1. Análisis de Componentes Principales (ACP)	32
8.4.5.2. Momentos Invariantes	33

8.4.5.3.	Patrones binarios locales (LBP)	34
8.4.5.4.	Descriptores de Haralick	34
9.	Marco de Referencia	37
10.	Propuesta de solución	46
10.1.	Conceptualización del Sistema	46
10.2.	Adquisición de conocimiento para el desarrollo del Sistema	47
10.2.1.	Adquisición de conocimiento: Características insectos descortezadores	48
10.3.	Interfaz de usuario del sistema: Menú Principal	50
10.4.	Diagramas de Caso de Uso	51
11.5.	Módulo por Reglas	62
11.5.1.	Validación de reglas	64
11.5.2.	Validación de interfaz de usuario	65
11.5.3.	Primer prototipo	66
11.5.4.	Segundo prototipo	66
11.5.5.	Descripción de la Interfaz de Usuario	67
11.5.6.	Resultados	69
11.5.7.	Evaluación del Sistema Experto	70
11.5.8.	Discusión	71
11.6.	Módulo de Visión por Computadora	71
11.6.1.	Adquisición de imágenes	72
11.6.2.	Creación de base de datos	72
11.6.3.	Extracción de características	74
11.6.4.	Vector de Características	74
11.6.5.	Clasificación de patrones	75
11.6.6.	Resultados del Experimento	75
11.6.7.	Evaluación del clasificador	76
11.6.8.	Descripción de la interfaz de usuario	76
11.6.9.	Discusión de los Resultados del Experimento	77
11.7.	Módulo de Inteligencia Artificial (IA)	78
11.7.1.	Adquisición de imágenes y observaciones	79
11.7.2.	Técnicas usadas para la identificación de especies de insectos	80
11.7.2.1.	Red Neuronal Feed-Fordward Backpropagation	80
11.7.2.2.	Máquinas de Soporte Vectorial	81
11.7.2.3.	Clasificador Bayesiano	81
11.7.3.	Descripción de la interfaz de usuario del Módulo de Inteligencia Artificial	81
11.7.4.	Resultados del Experimento y Evaluación de las Técnicas	85
11.7.5.	Discusión de los Resultados del Experimento	88
12.	Discusión General	89
13.	Conclusiones	90

14.	<i>Trabajos futuros</i>	90
15.	<i>Bibliografía</i>	91
16.	<i>Anexos</i>	97
	Anexo 1	97

Índice de tablas

Tabla 1. Descortezadores del género <i>Dendroctonus</i> , tamaño y huéspedes que habitan	12
Tabla 2. Especies del género <i>Ips</i> , tamaño y huéspedes que habitan.	13
Tabla 3. Listado de dispositivos probados.	14
Tabla 4. Coeficiente kappa	26
Tabla 5. Ejemplo de matriz de coeficientes	27
Tabla 6. Estructuras de referencia	30
Tabla 7. Características insectos descortezadores	49
Tabla 8. Caso de Uso: Menú Principal	52
Tabla 9. Caso de Uso: BUSLI	53
Tabla 10. Caso de Uso: BUSGUI	55
Tabla 11. Caso de Uso: LISESP	56
Tabla 12. Caso de Uso: Clasificador Supervisado	57
Tabla 13. Caso de Uso: Visión por Computadora	59
Tabla 14. Caso de Uso: Preprocesamiento	60
Tabla 15. Caso de Uso: EXTCAR	61
Tabla 16. Caso de Uso: CLAS	62
Tabla 17. Matriz de confusión del clasificador	76
Tabla 18. Registros por especie	86
Tabla 19. Matriz de confusión RNFF	86
Tabla 20. Matriz de Confusión CB	86
Tabla 21. Matriz de Confusión SVM	87
Tabla 22. Precisiones por clase	87
Tabla 23. Índice kappa por técnica	87

Índice de figuras

Figura 1. Mapa de la superficie de bosque en México.....	3
Figura 2. Superficie Afectada por Descortezadores.....	4
Figura 3. Montaje de celular en estereoscopio.....	15
Figura 4. Fotografías <i>Dendroctonus valens</i> : a) cabeza, b) elitro, c) pronoto, d) cuerpo, e) lateral y f) declive elitral.....	15
Figura 5. Fotografías <i>Ips integer</i> : a) cabeza, b) elitro, c) pronoto, d) cuerpo, e) lateral y f) declive elitral.....	16
Figura 6. Arquitectura de un SE.....	17
Figura 7. Estructura Red Neuronal Feed-Forward.....	21
Figura 8. Función de Transferencia Hard Limit.....	22
Figura 9. Función de Transferencia Lineal.....	22
Figura 10. Función de Transferencia Log-Sigmoid.....	22
Figura 11. Función de Transferencia Hyperbolic Tangent Sigmoid.....	23
Figura 12. Esquema de la arquitectura del Sistema.....	47
Figura 13. Menú Principal del Sistema.....	50
Figura 14. Diagrama de Caso de Uso: Menú Principal.....	51
Figura 15. Diagrama de Caso de Uso: BUSLI.....	51
Figura 16. Diagrama de Caso de Uso: BUSGUI.....	54
Figura 17. Diagrama de Caso de Uso: LISESP.....	54
Figura 18. Diagrama de Caso de Uso: Clasificador Supervisado.....	56
Figura 19. Diagrama de Caso de Uso: Visión por Computadora.....	58
Figura 20. Caso de Uso: Preprocesamiento.....	58
Figura 21. Caso de Uso: EXTCAR.....	60
Figura 22. Caso de Uso: CLAS.....	61
Figura 23. Ejemplo árbol de decisión para <i>D. valens</i> y <i>D. rhizophagus</i>	63
Figura 24. Representación de conocimiento por medio de reglas en Prolog de la especie <i>Dendroctonus adjunctus</i> Blandford.....	64
Figura 25. Interfaz de búsqueda libre.....	67
Figura 26. Interfaz de búsqueda guiada.....	68
Figura 27. Interfaz listado de especies.....	69
Figura 28. Aplicación de Filtro Gaussiano.....	72
Figura 29. Ecualización de imagen a escala de grises.....	73
Figura 30. Binarización de imagen ecualizada.....	73
Figura 31. Erosión de imagen.....	74
Figura 32. Arquitectura Red Neuronal.....	75
Figura 33. Entrenamiento de la Red.....	75
Figura 34. Búsqueda con Visión Artificial.....	77
Figura 35. Segmentación de la imagen.....	80
Figura 36. Estructura Red Neuronal Feed-Forward Backpropagation.....	81
Figura 37. Interfaz Módulo IA: Pregunta 1 a 3.....	82
Figura 38. Interfaz Módulo IA: Pregunta 4.....	83
Figura 39. Interfaz Módulo IA: Pregunta 5.....	84
Figura 40. Interfaz Módulo IA: Pregunta 6.....	85

1. Introducción

Debido a las condiciones climáticas que se presentan hoy en día la aparición y propagación descontrolada de plagas y enfermedades en los bosques de pino se ha incrementado considerablemente, lo que representa un riesgo para la conservación de los ecosistemas forestales, la provisión de los servicios ambientales y la productividad del sector forestal (Deschamps, 2016).

El Programa Nacional Forestal 2014-2018 identificó que una de las problemáticas en éste Sector son las plagas y enfermedades que atacan a los ecosistemas forestales ocasionando su deterioro, teniendo como consecuencia costos por pérdida de oportunidades económicas, de funcionamiento de los ecosistemas, de biodiversidad y de servicios ambientales (SEMARNAT, 2014).

Derivado de lo anterior, uno de los principales problemas a los que se enfrenta el sector forestal es la detección oportuna y combate de plagas y enfermedades en los ecosistemas forestales. Bajo este contexto, este trabajo abordará la problemática enfocándose en la detección oportuna de plagas de insectos descortezadores de los géneros *Dendroctonus* e *Ips* en los bosques de pino debido a que los bosques de pino son una de las principales superficies de vegetación forestal presentes en el país. Para lograr la identificación oportuna es necesario contar con conocimientos avanzados que permitan discernir entre lo normal y anormal, identificando las características de la plaga presente en los árboles de pino.

Con base en lo anterior surge la pregunta de investigación:

*¿De qué manera un sistema experto con visión artificial puede contribuir en la detección oportuna de plagas de insectos descortezadores de los géneros *Dendroctonus* e *Ips*?*

El propósito de la pregunta de investigación es conocer la contribución que realizaría un sistema experto (SE) con visión artificial al Sector Forestal; que identifique plagas de insectos descortezadores a través de una fotografía, sirviendo como directriz para emprender un plan de manejo evitando una propagación mayor que termine en epidemia ocasionando pérdida de superficie boscosa.

Contar con una herramienta computacional, como es el SE, que reúna los conocimientos y experiencia de expertos en el área, permitirá que cualquier persona interesada en la conservación y cuidado de los bosques, en el contexto de esta investigación, pueda desempeñar la tarea de identificación de plagas de insectos descortezadores, ya que se considera de suma importancia para su control y manejo en etapas tempranas.

Para responder la pregunta de investigación anteriormente planteada, se tiene como objetivo: *Analizar si la identificación de plagas con el uso de un sistema experto con visión artificial puede contribuir en la detección temprana de especies de insectos descortezadores de los géneros Dendroctonus e Ips.*

El objetivo de este trabajo de investigación es desarrollar un sistema experto con visión artificial que reconozca las especies de insectos descortezadores de los géneros *Dendroctonus* e *Ips*.

Con base en lo anterior, la hipótesis que se buscará comprobar en el transcurso de la investigación es la siguiente:

H1: El sistema experto con visión artificial reconocerá las especies de insectos descortezadores de los géneros Dendroctonus e Ips.

Para poder responder la pregunta, lograr el objetivo y comprobar la hipótesis planteada, partiendo de la naturaleza de la investigación que es mixta de tipo descriptiva y utilizando un método inductivo-deductivo para realizar el análisis y descripción partiendo de lo particular a lo general y viceversa, se realizarán las siguientes actividades que se encuentran descritas en la sección de metodología:

- Obtención de las muestras.
- Recolección de datos.
- Análisis de datos.
- Generación de base de conocimiento.
- Desarrollo de la herramienta computacional.

Se propone un capitulado tentativo en la sección de marco teórico que se utilizará como directriz para llevar a cabo la obtención de información que servirá como base de conocimiento; y la selección del tipo de herramienta computacional que se desarrollará como parte fundamental de esta investigación.

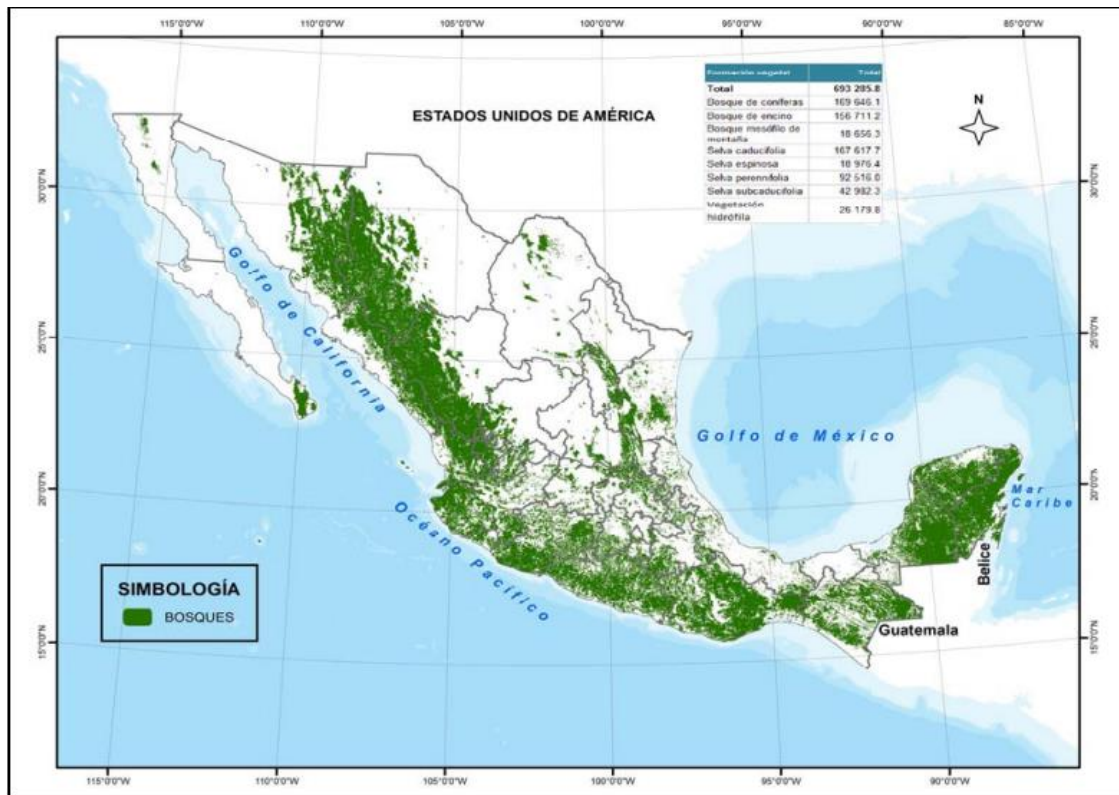
2. Planteamiento del Problema

México cuenta con alrededor de 138 millones de hectáreas con vegetación forestal, equivalente al 70% de la superficie nacional. Los principales ecosistemas que componen esta superficie son: los matorrales xerófilos (41.2%), los bosques templados (24.24%), ver Figura 1, las selvas (21.7%), manglares y otros tipos de asociaciones de vegetación forestal (1.06%) y otras áreas forestales (11.8%) (SEMARNAT, 2014).

Los bosques templados también conocidos como bosques de coníferas y encinos constituyen el segundo tipo de vegetación más importante, su conformación es,

principalmente, por pinos y encinos. Para el caso de los pinos, en México existen 49 especies con una distribución natural de las 120 existentes en el mundo, de las cuales 22 son endémicas del país (Gernandt & Pérez De La Rosa, 2014). La distribución de los bosques se encuentra a lo largo de la Sierra Madre Occidental, la Sierra Madre Oriental y el Eje Neo volcánico.

Figura 1. Mapa de la superficie de bosque en México



Fuente: Dirección General de Geografía y Medio Ambiente. Conjunto de Datos Vectoriales de Uso del Suelo y Vegetación escala 1:250 000 Serie V.

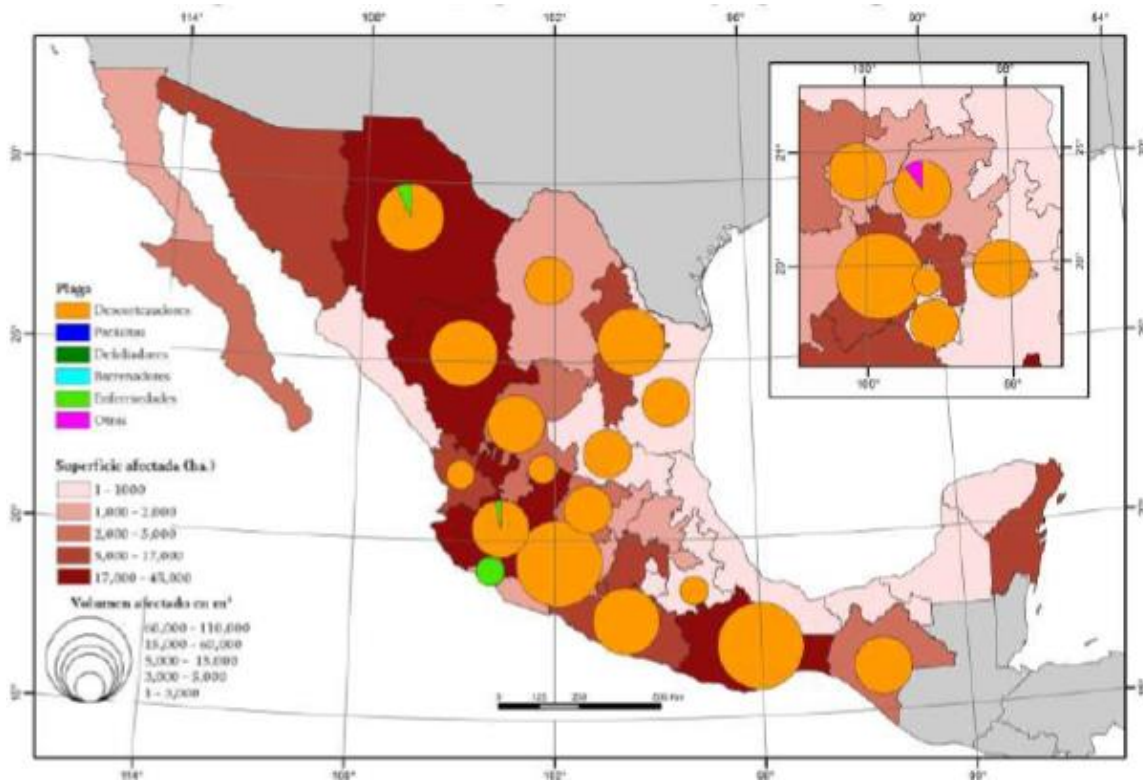
Nota: La superficie considera todos los tipos de bosque.

Cabe agregar que la superficie perturbada por plagas y enfermedades hasta 1998 era de 23 mil hectáreas anuales en promedio, entre 1999 y 2010 se observó un aumento de más del doble ascendiendo hasta 51 mil hectáreas en promedio afectadas por año (SEMARNAT, 2014). Por su parte en 2015 y 2016 se reportaron 139,294 y 91,094 hectáreas afectadas por plagas y enfermedades dentro de las que destacan descortezadores, plantas parásitas (muérdago), plantas epifitas, defoliadores, enfermedades, barrenadores y otras plagas (SEMARNAT, 2016) (SEMARNAT, 2017).

De acuerdo con el Anuario Estadístico de la Producción Forestal los insectos descortezadores representan el 36.9% de la superficie total afectada (SEMARNAT,

2016), ver Figura 2, siendo el principal problema de plaga forestal en México debido a su amplia distribución en los bosques de coníferas, se desarrollan bajo la corteza de los árboles, debilitándolos y provocando la muerte (Magallón, 2013). El daño ocasionado por insectos descortezadores implica pérdidas ambientales y económicas; ambientales por que conlleva la muerte masiva de árboles generando perdida de hábitat de especies afectando la biodiversidad, y económicas disminuye del 50 al 100% del valor comercial de la madera (CONAFOR, 2013).

Figura 2, Superficie Afectada por Descortezadores



Fuente: Anuario Estadístico de Producción Forestal 2016.

Es evidente entonces que la ocurrencia de plagas y enfermedades forestales ha ido en aumento en los últimos años, ante el contexto climático actual, lo que representa un riesgo para la conservación de los ecosistemas forestales, la provisión de los servicios ambientales y la productividad del sector forestal (Deschamps, 2016).

Por su parte las áreas con ecosistemas forestales están habitadas en la actualidad por más de 11 millones de personas, para quienes los recursos forestales representan un capital natural que debería contribuir a satisfacer sus necesidades básicas y mejorar su calidad de vida, sin embargo, lo anterior no se ha podido concretar y es en las zonas forestales donde se presentan los mayores niveles de marginación y pobreza (SEMARNAT, 2014).

Por otro lado, la falta de información para realizar la identificación de las plagas de insectos descortezadores que atacan a los bosques puede ocasionar: la propagación descontrolada de las mismas, la pérdida o reducción de las funciones ambientales de los bosques y pérdidas económicas considerables en el Sector Forestal del país.

Derivado de lo anterior, uno de los principales problemas a los que se enfrenta el sector forestal es la detección oportuna y combate de plagas y enfermedades que presentan los bosques de pino.

Si bien existen métodos de detección comunes: vigilancia de campo, detección sistemática y detección aérea, también los hay especializados que se realizan a través de: fotografía aérea, rastreadores multiespectrales, satélites orbitales, captura de feromonas que desprenden los insectos en estado de apareamiento, (Ciesla, Cadahía, & Robredo, 1976). Todos estos métodos concluyen que se necesita personal experto de plagas y enfermedades que puedan reconocer lo inusual o anormal, además en el caso de los métodos especializados se debe de contar con equipo sofisticado que es muy costoso y difícil de operar. Por lo que para llevar a cabo el diagnóstico fitosanitario es necesario contar con conocimientos avanzados en la materia, esto conlleva que se reduzca la cantidad de personas capacitadas que lo pueden realizar.

Por otro lado, la visión artificial ha sido utilizada en el sector agropecuario para el reconocimiento de plagas y enfermedades en cultivos, exponiendo buenos resultados. Sin embargo no existen esfuerzos aplicados en el sector forestal de la aplicación de este tipo de técnicas para el reconocimiento de plagas de insectos descortezadores, considerando los repositorios El Sevier, Redalyc, Scopus, IEEE, Springer, Scielo. En este mismo contexto en este trabajo de investigación se desarrollará un sistema experto con visión artificial que permita reconocer especies de insectos descortezadores.

Con base en lo anterior surge la siguiente pregunta de investigación:

¿De qué manera un sistema experto con visión artificial puede contribuir en la detección oportuna de plagas de insectos descortezadores de los géneros *Dendroctonus* e *Ips*?

Los resultados emitidos por el sistema experto serán evaluados a fin de verificar que sean acorde a lo que emiten los expertos del área, en el contexto de esta tesis, lo que permitirá conocer el nivel de contribución de dicho sistema.

3. Justificación

Ciesla et. al. (1976) indican, que la detección de plagas y enfermedades se origina en el descubrimiento de los niveles de actividad de insectos y enfermedades, de tal modo que pueda prevenirse el desarrollo de las principales epidemias, a través de programas de manejo de plagas cuidadosamente planeados y ejecutados.

Considerando el párrafo anterior y la problemática planteada, es importante contar con una herramienta computacional como lo es el sistema experto de fácil manejo en campo que permita la identificación temprana de plagas de insectos descortezadores presentes en las especies de Pino; como punto base en el tratamiento fitosanitario del género a nivel de ecosistema natural o plantación forestal.

El sistema experto apoyará a usuarios inexpertos en el manejo de plagas de insectos descortezadores, en la toma de decisiones respecto a las medidas que se deben emprender para mitigarlas o erradicarlas, evitando así un mal asesoramiento por técnicos que carecen de los conocimientos y experiencia de un experto, disminuyendo la necesidad de depender de un especialista para identificar la plaga que ataca a los pinos en campo.

La importancia de que el SE identifique la plaga de manera temprana, permitirá que el mismo genere un tratamiento fitosanitario adecuado con aplicación de medidas de control y manejo para la conservación de la vegetación de los bosques de pino en México. Cabe destacar, que, la economía del Sector Forestal del país se basa en productos madereros y los obtenidos de las vegetaciones forestales; no obstante para lograr obtener una mejor producción dependen en mayor medida de la preservación de la salud de los bosques.

4. Objetivo General

- Desarrollar un sistema experto con visión artificial que reconozca las especies de insectos descortezadores de los géneros *Dendroctonus* e *Ips*.

5. Objetivos Específicos

- Crear la base de conocimiento sobre los insectos descortezadores de los géneros *Dendroctonus* e *Ips* y sus características que se encuentran presentes en los estados más influyentes en la economía del Sector Forestal de México.

- Elegir las tecnologías y técnicas computacionales que pueden auxiliar en el reconocimiento de plagas de insectos descortezadores.
- Desarrollar el módulo del sistema experto que reconozca la especie de insecto descortezador con base en las características físicas observadas por el usuario.
- Desarrollar el módulo del sistema experto para reconocer con visión artificial las especie de insecto descortezador.
- Evaluar los resultados emitidos por el sistema experto con visión artificial con personal experto en sanidad forestal.

6. Hipótesis

H0: El sistema experto con visión artificial no reconocerá las especies de insectos descortezadores de los géneros *Dendroctonus* e *Ips*.

H1: El sistema experto con visión artificial reconocerá las especies de insectos descortezadores de los géneros *Dendroctonus* e *Ips*.

7. Metodología

La naturaleza de esta investigación es mixta de tipo descriptiva, debido a que se recolectará información que ayude a describir como se realiza la identificación de plagas de insectos descortezadores en los pinos por los expertos en el área con el acercamiento a las propiedades, características y rasgos importantes a considerar para proporcionar un diagnóstico certero, mediante técnicas de recolección y análisis de datos de investigación cualitativa y cuantitativa.

Por un lado se tienen las técnicas cualitativas logrando con éstas una mayor profundidad y comprensión en las respuestas y fenómeno de estudio respectivamente, permitiendo flexibilidad en la aplicación de los instrumentos y favoreciendo la generación de un vínculo con los sujetos participantes en el estudio. Por otro lado, las técnicas cuantitativas proporcionarán la medición y documentación con instrumentos que permitirán describir las variables de la investigación y explicar sus cambios y movimientos. Con ambas técnicas se analizará: el fenómeno, los procesos, las herramientas y el nivel de conocimiento o experiencia de los participantes para identificar las plagas de insectos descortezadores que atacan a los pinos. La información obtenida del análisis facilitará alimentar la base de conocimiento del SE para que esta pueda dar un diagnóstico certero de las plagas de insectos descortezadores que atacan a los pinos.

Derivado de lo anterior, se requiere obtener una muestra de los expertos en sanidad forestal que puedan apoyar en el trabajo de investigación aportando conocimiento del área.

La muestra de expertos en el área se obtuvo aplicando el muestreo no probabilístico intencional o de conveniencia; se seleccionó este tipo de muestreo debido a que la población de expertos con la que se podría contar para poder llevar la investigación es de fácil acceso por la cercanía donde residen.

Para llevar a cabo esta investigación se utilizará un método inductivo-deductivo para realizar el análisis y descripción partiendo de lo particular a lo general y viceversa, por medio de las siguientes actividades:

1. Recolección de datos
 - a. Selección de técnicas de recolección de datos
 - i. Técnicas cuantitativas
 1. Cuestionario
 2. Entrevista estructurada/semi estructurada
 3. Datos secundarios: implica la revisión de documentos, registros públicos y archivos físicos y electrónicos.
 - ii. Técnicas cualitativas
 1. Observación participante.
 2. Entrevista en profundidad.
 - b. Elaboración de los instrumentos de levantamiento de datos
 - i. Cuestionario.
 - c. Evaluación del instrumento de levantamiento de datos
 - i. Validez por juicio de experto
 - d. Aplicación de los instrumentos de levantamiento de datos
2. Análisis de datos
 - a. Reducción de datos
 - b. Transformación de datos
 - c. Análisis de contenido
3. Base de conocimiento de insectos descortezadores, generada del análisis de los datos.
4. Desarrollo del SE
 - a. Diseño del sistema
 - i. Especificación de los requerimientos del sistema
 - ii. Diseño de lo base de conocimiento
 - iii. Diseño de la arquitectura: relación entre cada elemento del sistema
 - iv. Diseño de la interfaz
 - b. Desarrollo del sistema

- i. Programación
 - 1. Base de conocimientos
 - 2. Interfaz de usuario
- c. Pruebas
 - i. Prueba de usabilidad.
 - ii. Pruebas de desempeño.
 - iii. Pruebas de fiabilidad.
- d. Mantenimiento
 - i. Correctivo: para corregir defectos encontrados durante la ejecución del sistema.
- e. Documentación
 - i. Del sistema
 - ii. Para usuarios

8. Marco Teórico

8.1. Economía del Sector Forestal

México es considerado como un país “megadiverso”, es de las naciones que cuenta con la mayor cantidad y diversidad de animales y plantas. Un ejemplo es el género *Pinus*, que de acuerdo con Gernandt & Pérez de la Rosa (2014) existen 120 especies de este género en el mundo y en México se pueden encontrar 49 especies equivalentes al 40% de las especies del mundo, su distribución altitudinal va de los 500 hasta los 4000 metros (Mirov, 1967), en México se encuentran desde la costa de Baja California hasta los límites arbóreos de los volcanes más altos del Eje Volcánico Transversal en el sur-centro del país. Tienen un alto valor económico porque son fuentes de la producción forestal, además de ofrecer servicios ambientales e influir en la regulación del clima regional (Sánchez, 2008).

La base de la producción del sector forestal se divide en dos categorías: maderable y no maderable. La primera de ellas está constituida por materiales leñosos (escuadría, celulosa, chapa y triplay; postes, pilotes y morillos; leña, carbón y durmientes), la segunda categoría está compuesta por actividades de recolección (resinas, fibras, gomas, ceras, rizomas, tierra de monte y otros) (SEMARNAT, 2017) (SEMARNAT, 2016) (SEMARNAT, 2015).

La producción maderable en 2016 tuvo un incremento en comparación con el año anterior del 9.7% reportando 6.7 millones de metros cúbicos en rollo (m³r). Los principales estados productores fueron: Durango (35.1%), Chihuahua (13.2%),

Veracruz (7.8%), Michoacán (6.7%) y Oaxaca (5.9%), que en conjunto contribuyeron con el 68.6% de la producción total. Los productos principales que se obtuvieron en 2016 fueron: madera para aserrío (escuadría y durmientes) con el 74.9% de la producción (5 millones de m³r), combustibles (leña y carbón) con el 9.4% (630.9 mil m³r) y el resto 15.7% (1.1 millones m³r) destinados a celulósicos, chapa y triplay, postes, pilotes y morillos. El valor total de la producción maderable fue de \$8,847,842,787.00 pesos (SEMARNAT, 2017).

La producción forestal maderable a nivel nacional se puede clasificar por género, dentro de los cuales se tiene que la producción a base de Pino fue de 5,040,049 m³r equivalente 76.7%. con un valor de \$7,266,111,524.00 pesos, el encino con 737,741 m³r aportando el 9.7% del total de la producción del año con un valor de \$545,608,397.00 pesos, y comunes tropicales con un total de 389,670 m³r con un valor de \$537,276,321.00 pesos.

Es por ello que mantener la salud de nuestros bosques es de suma importancia, ya sea para incrementar los servicios ambientales que proporcionan o para obtener una mayor producción forestal.

8.2. Insectos Descortezadores

El incremento de infestaciones por plagas y enfermedades se está convirtiendo en un problema público por sus afectaciones ambientales, sociales y económicas. Los impactos destacados que provocan son el incremento en la mortalidad de masas forestales, cambios en la diversidad de los ecosistemas, y pérdida en la cantidad y calidad de los productos y servicios ambientales, implicando procesos de recuperación que duran años. Las plagas y enfermedades forestales son factores biológicos determinantes de la productividad forestal (Castellanos Bolaños, Martínez Ruiz, Cárdenas Gómez, & Cubas González, 2013) por que disminuyen el valor comercial de la madera y disminuyen las superficies forestales.

De acuerdo con las estadísticas del Anuario de Producción Forestal, en 2016 los insectos descortezadores son la principal plaga que afecta la producción de madera, reportando en ese año el 96.7% del volumen afectado con un total de 12,868 hectáreas en todo el territorio nacional (SEMARNAT, 2017).

Los insectos descortezadores habitan de forma natural en los bosques teniendo una interacción biológica con los pinos, a nivel mundial México es el país que presenta mayor diversidad tanto de especies de pino como de escarabajos descortezadores. La función natural de los descortezadores en la mecánica de los bosques es servir como elemento depurador, donde solo suelen atacar y matar árboles adultos dañados, enfermos o viejos, de este modo se abre espacio en el dosel, provocando

la regeneración del bosque y manteniéndolo saludable y productivo. Las poblaciones de descortezadores suelen estar controladas por los árboles a través de defensas para repeler el ataque; por el clima debido a que son muy sensibles a los cambios de temperatura; por sus depredadores; y otros controles biológicos. Pero cuando presentan fluctuaciones poblacionales se les considera plaga debido a que su ataque es masivo infestando árboles de todas las edades ocasionándoles daños mecánicos o fisiológicos, causando debilitamiento, disminución de crecimiento o, incluso la muerte (Durán & Poloni, 2015).

En México se tienen detectados 12 especies de descortezadores del género *Dendroctonus*, de los cuales varios de ellos son considerados de importancia económica al grado de considerarlas como las plagas forestales más relevantes. El género *Dendroctonus* está asociado con plagas secundarias de descortezadores del género "*Ips*" detectándose en el territorio nacional 10 especies (Cibrian Tovar D. , Méndez Montiel, Campos Bolaños, Yates III, & Flores Lara, 1995), algunas de ellas también son consideradas plagas de importancia por si solas; su distribución abarca desde el norte hasta el sur del país (Castellanos Bolaños, Martínez Ruiz, Cárdenas Gómez, & Cubas González, 2013).

En la presente investigación solo se trabajó con 8 especies del género *Dendroctonus* (*mexicanus* Hopkins, *adjunctus* Blandford, *frontalis* Zimmermann, *brevicomis* LeConte, *approximatus* Hopkins, *parallelocolis* Chapuis, *valen* LeConte y *rhizophagus* Thomas & Bright), ver Tabla 1, y 9 del género *Ips* (*mexicanus* Hopkins, *bonanseai* Hopkins, *pini*, *integer*, *lecontei* Swaine, *confusus*, *cribricollis*, *grandicollis* y *calligraphus*), ver Tabla 2.

8.2.1. Género *Dendroctonus*

Las especies de este género son reconocidas como plagas forestales más dañinas del país, teniendo una división entre especies primarias y secundarias; las primarias son aquellas que inician la colonización de árboles susceptibles y son capaces de vencer la resistencia del árbol, mientras que las especies secundarias son aquellas que arriban a los árboles una vez que estos han sido colonizados por algún descortezador primario y rara vez provocan la muerte del árbol. Su distribución está estrechamente correspondida con las especies arbóreas que colonizan además de su intervalo de altitud en el que se desarrollan, siendo estos aspectos de importancia para realizar su identificación (Salinas, Vargas, Zuñiga, Victor, Ager, & Hayes, 2010).

Tabla 1. Descortezadores del género *Dendroctonus*, tamaño y huéspedes que habitan

Descortezadores		Tamaño	Especies de Pinus
Primarias	<i>D. mexicanus</i> Hopkins	2.3 a 4.5 mm	<i>P. arizonica</i> , <i>P. ayacahuite</i> , <i>P. cembroides</i> , <i>P. chihuahuana</i> , <i>P. durangensis</i> , <i>P. lawsonii</i> , <i>P. leiophylla</i> , <i>P. maximinoi</i> , <i>P. michoacana</i> , <i>P. montezumae</i> , <i>P. pseudostrobus</i> , <i>P. teocote</i>
	<i>D. adjunctus</i> Blandford	2.9 a 6.9 mm	<i>P. arizonica</i> , <i>P. ayacahuite</i> , <i>P. chihuahuana</i> , <i>P. durangensis</i> , <i>P. lawsonii</i> , <i>P. maximinoi</i> , <i>P. michoacana</i> , <i>P. montezumae</i> , <i>P. pseudostrobus</i>
	<i>D. frontalis</i> Zimmermann	2.2 a 3.2 mm	<i>P. arizonica</i> , <i>P. durangensis</i> , <i>P. maximinoi</i> , <i>P. oocarpa</i> , <i>P. pringlei</i> , <i>P. teocote</i>
	<i>D. valens</i> Le Conte	5.7 a 10 mm	<i>P. arizonica</i> , <i>P. ayacahuite</i> , <i>P. lawsonii</i> , <i>P. leiophylla</i> , <i>P. maximinoi</i> , <i>P. michoacana</i> , <i>P. montezumae</i> , <i>P. oocarpa</i> , <i>P. pseudostrobus</i> , <i>P. teocote</i>
Secundarias	<i>D. brevicomis</i> LeConte	2.5 a 4.8 mm	<i>P. arizonica</i> , <i>P. durangensis</i>
	<i>D. approximatus</i> Hopkins	4.5 a 7.0 mm	<i>P. ayacahuite</i> , <i>P. chihuahuana</i> , <i>P. durangensis</i> , <i>P. leiophylla</i> , <i>P. michoacana</i> , <i>P. montezumae</i> , <i>P. oocarpa</i> , <i>P. pseudostrobus</i> , <i>P. teocote</i>
	<i>D. parallelocolis</i> Chapuis	4.9 a 6.3 mm	<i>P. leiophylla</i> , <i>P. oocarpa</i> , <i>P. pringlei</i>
	<i>D. rhizophagus</i> Thomas & Bright	6.0 a 8.2 mm	<i>P. arizonica</i> , <i>P. ayacahuite</i> , <i>P. cembroides</i> , <i>P. chihuahuana</i> , <i>P. durangensis</i> , <i>P. leiophylla</i> , <i>P. oocarpa</i> , <i>P. teocote</i>

Fuente: Elaboración propia

La distribución de este género en América se extiende desde el oeste de Canadá hasta Nicaragua. En México se localiza en los principales sistemas montañosos, desde la Sierra de Baja California, Sierra Madre Occidental, Sierra Madre Oriental, Faja Volcánica Transmexicana, Sierra Madre del Sur, hasta la Sierra de Chiapas. Los ambientes en los que viven de acuerdo a su distribución son templados y templado-fríos, las altitudes a las que se encuentran oscilan entre 600 y 4000 msnm. (Salinas, Vargas, Zuñiga, Victor, Ager, & Hayes, 2010).

8.2.2. Género *Ips*

Este género se asocia principalmente a especies de pino y se localiza en todas las sierras Mexicanas en donde se distribuyen sus hospederos. Se clasifica generalmente como un descortezador secundario, que solo se establece en arbolado débil, muerto en pie o que han sido derribados; aunque en los últimos años se han venido comportando como especies primarias en algunas regiones.

Tabla 2. Especies del género *Ips*, tamaño y huéspedes que habitan.

Descortezadores	Tamaño	Especies de Pinus
<i>Ips mexicanus</i> Hopkins	3.6 a 5.0 mm	<i>P. ayacahuite</i> , <i>P. durangensis</i> , <i>P. leiophylla</i> , <i>P. michoacana</i> , <i>P. montezumae</i>
<i>Ips bonanseai</i> Hopkins	2.9 a 3.8 mm	<i>P. arizonica</i> , <i>P. ayacahuite</i> , <i>P. cembroides</i> , <i>P. chihuahuana</i> , <i>P. durangensis</i> , <i>P. leiophylla</i> , <i>P. montezumae</i> , <i>P. oocarpa</i> , <i>P. pseudostrobus</i>
<i>Ips pini</i>	3.3 a 4.3 mm	<i>P. arizonica</i> , <i>P. durangensis</i>
<i>Ips integer</i>	4.6 a 5.7 mm	<i>P. arizonica</i> , <i>P. chihuahuana</i> , <i>P. durangensis</i> , <i>P. leiophylla</i> , <i>P. maximinoi</i> , <i>P. montezumae</i> , <i>P. pseudostrobus</i> , <i>P. teocote</i>
<i>Ips lecontei</i> Swaine	4.0 a 4.7 mm	<i>P. arizonica</i> , <i>P. ayacahuite</i> , <i>P. durangensis</i> , <i>P. leiophylla</i> , <i>P.</i> <i>montezumae</i> , <i>P. oocarpa</i> , <i>P. pseudostrobus</i>
<i>Ips confusus</i>	3.5 a 4.2 mm	<i>P. cembroides</i> , <i>P. oocarpa</i>
<i>Ips cribricollis</i>	2.9 a 3.6 mm	<i>P. durangensis</i> , <i>P. lawsonii</i> , <i>P. montezumae</i> , <i>P. oocarpa</i> , <i>P.</i> <i>pseudostrobus</i> , <i>P. teocote</i>
<i>Ips grandicollis</i>	2.8 a 4.6 mm	<i>P. arizonica</i> , <i>P. durangensis</i> , <i>P. leiophylla</i> , <i>P. montezumae</i> , <i>P. oocarpa</i> , <i>P. pseudostrobus</i> , <i>P. teocote</i>
<i>Ips calligraphus</i>	3.9 a 5.9 mm	<i>P. maximinoi</i> , <i>P. michoacana</i> , <i>P. montezumae</i> , <i>P. oocarpa</i> , <i>P. pseudostrobus</i>

Fuente: Elaboración propia.

8.3. Materiales y métodos

Parte importante de esta investigación es identificar a través de una fotografía tomada directamente en campo al insecto descortezador que se encuentre en un árbol, para lograr esto es indispensable tener un dispositivo que permita tomar una fotografía que visualice las características físicas que lo diferencian de otras especies con nitidez y con un acercamiento potente, con una resolución mínima de 640x480 pixeles.

Para poder observar a los insectos descortezadores, en campo al momento de tomar la muestra de insectos estos deberán ser colocados en un frasco con alcohol al 70%, para que el insecto muera y sea más fácil su manipulación.

En los laboratorios de entomología del Departamento de Ciencias Forestales de la Universidad Autónoma Chapingo, utilizan microscopios estereoscópicos leica zoom 2000, con aumento de 10.5x a 45x, para poder observar los insectos y determinar la especie de la que se trata. Para el propósito de esta investigación llevar este tipo de dispositivos a campo no es viable porque no permiten obtener una fotografía de lo que se observa y además tiene un costo elevado.

Derivado de lo anterior se realizaron pruebas con diferentes dispositivos, desde teléfonos celulares de gama media y alta de distintas marcas, cámaras digitales y microscopios digitales. Las pruebas se realizaron con ejemplares de diferentes especies de descortezadores del género *Dendroctonus* e *Ips*, prestadas del material de enseñanza del Departamento de Ciencias Forestales de la Universidad Autónoma Chapingo. En la Tabla 3 se describen las características de los dispositivos probados.

Tabla 3. Listado de dispositivos probados.

Dispositivo		Características			
Tipo	Modelo	Megapíxeles	Apertura	Zoom	Lente
Teléfonos Celulares	Moto G4 Plus	16	f/2.0	4X	SE ¹
	Iphone 4	5	f/2.8	SE ¹	30 mm
	Iphone 5s	8	f/2.4	SE ¹	33 mm
	Iphone 6	8	f/2.2	SE ¹	33 mm
	Xperia Z5	23	f/2.0	8X	24 mm
Cámaras digitales	Sony DSC	7	f/2.0	6X	50 mm
	Rebel T5	18	f/3.5-5.6	5.6X	70 mm
Microscopios digitales	U500X	1.3	SE ¹	5X	55 mm
	Celestron 44306	2	SE ¹	20x	SE ¹

Fuente: Elaboración propia.

Nota: ¹ SE - Sin Especificación

De los anteriores descritos los microscopios digitales, las cámaras digitales y teléfonos celulares no alcanzan a enfocar el insecto por si solos, ninguno de los dispositivos probados permitió obtener la calidad de fotografía deseada, la idea de probar este tipo de dispositivos era encontrar uno que sea compacto, fácil de transportar, ligero y que tenga la capacidad de calidad fotografías mínima requerida tomando la fotografía directamente en campo.

Debido a lo anterior se decidió utilizar el teléfono celular Moto G4 Plus montado en el ocular de un Microscopio Estereoscópico marca leica zoom 2000, con aumento de 10.5x a 45x (Figura 3) y para la iluminación se utilizaron 2 lámparas de luz led de 120 watts. En total se tomaron 6 fotografías en diferentes posiciones de cada insecto (frontal, lateral, cabeza, pronoto, declive y dorsal), con una calidad de 6 megapíxeles, con una resolución de 1836 x 3264 pixeles y un peso de 2 MB por imagen aproximadamente.

Se fotografiaron 4 especies del género *Dendroctonus*: *adjuntus*, *frontalis*, *mexicanus* y *valens*, con un total de 9, 5, 4 y 6 insectos por especie respectivamente; y 5

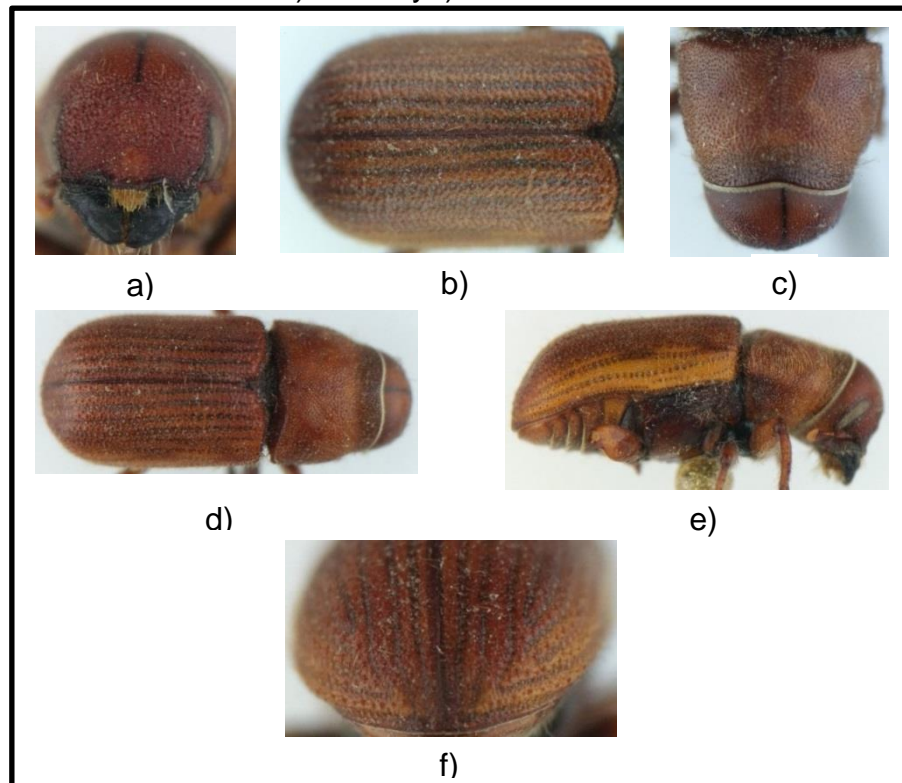
especies del género *Ips*: *bonanseai*, *calligraphus*, *grandicollis*, *integer* y *mexicanus*, con 3, 4, 3, 6 y 6 insectos por especie respectivamente.

Figura 3. Montaje de celular en estereoscopio



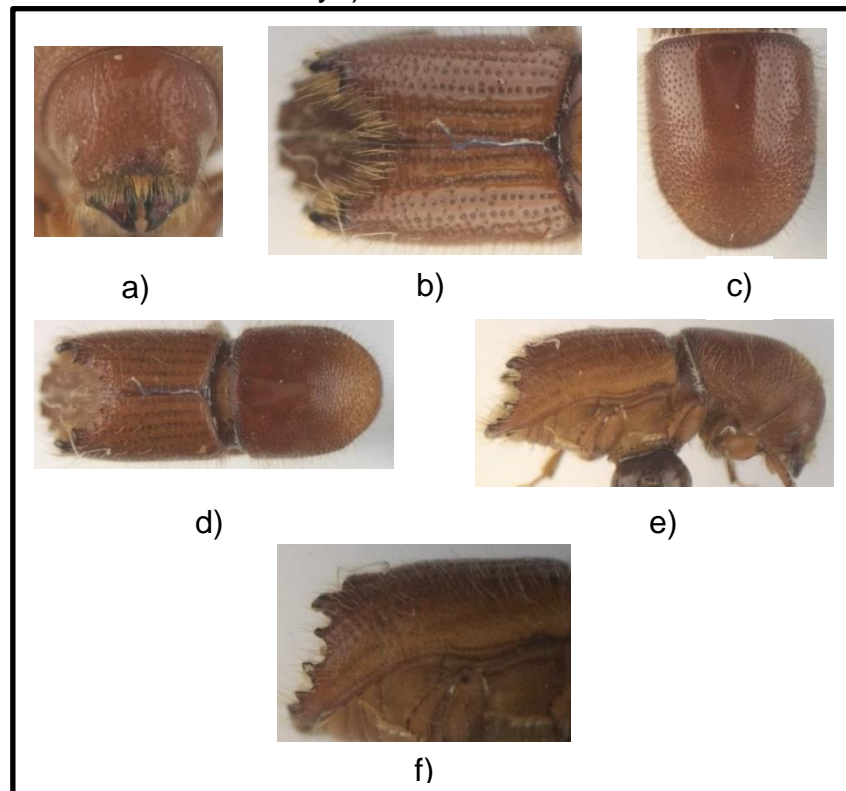
Fuente: Elaboración propia

Figura 4. Fotografías *Dendroctonus valens*: a) cabeza, b) elitra, c) pronoto, d) cuerpo, e) lateral y f) declive elitral



Fuente: Elaboración propia

Figura 5. Fotografías Ips integer: a) cabeza, b) elitro, c) pronoto, d) cuerpo, e) lateral y f) declive elitral



Fuente: Elaboración propia

8.4. Herramientas Computacionales

Derivado de la finalidad, del presente trabajo, de identificar las especies de insectos descortezadores del género *Dendroctonus* e *Ips* que atacan a las especies de pino que se distribuyen en México, se evaluaron tres formas que sirven para realizar esta labor, dentro de los cuales se consideraron: a) los Sistemas Expertos, donde se generó la representación de conocimiento sobre las especies de insectos; b) técnicas de IA para generar un modelo de conocimiento a partir de las características morfológicas observadas por usuarios; y c) la visión por computadora para identificar por medio de una imagen la especie de insecto descortezador.

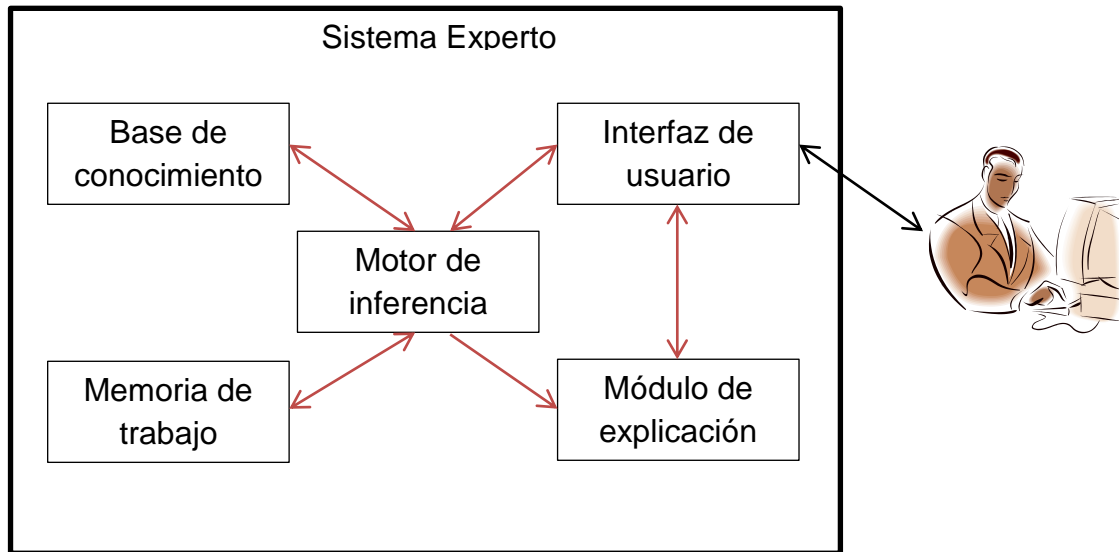
8.4.1. Sistema Experto

Los Sistemas Expertos (SE) forman parte de la Inteligencia Artificial debido a que emplea conocimiento humano codificado en una computadora para resolver problemas sobre algún dominio que normalmente requieren expertos humanos, a grandes rasgos cuando los sistemas están bien desarrollados imitan el

comportamiento que un humano experto ejecutaría para tomar decisiones en una específica y acotada área de pericia.

La arquitectura básica de un sistema experto se puede ver en Figura 6.

Figura 6. Arquitectura de un SE.



Fuente: Elaboración propia.

Existen diferentes tipos de SE, los más conocidos son los basados en reglas su funcionamiento básicamente está dado por la aplicación de reglas, comparación de resultados y la aplicación de nuevas reglas modificadas, pudiendo también trabajar con inferencia lógica con encadenamiento hacia adelante o hacia atrás. Otros SE son los basados en casos, que soluciona nuevos problemas basándose en soluciones de problemas pasados; SE basados en redes bayesianas, que es un modelo gráfico probabilístico que representa a través de un gráfico a cíclico dirigido un conjunto de variables aleatorias y sus dependencias condicionales; y los SE difusos utilizan la lógica difusa, que trabaja con incertidumbre, esto es debido a que la toma de decisiones no siempre es verdadero o falso, a veces utiliza el término “quizás”.

La estructura básica de un SE considera los siguientes componentes:

- Base de conocimiento: contiene la acumulación, transferencia y transformación del conocimiento del dominio particular, es la experiencia para resolver problemas
- Memoria de trabajo: contiene los hechos sobre el problema que se desea tratar.
- Motor de inferencia: esencialmente realiza el razonamiento de la información en la base de conocimiento.

- Módulo de explicación: explica el comportamiento del SE cuando encuentra una solución.
- Interfaz de usuario

Una de las partes fundamentales dentro de los SE es el desarrollo de la Base de Conocimiento que incluye los elementos fundamentales para comprender, formular y resolver problemas, básicamente contiene la representación del conocimiento sobre el dominio particular (Pacheco) (Torres) (Badaro, Ibañez, & Agüero, 2013).

Para el caso particular de este trabajo se utilizó la representación de conocimiento por medio de reglas debido a que de esta forma se puede plasmar de mejor manera las claves dicotómicas para identificar las especies de insectos descortezadores.

8.4.1.1. Metodología del SE

Existen fases básicas definidas en el desarrollo del SE, a continuación se describen (Durkin, 1994):

- Evaluación del problema

Se enfoca en determinar si el SE puede ser una solución del problema, si es factible, si el problema requiere de decisión humana, si se maneja conocimiento heurístico, si está delimitado el dominio, además de saber si hay otros SE que se hayan aplicado al mismo problema.

- Adquisición del conocimiento

Es un proceso cíclico, de vital importancia para la creación del SE que puede resultar difícil, si no es tratado con cuidado puede ocasionar pérdida de conocimiento, generar conocimientos irrelevantes o incompletos lo que puede generar una base de conocimiento inconsistente. Esta consiste de:

- Recolección: obtener información del experto, así como de diversas fuentes de información.
- Análisis: estudiar las piezas claves de la información recolectada.
- Interpretación: detectar las piezas clave de la información recolectada.
- Diseño: preparación para las sesiones con los expertos, para explorar conceptos.
- Diseño del sistema

Esta fase inicia cuando se entiende la magnitud del problema. Partiendo de la selección de la técnica de representación de conocimiento. Posteriormente se debe seleccionar la herramienta para construir el SE, construyendo un prototipo que pueda ser validado por los expertos sirviendo como guía para el resto del proyecto. Por

último se encuentra el desarrollo el cual se debe refinar para cumplir con los objetivos del proyecto.

- Pruebas y evaluación

Las pruebas y evaluación se deben de realizar de forma periódica durante el desarrollo del SE para confirmar o rectificar el avance del SE. La evaluación se relaciona con la validación del sistema y la aceptación del usuario. Tanto la validación y aceptación del usuario son subjetivas lo que lo convierte en una tarea compleja.

- Documentación

Es importante documentar el conocimiento, gráficos del conocimiento, código fuente, pruebas y el glosario de términos específicos del dominio del problema. Con lo anterior se busca que la entrada de nuevo conocimiento así como la modificación resulte una tarea fácil.

- Mantenimiento

Debido a que el SE puede estar sujeto a cambios, carencias o dificultades, se debe tener consideraciones para facilitar la tarea de mantenimiento como:

- Documentación completa y organizada.
- Portabilidad del SE.
- Considerar una estructura modular

8.4.1.2. Representación de conocimiento con reglas

Hace referencia a la lógica de primer orden para representar y resolver problemas creando enunciados declarativos, que tratándose de lenguaje natural, pueden ser verdaderas, o bien falsas. Lo que se busca es describir una situación en términos de objetos y relaciones entre ellos para obtener conclusiones aplicando reglas de razonamiento. Existiendo dos mecanismos para inferir conocimiento a partir de un conjunto de reglas: 1) el razonamiento hacia atrás (comienza con la definición de las metas) y 2) hacia adelante (se inicia definiendo los datos iniciales) (Guerra, 2011).

En el mercado existen diversas herramientas para construir SE por medio de reglas, como por ejemplo: CLIPS, JESS, Drools, JEOPs, OpenCyc, Prolog, por mencionar algunos. En este trabajo se seleccionó Prolog para la representación del conocimiento con reglas.

8.4.1.3. PROLOG

PROLOG utiliza el paradigma de programación basado en los principios de la lógica de predicados, el programador especifica relaciones entre valores de datos y luego plantea consultas al entorno de ejecución para ver si ciertas relaciones se mantienen. Un programa lógico a través de hechos y reglas explícitos, define una base de conocimiento a partir del cual se puede extraer conocimiento implícito. PROLOG tiene similitudes con los lenguajes de programación funcional, la diferencia radica en la forma en que trabajan, en los funcionales se definen funciones, y en PROLOG se definen relaciones. Otra gran diferencia se presenta en el motor de ejecución, en el lenguaje funcional se evalúa una expresión convirtiéndola en un gráfico a cíclico y luego se reduce el gráfico a una forma normal que representa el valor calculado. Por otra parte PROLOG deduce una respuesta a partir de las relaciones dadas, en lugar de recibir una expresión para evaluar, recibe una expresión que interpreta como una pregunta.

La composición de PROLOG está dada por: hechos que declaran relaciones verdaderas entre objetos; variables instanciadas o no, que representan objetos; reglas que expresan que un hecho depende de uno o más hechos, representándose con una implicación lógica del tipo $p \rightarrow q$ (p implica q); y operadores para operaciones matemáticas básicas, utilizando infijos ó prefijos, también se manejan operadores de igualdad y desigualdad (Penagos, 2004) (Morales, 2017).

Su selección se derivó de la aplicación que ha tenido en una gran cantidad de campos de forma satisfactoria, además de contar con una sintaxis y semántica simple interpretando gramáticas en lenguaje natural, donde su mecanismo de resolución de problemas genera objetivos y subobjetivos buscando relaciones entre los objetos creados, variables y listas, realizando preguntas (objetivos) de forma interactiva donde el intérprete generará por inferencia los resultados que se deducen a partir del contenido en la base de conocimiento (Rauch, 1989).

También por su metodología de diseño top-down que permite descomponer el problema en subproblemas, resolviendo estos últimos primero (bottom-up) lo que permite una mejor depuración ya que se puede ser probado inmediatamente y ser corregido si lo necesita.

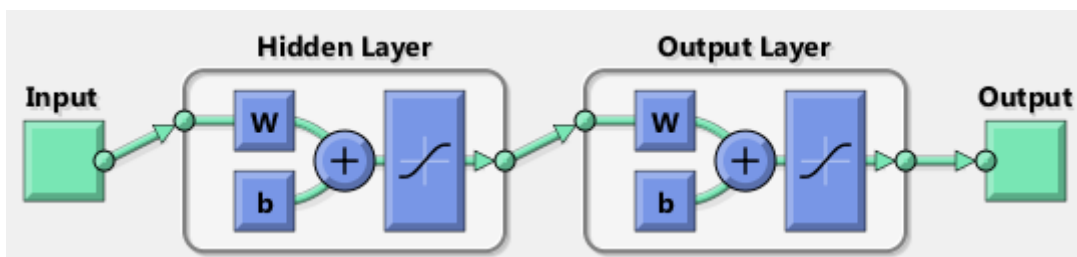
8.4.2. Técnicas de IA

En este apartado se detallan las tres técnicas que se utilizaron para clasificar los insectos: Red Neuronal Feed-Fordward Backpropagation, Red Bayesiana y Máquinas de Soporte Vectorial.

8.4.2.1. Red Neuronal Feed-Forward Backpropagation (RNFF)

Las Redes Neuronales Feed-Forward están conformadas por al menos 3 capas de neuronas: una capa de entrada, al menos una capa oculta intermedia y una capa de salida; las neuronas son conectadas hacia adelante, esto quiere decir que las neuronas de entrada se conectan a las neuronas de la capa oculta y esta a su vez se conecta a las neuronas de la capa de salida (Figura 7). El método de entrenamiento Backpropagation es utilizado en este tipo de red para adaptar los pesos de conexión entre neuronas a partir del cálculo del error comparando la salida obtenida con la esperada, lo que permite adquirir nuevos conocimientos (Amardeep, 2017).

Figura 7. Estructura Red Neuronal Feed-Forward



Fuente: Elaboración propia.

La RNFF utiliza una capa denominada feedforward donde realiza una correlación entre cada uno de los patrones prototipo y el patrón de entrada. Para realizar esta correlación, las filas de la matriz de pesos de la capa feedforward, representada por la matriz de conexión W^1 , son un conjunto de patrones prototipo (Hagan, Demuth, & Beale, 1995).

$$W^1 = \begin{bmatrix} P_1^T \\ P_2^T \end{bmatrix} \quad (1)$$

Asimismo utiliza una función de transferencia lineal, y cada elemento del vector de bias es igual a R , donde R es el número de elementos en el vector de entrada.

$$b^1 = \begin{bmatrix} R \\ R \end{bmatrix} \quad (2)$$

La salida de la capa feedforward con la matriz de pesos y vector de bias, es:

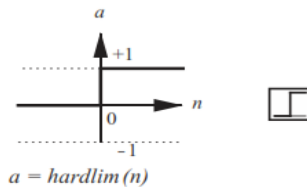
$$a^1 = W^1 p + b^1 = \begin{bmatrix} P_1^T \\ P_2^T \end{bmatrix} P + \begin{bmatrix} R \\ R \end{bmatrix} = \begin{bmatrix} P_1^T P + R \\ P_2^T P + R \end{bmatrix} \quad (3)$$

Como se puede observar en la ecuación anterior, las salidas de la capa feedforward son iguales a los productos internos de cada patrón prototipo con la entrada, mas R.

Las salidas de la red dependen de la función de transferencia o activación seleccionada, los parámetros w y b son ajustados por alguna regla de aprendizaje para que la relación de entrada-salida de la neurona cumpla con un objetivo específico (Hagan, Demuth, & Beale, 1995). Las funciones de transferencia comúnmente utilizadas son:

- Función de Transferencia Hard Limit (Figura 8): Arroja 0 como salida de la neurona si el argumento de la función es menor que 0, o 1 si su argumento es mayor o igual a 0.

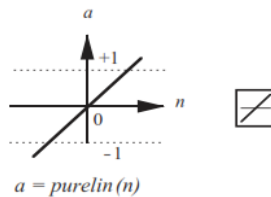
Figura 8. Función de Transferencia Hard Limit



Fuente: (Hagan, Demuth, & Beale, 1995)

- Función de Transferencia Lineal (Figura 9): La salida es igual a su entrada.

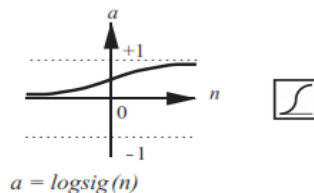
Figura 9. Función de Transferencia Lineal



Fuente: (Hagan, Demuth, & Beale, 1995).

- Función de Transferencia Log-Sigmoid (Figura 10): Toma la entrada y ajusta la salida en un rango entre 0 y 1, utilizando la expresión: $a = \frac{1}{1+e^{-n}}$

Figura 10. Función de Transferencia Log-Sigmoid

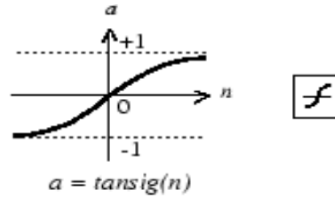


Fuente: (Hagan, Demuth, & Beale, 1995).

- Función de Transferencia Hyperbolic Tangent Sigmoid (Figura 11): Toma la entrada y ajusta la salida en un rango entre -1 y 1, utilizando la expresión:

$$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$$

Figura 11. Función de Transferencia Hyperbolic Tangent Sigmoid



Fuente: (Hagan, Demuth, & Beale, 1995).

La RNFF usa el método Backpropagation para el entrenamiento, el cual tiene tres pasos (Castillo, Cobo, Gutiérrez, & Pruneda González, 1999) (Freeman & Skapura, 1993) (Hilera & Martínez, 2000). El primero paso es propagar la entrada de la red hacia adelante hasta llegar a la capa de salida:

$$\begin{aligned} a^0 &= P, \\ a^{m+1} &= f^{m+1}(W^{m+1}a^m + b^{m+1}) \text{ para } m = 0, 1, \dots, M - 1 \\ a &= a^M \end{aligned} \quad (4)$$

El segundo paso es propagar la sensibilidad hacia atrás a través de la red, obteniendo el error de la salida de la red con el valor esperado:

$$\begin{aligned} s^M &= -2F^M(n^M)(t - a) \\ s^m &= F^m(n^m)(W^{m+1})^T s^{m+1}, \text{ para } m = M - 1, \dots, 2, 1 \end{aligned} \quad (5)$$

El tercer paso es ajustar los pesos y bias capa por capa hasta llegar a la primera, realizando una retropropagación del error con la finalidad de que todas las neuronas de la red hayan recibido un error que describa su aportación relativa al error total. Con el valor del error se modifican los pesos de las conexiones de la red para cada patrón de aprendizaje, con el propósito de que en la siguiente ocasión que se presente el mismo patrón, la salida esté más cerca a la deseada.

Para realizar el ajuste se utilizan las siguientes ecuaciones:

$$W^m(k + 1) = W^m(k) - \alpha s^m (a^{m-1})^T \quad (6)$$

$$b^m(k + 1) = b^m(k) - \alpha s^m \quad (7)$$

8.4.2.2. Clasificador Bayesiano (CB)

Este tipo de clasificadores modelan un fenómeno mediante un conjunto de variables y sus relaciones de dependencia entre ellas, permitiendo estimar la probabilidad posterior de variables con base en variables conocidas. Aprenden de un conjunto de entrenamiento la probabilidad condicional de cada atributo o característica X_1, X_2, \dots, X_n dada la etiqueta de una de m clases posibles C_1, C_2, \dots, C_n , tal que la probabilidad de la clase dados los atributos se maximiza:

$$\text{Arg}_c[\text{Max}P(C|X_1, X_2, \dots, X_n)] \quad (8)$$

La clasificación se realiza aplicando la regla de Bayes para calcular la probabilidad de C dada la instancia particular de X_1, \dots, X_n , y luego predice la clase con la probabilidad más alta:

$$P(C|X_1, X_2, \dots, X_n) = \frac{P(C)P(X_1, X_2, \dots, X_n|C)}{P(X_1, X_2, \dots, X_n|C)} \quad (9)$$

Escribiendo de forma más compacta, considerando $X = \{X_1, X_2, \dots, X_n\}$, se puede expresar como:

$$\text{Arg}_c \left[\text{Max} \left[P(C|X) = \frac{P(C)P(X|C)}{P(X)} \right] \right] \quad (10)$$

Para resolver un problema de clasificación, se requiere calcular la probabilidad a priori de cada clase, $P(C)$, y la verosimilitud de cada atributo dada la clase, $P(X | C)$; para obtener la probabilidad a posteriori $P(C | A)$ (Castrillón, Giraldo, & Sarache, 2008) (Friedman, Geiger, & Goldszmidt, 1997) (Sucar, Clasificadores Bayesianos: de Datos a Conceptos) (Sucar, Redes Bayesianas).

8.4.2.3. Máquina de Soporte Vectorial (SVM)

Las Máquinas de Soporte Vectorial son un sistema para entrenar de forma eficiente máquinas de aprendizaje lineal en hiper-planos de características inducidas por kernels, permitiendo construir modelos plasmando vectores de entrada en un espacio multidimensional, construyendo una superficie de decisión lineal llamada vector de soporte que permite separar las clases aplicando funciones de kernel, permitiendo predecir si un punto pertenece a una clase. Para realizar una clasificación, la SVM debe de obtener un grupo de datos pertenecientes a la clase, llamado de entrenamiento, este grupo debe contener valores que pertenecen a la clase (1) y valores que no pertenecen a la clase (-1). Cuando se tiene el grupo de entrenamiento este se transforma a un espacio de n dimensiones, buscando el hiper-plano de

separación que maximice el margen de separación entre la clase positiva y negativa, respetando los puntos de vista de la teoría de generalización y explotando la teoría de optimización (Cortes & Vapnik, 1995) (Ebrahimi, Khoshtaghaza, Minaei, & Jamshidi, 2017). Algunos de los kernel más utilizados son:

- Grado polinomial P^{th}

$$K(x_i, x_j) = (x_i \cdot x_j + 1)^p \quad (11)$$

- Función Base Radial

$$K(x_i, x_j) = e^{-\| \frac{x_i - x_j}{2\sigma^2} \|^2} \quad (12)$$

- Perceptron multi capa (MLP)

$$K(x_i, x_j) = \tanh(\beta x_i \cdot x_j + \delta) \quad (13)$$

Las funciones aprendidas por las máquinas kernel se pueden representar como combinaciones lineales de núcleos calculadas en los puntos de entrenamiento (Cortes & Vapnik, 1995) (Cristianini & Schölkopf, 2002):

$$f(x) = \sum y_i \alpha_i k(x_i, x) + b \quad (14)$$

La función objetivo fue:

$$W(\alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (15)$$

Donde $i, j \in \{1, \dots, n\}$, sujeto a: $\alpha \geq 0$, $\sum \alpha_i y_i = 0$.

Después de resolver, se encuentra la matriz de α . Por lo tanto, el bias y la función de decisión se describen como:

$$b = \frac{1}{|S|} \sum_i \left[y_i - \sum_j \alpha_j y_j k(x_j, x) \right] \quad (16)$$

$$y = \text{Sing} \left(\sum y_i \alpha_i k(x_i, x) + b \right) \quad (17)$$

Donde x_i , y_i , k son puntos de entrenamiento, S describe los valores que no son cero de la matriz α , mientras que x es la matriz de entrada. El signo de la ecuación determina la clase de los datos de entrada.

8.4.3. Evaluación de los modelos

Para llevar a cabo la evaluación de los modelos obtenidos por cada técnica antes mencionada, se utilizó la matriz de confusión para calcular el índice kappa y el porcentaje de precisión general. Además en el caso de la SVM se aplicó una validación cruzada de 20 veces.

8.4.3.1. Índice kappa

El índice kappa es una medida de concordancia entre observadores, siendo la diferencia entre la exactitud lograda por un clasificador automático y por un clasificador aleatorio, se calcula a partir de la matriz de confusión como sigue (Congalton & Green, 2009):

$$K = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} \cdot x_{+i})} \quad (18)$$

Dónde:

r = número de filas en la matriz.

x_{ij} = número de observaciones en el elemento de la fila i y la columna j.

x_{i+} = total de observaciones en la fila i (total marginal a la derecha de la matriz).

x_{+i} = total de observaciones en la columna i (total marginal al pie de la matriz).

N= número total de observaciones incluidas en la matriz.

El valor obtenido de K variará entre 0 y 1, siendo 1 el valor óptimo. El grado de concordancia según el resultado de kappa se mide como en la Tabla 4, determinando el grado de acuerdo entre los observadores.

Tabla 4. Coeficiente kappa

Kappa	Grado de acuerdo
< 0	Sin acuerdo
0 - 0.2	Insignificante
0.2 – 0.4	Bajo
0.4 – 0.6	Moderado
0.6 – 0.8	Bueno
0.8 – 1	Muy bueno

Fuente: (Congalton & Green, 2009).

8.4.4. Técnicas de procesamiento de imágenes

La visión por computadora es una rama de la inteligencia artificial que busca simular computacionalmente la forma en que los seres humanos perciben los objetos por medio del sentido de la vista, las técnicas utilizadas para lograr el reconocimiento de los objetos observados han mejorado a tal grado que al día de hoy existen muchas aplicaciones que utilizan la visión artificial para realizar labores que normalmente humanos desempeñan logrando una reducción de errores y aumentando la efectividad.

Se utilizaron diferentes técnicas y métodos para mejorar las imágenes, a continuación se describen.

8.4.4.1. Redimensión de imágenes

Debido a que el tamaño de las imágenes segmentadas manualmente no es uniforme, al momento de cargar las imágenes se redimensionaron a un tamaño de 640 x 480 píxeles, para que todas manejaran las mismas dimensiones al momento de realizar el procesamiento por lote para posteriores cálculos.

Para lograrlo se utilizó el comando de MATLAB:

`imresize("nombre imagen",[640,480]).`

8.4.4.2. Filtro Gaussiano

Es un filtro lineal que pertenece a la clase de filtros de suavizado especificados por el valor de su matriz de coeficientes existiendo un sinnúmero de combinaciones, realiza dos convoluciones unidimensionales; una en sentido horizontal y sentido vertical, produciendo un suavizado más uniforme.

$$G(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (19)$$

Donde la desviación estándar σ representa el radio de cobertura de la función de Gauss (Cuevas, Zaldívar, & Pérez, 2010).

Tabla 5. Ejemplo de matriz de coeficientes

1	2	1
2	4	2
1	2	1

Fuente: Elaboración propia.

8.4.4.3. Transformación a escala de grises

En forma sencilla se calcula el promedio de los tres planos de color, tal que:

$$E_p(x, y) = \frac{R(x, y) + G(x, y) + B(x, y)}{3} \quad (20)$$

Debido a cambios de iluminación, si el valor del componente rojo o verde es alto se tendrá una apariencia oscura, caso contrario si el componente azul es alto la apariencia será más clara. Para tratar de solventar esto se considera calcular una combinación lineal de todos los planos:

$$E_p(x, y) = W_R R(x, y) + W_G G(x, y) + W_B B(x, y) \quad (21)$$

Donde W_R , W_G y W_B son los coeficientes que definen la transformación, considerándose como:

$$W_R = 0.299 \quad W_G = 0.587 \quad W_B = 0.114 \quad (22)$$

MATLAB utiliza la función (21) con los coeficientes (22) a través del comando: `rgb2gray(RGB)`, devolviendo una matriz bidimensional del tamaño de la imagen con el nuevo valor de cada pixel en escala de grises (Cuevas, Zaldívar, & Pérez, 2010).

8.4.4.4. Ecuación de histograma

El histograma puede ser considerado como medidas estadísticas de la imagen, describiendo la distribución de frecuencias de intensidades (pixeles) de la imagen. Son mejor entendidos en imágenes a escala de grises debido a que en una imagen $I(u, v)$ con intensidades en el intervalo $[0, K-1]$ contendrá el histograma H exactamente K valores diferentes, considerando una imagen de 8 bits $H=2^8 = 256$ intensidades. Cada valor de $h(i)$ será igual al número de pixeles de I con el valor de intensidad i para todos los valores $0 \leq i < K$.

Para lograr una mejor distribución de todos los niveles de intensidad de un histograma se realiza la ecualización del histograma modificando mediante una operación de pixel la imagen. La operación de pixel que se requiere $f_{eq}(p)$ para equilibrar el histograma de una imagen se calcula a partir de su histograma acumulativo que es una variante del histograma normal, definido como:

$$H(i) = \sum_{j=0}^i h(j) \quad \text{para } 0 \leq i < K \quad (23)$$

Donde el valor de $H(i)$ es la suma de todos los valores abajo del valor especificado i del histograma normal $h(j)$ con los valores $j=0 \dots i$ o bien , el obtenido considerando el valor inmediato anterior

$$H(i) = \begin{cases} h(0) & \text{para } i = 0 \\ H(K-1) + h(i) & \text{para } 0 \leq i < K \end{cases} \quad (24)$$

Con lo anterior es posible desplazar las líneas del histograma de tal forma que el histograma acumulativo muestre de manera aproximada una función lineal creciente.

Para una imagen de resolución $M \times N$ pixeles en el intervalo de $[0 \dots K-1]$ quedaría la operación de pixel definida como:

$$f_{eq}(p) = \left[H(p) \cdot \frac{K-1}{MN} \right] \quad (25)$$

La función (25) es monótona creciente debido a que también lo es la función del histograma acumulativo $H(p)$ y a que los parámetros K , M y N son constantes (Cuevas, Zaldívar, & Pérez, 2010).

MATLAB posee una función capaz de realizar la ecualización, llamada `histeq` la cual permite equilibrar el histograma aumentando el contraste mediante la transformación de los niveles de intensidad de la imagen.

8.4.4.5. Erosión

La erosión es el resultado de comprobar si el elemento estructurante b está completamente incluido dentro del conjunto f . Supone tomar el valor mínimo de la imagen en el entorno de vecindad definido por el elemento estructurante, definiéndose como:

$$(f \ominus b)(s, t) = \min\{f(s+x, t+y) : (s+x, t+y) \in D_f, (x, y) \in D_b\} \quad (26)$$

Donde D_f y D_b son los dominios de f y b respectivamente.

Siendo una operación morfológica, su utilidad consiste en definir una geometría determinada al elemento estructurante y pasarlo sobre la imagen. Los objetos menores al elemento estructurante no aparecerán en la imagen resultante. Los objetos que queden de la transformación habrán sido degradados. Esta operación es definida en la notación de conjuntos de puntos como:

$$I \ominus H = \{(x, y) = (x, y) | (x+i, y+j) \in P_I, (i, j), \forall (i, j) \in P_H\} \quad (27)$$

Como lo muestra la expresión (27), para cada punto (x, y) de la imagen, el resultado se compone por los puntos (x, y) para los cuales todos los posibles valores $(x+i, y+j)$ se encuentran en I (Cuevas, Zaldívar, & Pérez, 2010).

En MATLAB el proceso de erosión puede ser calculado con la función `imerode`("Imagen", "estructura de referencia"), la estructura de referencia se define con la función `strel`, donde dependiendo la forma se tienen que colocar parámetros para su construcción como se muestra en la Tabla 6 (Cuevas, Zaldívar, & Pérez, 2010).

Tabla 6. Estructuras de referencia

Estructura	Descripción
strel('diamond', r)	Estructura en forma de diamante, donde r es la distancia del extremo de la figura al centro.
strel('disk', r)	Estructura de disco, con radio r.
strel('line', long, grad)	Estructura de línea, donde long es la longitud de la línea y grad es la dirección.
strel('octagon', r)	Estructura octagonal, donde r es la distancia del centro a uno de los lados.
strel('square', w)	Estructura cuadrada, donde w es el ancho.
strel('arbitrary', Matriz)	Estructura arbitraria definida por Matriz con elementos ceros y unos.

Fuente: (Cuevas, Zaldívar, & Pérez, 2010).

8.4.4.6. Binarización de la imagen a escala de grises

Es una técnica que realiza un recorrido de la matriz de los valores de cada pixel de una imagen en escala de grises, con el objetivo de reducir cada valor a un valor de los dos posibles, negro (0) y blanco (1), comparando cada pixel con un umbral de sensibilidad (threshold). Por ejemplo valores mayores que el umbral obtienen un valor de 1 y menores 0.

8.4.4.6.1. Otsu

Es una técnica que utiliza la varianza para medir la dispersión de los valores de los niveles de grises, calculando el valor umbral de forma que la dispersión dentro de cada clase sea lo más pequeña posible, pero al mismo tiempo la dispersión sea lo más alta posible entre clases diferentes (Otsu, 1979).

Los pasos para llevarlo a cabo son:

1. Los pixeles se dividen en dos clases C_A y C_B por un umbral en el nivel k .
2. El umbral optimo k^* es aquel que maximiza la varianza entre clases:

$$\sigma_B^2(k^*) = \max_{1 \leq k < L} \sigma_B^2(k) \quad (28)$$

En donde:

$$\sigma_B^2(k) = \frac{[\mu_T \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]} \quad (29)$$

3. Sea L la cantidad de niveles de gris.

- El número de píxeles con nivel de gris i se denota como n_i y el total de píxeles de la imagen por $N = n_1 + n_2 + \dots + n_L$.
- El histograma es normalizado como una distribución de probabilidad:

$$p_i = \frac{n_i}{N}, p_i \geq 0, \sum_{i=1}^L p_i = 1 \quad (30)$$

4. Se calculan las probabilidades y media de la clase.

$$\omega(k) = \sum_{i=1}^k p_i \quad (31)$$

$$\mu(k) = \sum_{i=1}^k i \cdot p_i \quad (32)$$

Dónde:

$$\mu_T = \mu(L) = \sum_{i=1}^L i \cdot p_i \quad (33)$$

Otra forma es encontrar el nivel k que maximice η :

$$\eta(k^*) = \max_{1 \leq k < L} \eta(k) \quad (34)$$

En donde:

$$\eta(k) = \frac{\sigma^2(k)}{\sigma_T^2} \quad (35)$$

Dónde:

$$\sigma^2(k) = \sum_{i=1}^k (i - \mu(k))^2 p_i \quad (36)$$

$$\sigma_T^2 = \sum_{i=1}^L (i - \mu_T)^2 p_i \quad (37)$$

8.4.5. Extracción de Características

8.4.5.1. Análisis de Componentes Principales (ACP)

El ACP permite transformar un número de variables correlacionadas en un número menor de variables no correlacionadas, llamadas componentes principales. Es un método utilizado para reducir dimensionalidad de un conjunto de datos multivariados o identificar nuevas variables subyacentes al conjunto de datos que permitan una mejor interpretación de los mismos. Las nuevas variables son combinaciones lineales de las anteriores y se van construyendo de acuerdo al orden de importancia en cuanto a la variabilidad total de los datos, buscando que los nuevos componentes reúnan la mayor cantidad de información. Los componentes principales se obtienen como sigue:

- Sea $\{\phi_1, \dots, \phi_m\} \subset \mathbb{R}^n$ un conjunto de datos, a partir del cual se calcula la media:

$$\mu = \frac{1}{n} \sum_{i=1}^n \phi_i \quad (38)$$

- Se construye la matriz de covarianza

$$\Omega = \frac{1}{m} \Phi \Phi^T, \text{ donde } \Phi = [\Phi_1 - \mu, \dots, \Phi_m - \mu] \quad (39)$$

- Se calculan los valores propios $\{\lambda_1, \dots, \lambda_m\}$ y vectores propios $\{v_1, \dots, v_m\}$ de la matriz de covarianza. Los vectores propios junto con su correspondiente valor propio se ordenan de mayor a menor y se colocan en la matriz V como columnas. Los vectores propios que forman la matriz V se les da el nombre de componentes principales y forman el espacio propio.
- Se proyectan los vectores iniciales al espacio que forman los componentes principales:

$$\varphi_i = V^T (\phi_i - \mu), i = 1, \dots, m \quad (40)$$

Solamente se consideran vectores propios dependiendo de su respectivo valor propio, donde los valores propios son representativos de la varianza de cada variable, con esto se busca que los valores propios cuya suma aporte un porcentaje significativo de la varianza total de los datos sean seleccionados. Se consideran características cuya varianza acumulada sea mayor o igual que el 95% del total.

Para reconocer un nuevo valor θ , se debe proyectar al espacio:

$$\varphi_{\theta} = V^T(\theta - \mu) \quad (41)$$

La proyección obtenida se compara con los vectores del conjunto de entrenamiento para su clasificación.

8.4.5.2. Momentos Invariantes

Pertencen a los descriptores de forma basados en contorno, siendo un conjunto de siete descriptores que permiten modelar objetos. Se basa en los momentos centrales normalizados de la distribución de coordenadas. Sea $O = \{(x_1, y_1), \dots, (x_n, y_n)\}$ el conjunto de coordenadas de los pixeles que conforman un objeto segmentado, los momentos invariantes se obtienen como (Hu, 1962) (Gonzalez & Woods, 2002):

$$\begin{aligned} H_1 &= \eta_{20} + \eta_{02} \\ H_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ H_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ H_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} - \eta_{03})^2 \\ H_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ H_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ &\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ H_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (42)$$

Donde η_{pq} es el momento centralizado y normalizado obtenido con la siguiente expresión:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{c+1}} \quad (43)$$

Calculándose c y μ_{pq} se calculan de la siguiente forma:

$$c = \frac{(p+q)}{2}, p+q = 2, 3, \dots \quad (44)$$

$$\mu_{pq}(O) = \sum_{(x,y) \in O} (x - \bar{x})^p (y - \bar{y})^q \quad (45)$$

Siendo (\bar{x}, \bar{y}) la coordenada del centroide del objeto, obteniéndose con:

$$\bar{x} = \frac{1}{n} \sum_{(x,y) \in O} x \quad (46)$$

$$\bar{y} = \frac{1}{n} \sum_{(x,y) \in O} y \quad (47)$$

8.4.5.3. Patrones binarios locales (LBP)

Es un operador de textura, etiqueta cada pixel de la imagen analizando su vecindad. Consiste en la comparación de un pixel central con los vecinos, el pixel central es tomado como el umbral con respecto a sus vecinos, al realizar la comparación del pixel central con el vecino se le asigna un valor de uno si el vecino es mayor o igual, en caso contrario se le asigna un valor de cero, a cada resultado del umbral se le asigna un peso de 2^n , donde n depende de la posición del vecino con respecto al pixel central, por último se realiza una suma de los diferentes pesos obteniendo la representación LBP del pixel (Harwood, Ojala, Pietikäinen, Kelman, & Davis, 1993) (Alvarado & Fernández, 2012) (García & Alegre).

Para calcular LBP en una imagen a escala de grises se utiliza la siguiente ecuación:

$$LBP_{P,R} = \sum_{p=0}^{P-1} S(g_p - g_c) 2^p, \quad S(X) = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} \quad (48)$$

Dónde:

P es el número de vecinos que se van a considerar.

R es el tamaño del vecindario.

g_c y g_p son los valores de gris del pixel central y cada uno de los p pixeles del vecindario.

Una vez obtenido el valor LBP para cada uno de los píxeles de la imagen, se obtiene de la imagen, dando como resultado un vector de 256 elementos.

8.4.5.4. Descriptores de Haralick

Este tipo de descriptores son un conjunto de medidas de textura calculados a partir de la matriz de co-ocurrencia. Para calcularlos se debe asumir que la información textural de una imagen está contenida en las relaciones espaciales que se dan entre los distintos niveles de gris de un objeto.

La matriz de co-ocurrencia describe la frecuencia de un nivel de gris que aparece en una relación espacial específica con otro valor de gris dentro del área de una ventana

determinada. Es un resumen de la forma en que los valores de los pixeles ocurren al lado de otro valor en una pequeña ventana. La ventana debe ser cuadrada y con un número impar de pixeles. ,

En su artículo Haralick propone 14 descriptores de textura que a continuación se describen (Haralick, Shanmugan, & Dinstein, 1973).

Notación

$p(i, j)$ (i,j) i-esimo entrada en una matriz de dependencia espacial de tonos de grises normalizada

$p_x(i)$ i-esimo entrada en la matriz de probabilidad marginal obtenida de sumar las columnas de $p(i, j) = \sum_{j=1}^{N_g} P(i, j)$

N_g Cantidad de niveles de grises en la imagen cuantificada

$$\sum_i, \sum_j = \sum_{i=1}^{N_g}, \sum_{j=1}^{N_g} \text{ respectivamente}$$

$$p_y(i) = \sum_{j=1}^{N_g} p(i, j)$$

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ i+j=k}}^{N_g} p(i, j), \quad k = 2, 3, \dots, 2N_g$$

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ |i-j|=k}}^{N_g} p(i, j), \quad k = 0, 1, \dots, N_g - 1$$

1. Momento Angular de Segundo Orden

$$f_1 = \sum_i \sum_j \{p(i, j)\}^2 \quad (49)$$

2. Contraste

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ |i-j|=n}}^{N_g} p(i, j) \right\} \quad (50)$$

3. Correlación

$$f_3 = \frac{\sum_i \sum_j (ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (51)$$

4. Suma de Cuadrados: Varianza

$$f_4 = \sum_i \sum_j (i - \mu)^2 p(i,j) \quad (52)$$

5. Momento de Diferencia Inversa

$$f_5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i,j) \quad (53)$$

6. Suma de promedios

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i) \quad (54)$$

7. Suma de varianza

$$f_7 = \sum_{i=2}^{2N_g} (i - f_8)^2 p_{x+y}(i) \quad (55)$$

8. Suma Entropía:²

$$f_8 = \sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\} \quad (56)$$

9. Entropía

$$f_9 = - \sum_i \sum_j p(i,j) \log(p(i,j)) \quad (57)$$

10. Diferencia de la Varianza

$$f_{10} = \text{varianza de } p_{x-y} \quad (58)$$

11. Diferencia de la Entropía

$$f_{11} = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log\{p_{x-y}(i)\} \quad (59)$$

12., 13. Medidas de Información de Correlación

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}}$$

$$f_{13} = (1 - \exp[-2.0(HXY2 - HXY)])^{1/2}$$

$$HXY = - \sum_i \sum_j p(i, j) \log(p(i, j)) \quad (60)$$

Donde HX y HY son entropías de p_x y p_y , y

$$HXY1 = - \sum_i \sum_j p(i, j) \log\{p_x(i)p_y(i)\}$$

$$HXY2 = - \sum_i \sum_j p_x(i)p_y(j) \log\{p_x(i)p_y(j)\}$$

14. Coeficiente de Correlación Máxima

$$f_{14} = \left(\sum_k \frac{p(i, k)p(j, k)}{p_x(i)p_y(k)} \right)^{1/2} \quad (61)$$

El vector resultante de la fase de extracción de características que contempla características locales como globales servirá para entrenar el clasificador.

9. Marco de Referencia

En este apartado se abordarán técnicas que se utilizan o se podrán utilizar en el sector forestal para realizar la identificación de plagas.

De los primeros acercamientos formales al sujeto de estudio, se llevó a cabo la revisión de literatura que servirá como soporte de esta investigación. La búsqueda de información se realizó en revistas especializadas, memorias de congresos, bases de datos electrónicas y tesis para obtención de grado a nivel maestría y doctorado; se obtuvo una lista de más de 50 documentos, de éstos una tercera parte están enfocados en técnicas de detección de plagas y enfermedades, para efectos del marco referencial se filtraron para obtener los de mayor relevancia considerando como base fundamental la utilización de herramientas computacionales para realizar la identificación de plagas y enfermedades.

Nascimento et al (2016) realizaron un sistema experto llamado ENTOTECA para identificar insectos de importancia económica que atacan las plantaciones comerciales de teka en Brasil, por medio de dispositivos móviles con Android. El sistema se desarrolló con HTML 5, jQuery y PhoneGap, cuenta con dos bases de conocimiento: una de los daños en las plantas de teka causados por 23 especies de

insectos, y la segunda de la organización de reglas jerárquicamente codificadas por agente causal de los daños en las plantas creada con ayuda de dos expertos en entomología. También se construyó simultáneamente una base de datos de imágenes, obtenida por los autores fotografiando los daños en las plantas, buscando información en literatura sobre insectos, colecciones entomológicas y visitando plantaciones comerciales. El resultado de efectividad de los diagnósticos del sistema experto fue del 98.44%.

Asimismo, Shiguihara y Valverde (2009) desarrollaron un sistema experto basado en el modelo probabilístico de redes Bayesianas para diagnosticar plagas y enfermedades del espárrago llamado SEDFE, que utilizando una red bayesiana maneja el conjunto de síntomas y causas (enfermedades), en donde denota una relación directa entre efecto-origen. Manejan dos tipos de probabilidad de los síntomas: la primera es la sensibilidad que se refiere a la probabilidad de correlación entre la aparición de la enfermedad y el síntoma, la segunda se refiere a la especificidad que es la probabilidad de no tener el síntoma cuando no está la enfermedad presente. Para realizar una selección adecuada de síntomas que expliquen acertadamente la aparición de una enfermedad, se tienen que seleccionar síntomas cuya relación de sensibilidad y especificidad sea elevada para una enfermedad o plaga. La inferencia es basada en el algoritmo de paso de mensajes de Kim y Pearl, considera la comunicación bidireccional entre nodos padres (enfermedad o plaga) e hijos (síntomas) y viceversa, lo que permite razonamiento de diagnóstico y predicción. El algoritmo almacena valores parciales recibidos de los padres e hijos permitiendo realizar un pre cálculo de probabilidades. Como resultado obtenido de su investigación el sistema experto logra establecer resultados coherentes de acuerdo a los patrones convencionales de cada patógeno del espárrago y sus manifestaciones.

Kaloudisa, Anastopoulou, Yialouris, Lorentzos y Sideridis (2005) desarrollaron un sistema experto que identifica 43 plagas presentes en especies de árboles (pinos, abetos, castañas) por medio de reglas IF-THEN. La base de conocimiento para diagnosticar las plagas se construyó con ayuda de forestales expertos y fue complementada con literatura relacionada: libros, revistas científicas, fotografías, dibujos de insectos y fotografías de los daños que causan los insectos al árbol. Este sistema da la opción de iniciar la identificación: por características del insecto o por los síntomas observados. Si el usuario opta por características del insecto el sistema iniciara la inferencia preguntando al usuario sobre el insecto (ciclo de vida, color del cuerpo, color de alas, y tamaño, por mencionar algunos). Si decide empezar por los síntomas observados el sistema preguntara al usuario sobre los daños que observa en el árbol. Si el sistema llega a una conclusión se la mostrara al usuario junto con un tratamiento. La evaluación del sistema se realizó en dos fases, verificación y

validación. En la fase de verificación se buscó determinar los posibles errores y garantizar que los resultados del sistema experto fueran correctos, para ello fue probado por dos de los desarrolladores de forma independiente en campo, realizando 86 experimentos. Para la fase de validación se llevó a cabo en tres grupos de usuarios, uno de ellos fue formado por expertos en el área y los dos restantes por usuarios potenciales. Los resultados arrojaron que la utilidad, amigabilidad, facilidad y entendimiento del sistema fue satisfactoria.

Bula, Aramendiz, Salas, Vergara y Villadiego (2012) realizaron un sistema experto llamado SEFI capaz de diagnosticar ocho plagas y nueve enfermedades de los cultivos de berenjena, por medio de un motor de inferencia que toma en cuenta los daños presentes en las diferentes partes de la planta, así como la presencia de algún tipo de insecto. Se desarrolló con la combinación de tecnologías SWI-PROLOG, JAVA, PostgreSQL y XML, trabajando en un ambiente web de tres capas: modelo, vista, controlador. La base de conocimientos se construyó con ayuda de un experto, literatura, artículos y revistas, seleccionando la información relevante para el sistema: fue dividida en un módulo de plagas y otro de enfermedades. Su funcionamiento se basa en una lista de chequeo formulada con la ayuda del experto, donde el usuario selecciona una opción de cada módulo respondiendo si o no, respuestas que son analizadas por el motor de inferencia. El motor de inferencia maneja dos tipos de reglas: la primera de ellas corresponde a plagas y la segunda corresponde a enfermedades, ambas trabajan con las entradas que digite el usuario en las preguntas u opciones de la lista de chequeo, analizando y comparando si coinciden con las reglas del motor de inferencia para cada tipo de regla, para lanzar después un diagnóstico. La evaluación del sistema se realizó con ayuda de expertos, y fue probado por ingenieros y estudiantes de agronomía, considerando la facilidad de uso del sistema, número de aciertos y tiempo de respuesta.

Ochoa y Kemper (2009) presentaron un sistema inteligente capaz de predecir y alertar los niveles de población de una plaga, así mismo realiza un diagnóstico de enfermedades basada en variables obtenidas en tiempo real por una estación Agro Climática, brindando una recomendación de tratamiento adecuado al problema detectado con indicación de productos, dosis, época y forma de aplicación, así como precauciones que se deben de tener en la manipulación de los productos fitosanitarios, enfocándose en los cultivos de jitomate y mango. El sistema cuenta con tres módulos: el módulo de monitoreo y control preventivo de enfermedades, el módulo de diagnóstico y control de enfermedades y el módulo de diagnóstico y control de plagas. Fue desarrollado en Visual C# con tecnología orientada a objetos, cada módulo cuenta con una base de conocimiento que alberga las reglas que permiten la identificación de la plaga o enfermedad y una base de datos que almacena los valores de las variables que forman parte de las reglas, los síntomas y

las fotografías de los síntomas. Como resultado permite determinar plagas y enfermedades con base en las variables climatológicas en cultivos de jitomate y mango, también recomienda inspecciones en campo para prevenir o detectar en etapas temprana el surgimiento de alguna enfermedad o plaga. Cuando proporciona una recomendación el sistema, da la oportunidad de que el productor guarde que producto utilizará para dar seguimiento y crear una bitácora del cultivo.

Es evidente que los sistemas expertos como herramientas computacionales para identificar plagas y enfermedades son los más utilizados debido a que emplean conocimientos humanos para resolver problemas específicos que normalmente requieren de expertos en el área, llegando a imitar el proceso de razonamiento que los expertos utilizan, tomando decisiones en la resolución de problemas.

Otra de las técnicas que hace uso de herramientas computacionales es la utilización de satélites para tomar fotografías como lo indican Blanco, Bonet y Eizaguirre (2009), en su trabajo proponen el uso de imágenes Landsat para detectar escolítidos en rodales de pino, los escolítidos son una plaga insectil que cuando está presente en los pinos provoca signos evidentes, modificando la coloración del follaje pasando de un color amarillento (nivel 1), rojizo (nivel 2) y grisáceo (nivel 3) cuando el árbol se encuentra muerto en pie por el ataque de esta plaga. En su estudio utilizaron 34 grupos de árboles (rodales) considerados como áreas de entrenamiento seleccionados aleatoriamente, 17 de ellos con presencia de escolítidos mostrando niveles de afectación 2 y 3, y los restantes sin presencia de la plaga (sanos) que sirvieron de control. Para evaluar esta técnica utilizaron una imagen que provino del satélite Landsat 5 sensor TM multibanda con una resolución de 30 metros, y con ayuda del programa computacional ERDAS IMAGINE 8.x realizaron la identificación de las áreas de entrenamiento de manera supervisada. Los resultados obtenidos de los niveles digitales (valor numérico de cada pixel obtenido de la radiación recibida por el sensor para una determinada parcela de la superficie terrestre en una determinada banda del espectro) mostraron que existe un patrón similar de comportamiento para los 34 rodales estudiados y las 7 bandas analizadas. Las bandas con las que se obtuvieron mayor resolución espectral a la discriminación de los dos grupos de rodales fueron la TM5 y TM7, en las cuales los 17 rodales afectados por escolítidos quedaron separados en su mayoría de los rodales sanos. Los resultados obtenidos confirman la viabilidad de utilizar imágenes Landsat para la detección de plagas forestales debido a la cantidad de rodales identificados a partir de una imagen satelital.

El uso de este tipo de técnicas resulta costoso, además de que se necesita personal capacitado y con conocimientos avanzados en la materia para poder operar el software con el que hacen el reconocimiento en la imagen satelital. La limitante de esta técnica es que identifica los rodales atacados pero no la especie de escolítido

que lo hace, para hacer esto se tiene que ir a la zona en campo, la ventaja es que con una imagen satelital pueden abarcar una zona extensa.

Por otro lado se cuenta con técnicas de procesamiento de imágenes por computadora, una muestra de ello es el “procesamiento de imágenes para reconocimiento de daños causados por plagas en el cultivo de *Begonia semperflorens* (flor de azúcar)” que realizaron Cáceres C., Amaya D. & Ramos O. (2014), cuya forma de obtener las fotografías se realizó por medio de un dron equipado con una cámara, las imágenes fueron procesadas con filtros morfológicos, difuminado gaussiano y HSL, con el fin de identificar el posible ataque de plagas en el cultivo. Como resultado principal se detectaron perforaciones de hojas ocasionadas por plagas.

Diversos autores han utilizado la visión por computadora para identificar insectos como lo muestra Martineau et al. (2017) realizando una revisión de cuarenta y cuatro artículos, cuyo objetivo fue realizar una revisión de los trabajos existentes para clasificar insectos utilizando imágenes en 2D, centrandó su interés en métodos de procesamiento de imágenes con condiciones de iluminación y disparo normales, adquiridas por cámaras estándar o dispositivos móviles. Se analizaron los trabajos considerando un sistema de imágenes clásico, que puede descomponerse en tres fases: captura de imágenes, extracción de características y clasificación. Una aportación importante de los autores es la comparación de la obtención de la fotografía en laboratorio y en campo, concluyendo que la captura de imágenes más utilizada es en laboratorio donde se tiene un mayor control de las condiciones para tomar la fotografía; y el posicionamiento del insecto se realiza de forma manual con pose restringida, debido a que en campo abierto es más difícil el reconocimiento. Para el caso de la extracción de características se debe aislar el insecto aplicando diferentes métodos de segmentación, para posteriormente obtener características locales o globales que permitan identificar el insecto.

Wang, Lin, Ji, & Liang (2012) realizan una comparativa de dos métodos de reconocimiento de patrones: redes neuronales artificiales y máquinas de soporte vectorial, para identificar insectos de acuerdo al orden al que pertenecen. Trabajaron con 255 imágenes de 10 ordenes; 25 por cada una. Desarrollaron un sistema con dos módulos: proceso de entrenamiento y proceso de reconocimiento, bajo la plataforma .NET versión 4.0, el sistema cuenta con un preprocesamiento de las imágenes donde se cambia el color de fondo por un color no presente en el insecto pudiendo realizar esta tarea de forma manual o de forma automática. Para la extracción de características se hizo una comparación de cada orden de insecto para encontrar las características taxonómicas y seleccionar o crear características que pudieran ser extraídas con el procesamiento de la imagen; se crearon siete características que permiten expresar la estructura del cuerpo del insecto, además

de que éstas se ven poco afectadas por la calidad de la imagen. Asimismo, se creó una interfaz de usuario con la finalidad de corregir los valores obtenidos y evitar errores en el cálculo automático de las características. Para el reconocimiento de patrones se implementaron dos métodos: una red neuronal back-propagation con la función sigmoide; y una C-SVM con RBF, utilizando los métodos de grid-search y crossvalidation (CV) implementándolo con la ayuda de LibSVM tool versión 2.89. Comparando ambos métodos la SVM tuvo más porcentaje de respuesta en comparación con la ANN, dado que en las pruebas realizadas, para el caso de la ANN, se obtuvo una exactitud del 94% con 8 órdenes de insectos donde se excluyó el orden Coleoptera; para el caso de SVM los mejores resultados se obtuvieron con 8 órdenes excluyendo al orden Heteroptera con un porcentaje de exactitud del 98% y de 88% de porcentaje de CV. Los autores también concluyen que es mejor utilizar ANN cuando se requiera realizar pruebas con más características y seleccionar una SVM para implementar el sistema en una aplicación real.

Qing et. al. (2012) desarrollaron un sistema para identificar cuatro especies del género Lepidoptera que atacan los cultivos de arroz, mediante la captura de imágenes de la parte superior e inferior de cada insecto con dos cámaras obteniendo un número mayor de características. Se extrajeron un total de 156 características de color, forma y textura, que sirvieron para entrenar una máquina de soporte vectorial (SVM) con función de núcleo de base radial, utilizando una validación cruzada de siete veces para mejorar la tasa de precisión en la identificación de los insectos. La tasa promedio de precisión fue de 97.5%, lo que se observó es que aplicando la validación cruzada aumentó la precisión en 7% en comparación con una SVM sin validación cruzada.

Larios et. al. (2008) describen un enfoque para la identificación taxonómica automatizada de larvas de mosca de piedra, este tipo de reconocimiento es un desafío porque los insectos presentan muchas articulaciones, exhiben un alto grado de variación intraespecie en tamaño y color, y algunas especies son difíciles de diferenciar visualmente. Las imágenes son adquiridas con la ayuda de un microscopio para crear un ambiente controlado altamente repetible. La clasificación se realiza mediante un proceso que implica: identificación de regiones de interés; representación de esas regiones en vectores SIFT; clasificación de los vectores SIFT en características para formar un histograma; y clasificación del histograma de características a través de algoritmos de clasificación el modelo de mezcla Gaussiana, el algoritmo de maximización de expectación y el árbol de modelo logístico. Se utilizaron tres detectores: detector de región de curvatura principal (PCBR), el detector Hessian-affine y el detector de entropía Kadir. Para evaluar los resultados se aplicó el método de Opelt. Los resultados que presentan utilizando tres detectores y cuatro clases de insectos son del 82%, con tres clases del 95%. El

nuevo método PCBR es eficiente para discriminar entre dos especies visualmente similares logrando una mayor precisión.

Xie et. al. (2015) proponen un modelo donde se extraen características tanto locales (HOG, SIFT y textura) como globales (color y forma) de los insectos con el propósito de identificar 24 especies presentes en cultivos de campo. En la fase de entrenamiento se crea un diccionario con cada una de las características, que posteriormente sirven para reconocer por medio de un sistema de representación de múltiples tareas dispersas y técnicas de aprendizaje de kernel múltiple (MKL). La representación de múltiples tareas dispersas puede combinar múltiples características para mejorar el rendimiento de reconocimiento y el MKL es propuesto para fusionar múltiples características efectivamente. Como resultado se obtuvo un 97.2% de exactitud de categorización.

Shin et. al. (2012) proponen un modelo de 4 niveles para identificar las pisadas de insectos. Para la toma de las huellas se utilizó un túnel de rastreo. Las pisadas se recopilan en tarjetas de seguimiento de color blanco que tienen un lado entintado, los animales son traídos con señuelos para que caminen a través del túnel y dejen sus huellas en la tarjeta de seguimiento. Para lograr la identificación se requiere que las huellas individuales se extraigan y luego se agrupen en patrones de seguimiento. Para la clasificación se propone un fuzzy weight decision de seis capas, donde los pesos son basados en grados de contribución utilizando una función triangular como la de membresía en el intervalo $[0, 1]$. La clasificación se puede realizar de dos formas; la primera de ellas utiliza el modelo de la media ponderada difusa donde se obtiene n resultados de clasificación, iguales o diferentes; la segunda forma se desprende de la primera cuando los resultados son diferentes, se emplea un enfoque de votación por mayoría. Para la evaluación del reconocimiento se utilizaron ocho características, generando dos tipos de patrones, donde cada una de ellas muestra el porcentaje de reconocimiento de la clase y dependiendo de cuál clase obtenga el mayor valor se vota como resultado de la clasificación.

Wen & Guyer (2012) propusieron un método que incluye tres modelos para identificar ocho especies de insectos. El primer modelo se basa en características locales invariables a escalas, rotación, iluminación y cambio de punto de vista, produciendo información local alrededor del punto de interés; para lograrlo utilizaron el detector Hessian-Affine para detectar puntos de interés y sus vecinos, también un vector SIFT de 128 elementos para describir el vecindario local de cada punto de interés, debido a que se pueden generar diversos puntos de interés se utilizó el método de bolsa de palabras (BOW) para convertir las matrices construidas con SIFT por punto de interés en vectores de histograma. El segundo modelo corresponde a la extracción de características globales utilizadas para describir imágenes completas, utilizando el método de segmentación por conglomerados basado en color para extraer el insecto

del fondo obteniendo información del contorno y forma, la segmentación se realizó en el espacio de color HSV, aplicando K-means clustering se clasificó los valores en dos clusters utilizando distancia euclidiana, la salida de K-means es una imagen binaria que separa el insecto del fondo, posteriormente se extraen características: geométricas relacionadas con las propiedades morfológicas del insecto, área, perímetro, excentricidad, relación de circularidad, longitud del eje principal, longitud del eje menor, área convexa, solidez y diámetro de equivalencia; funciones de contorno aplicando el descriptor de Fourier; momentos invariables derivando del segundo y tercer momento; características de textura útiles para el reconocimiento y clasificación de insectos, siendo uno de los índices distintivos por los expertos en entomología, aplicando matriz de co-ocurrencia de nivel de gris (GLCM); características de color, tono medio en el espacio de color HSV, desviación estándar del tono en el espacio HSV, saturación media en el espacio de color HSV, desviación estándar de la saturación en el espacio de color HSV, tono medio en el espacio de color LCH, desviación estándar del tono en el espacio de color LCH saturación media en el espacio de color de HCL, desviación estándar de saturación en el espacio de color de HCL, luminancia media en el espacio de color de LAB y desviación estándar de saturación en el espacio de color de LAB. El tercer modelo llamado clasificación jerárquica que es la combinación de características locales y globales, cuenta con dos niveles, en el primer nivel se seleccionó el mejor clasificador para las características globales donde se encontró un valor umbral de probabilidad de clasificación, si los valores de probabilidad son más altos que el valor del umbral se envían a un clasificador de segundo nivel, en este se selecciona el mejor clasificador de rendimiento para características locales. Se obtuvo una combinación de 54 características globales y 100 locales por imagen, se utilizaron cinco clasificadores, siendo el mínimo lineal (MLSLC), el clasificador lineal basado en densidades normales (NDLC), clasificador de vecino más cercano (KNNC), el clasificador de medias más cercano (NMC) y el árbol de decisión (DT). Para el clasificador basado en características globales usando NDLC dio la mejor precisión con un 96.9%, para las características locales el NMC dio la mejor precisión con el 71.3%. Las características globales brindan descripciones directas y significativas similares a los biólogos, pero son muy sensibles a la iluminación, la rotación y la oclusión. Las características locales superan estas carencias pero son más difíciles de calcular. Los resultados experimentales en la clasificación de imágenes de insectos recogidos en el campo por modelo de combinación lograron la tasa de clasificación del 86.6%.

Diello et. al. (2016) proponen un enfoque para el reconocimiento de enfermedades de soja basado en técnicas de descriptores locales (SURF, DSIFT, SIFT y PHOW) y bolsa de palabras (BOVM) aplicados sobre imágenes en escala de grises, para realizar la clasificación se utilizó una máquina de soporte vectorial que tiene como entrada los vectores BOVM, se evaluó la salida de la clasificación con una validación

cruzada de diez veces y la métrica de tasa de verificación correcta (CCR). En la validación cruzada se dividen en 10 conjuntos lo que garantiza que cada uno contenga la misma proporción de cada clase, luego se usa un conjunto para probar y los nueve restantes sirven para entrenar, esto se repite diez veces con cada conjunto una vez, el CCR corresponde al promedio de la ejecución de las 10 rondas. Los resultados experimentales mostraron que el uso de descriptores locales, junto con la técnica BOVW, fueron efectivos (más del 98% de precisión) y eficientes (un tiempo promedio de 0.1 s por imagen) en la tarea de detectar computacionalmente enfermedades de soja. Estos resultados mostraron que el descriptor PHOW proporcionaba el mejor resultado, seguido de los descriptores SIFT y DSIFT. Se observó que el descriptor PHOW funciona mejor para espacios de color si se compara con la escala de grises, alcanzando tasas de clasificación correctas que van desde 96.25% (± 1.53) a 99.83 (± 0.35).

Sossa et. al. (2005) utilizan siete momento geométricos invariantes de Hu, probando ocho clasificadores basados en distancia: Euclidiana, Manhattan, Manhattan modificado, Chi, Canberra, Coeficiente de correlación, Ángulo y SSE modificado. El desempeño de los clasificadores se comparó con un clasificador Bayesiano. Utilizando 100 imágenes en diferentes posiciones y orientaciones se observó que el mejor desempeño de clasificación correspondió al clasificador Bayesiano combinado con los rasgos descriptores geométricos de Hu dando un porcentaje del 96.75, seguido de los clasificadores basados en la distancia de Manhattan modificada y de Canberra, con un 92% de reconocimiento en ambos casos.

Ebrahimi et. al. (2017) presentaron un sistema para identificar parásitos en invernaderos utilizando fotografías adquiridas con la ayuda de un robot que recorre el cultivo controlado con LabVIEW, montando una cámara en un brazo del robot a una distancia de 80 centímetros con respecto a las flores. Dentro del tratamiento de imágenes se eliminan las regiones que no son flores consideradas como fondo aplicando el operador gamma, actualización de histograma y estiramiento de contraste. Para la detección se utilizó una SVM con la función de kernel base radial RBF, evaluando el rendimiento del modelo con: MSE (error cuadrático medio), RMSE: (raíz del error cuadrático medio), MAE (error absoluto medio) y MPE (error porcentual medio). Demostrando que el procedimiento de procesamiento de imágenes desarrollado es capaz de identificar parásitos en el entorno de invernadero con un error menor al 2.5%.

Hongtao et. al. (2011) propone un método para clasificar 15 especies de insectos en grano almacenado, usando una SVM con un kernel con función de base radial con un porcentaje de precisión del 94.67%, utilizando una clasificación uno contra uno lo que permite transformar una multi clasificación en una clasificación binaria clásica, lo que genero 105 clasificadores binarios.

10. Propuesta de solución

A continuación se describe en forma concisa el área en que el Sistema apoyará a los entomólogos.

10.1. Conceptualización del Sistema

Analizando la manera en que el Sistema pudiera brindar apoyo a los expertos en el área en la tarea de identificar insectos descortezadores del género *Dendroctonus* e *Ips* tanto en las inspecciones de campo como en laboratorio, se concluyó que el SE a desarrollar debería ayudar en:

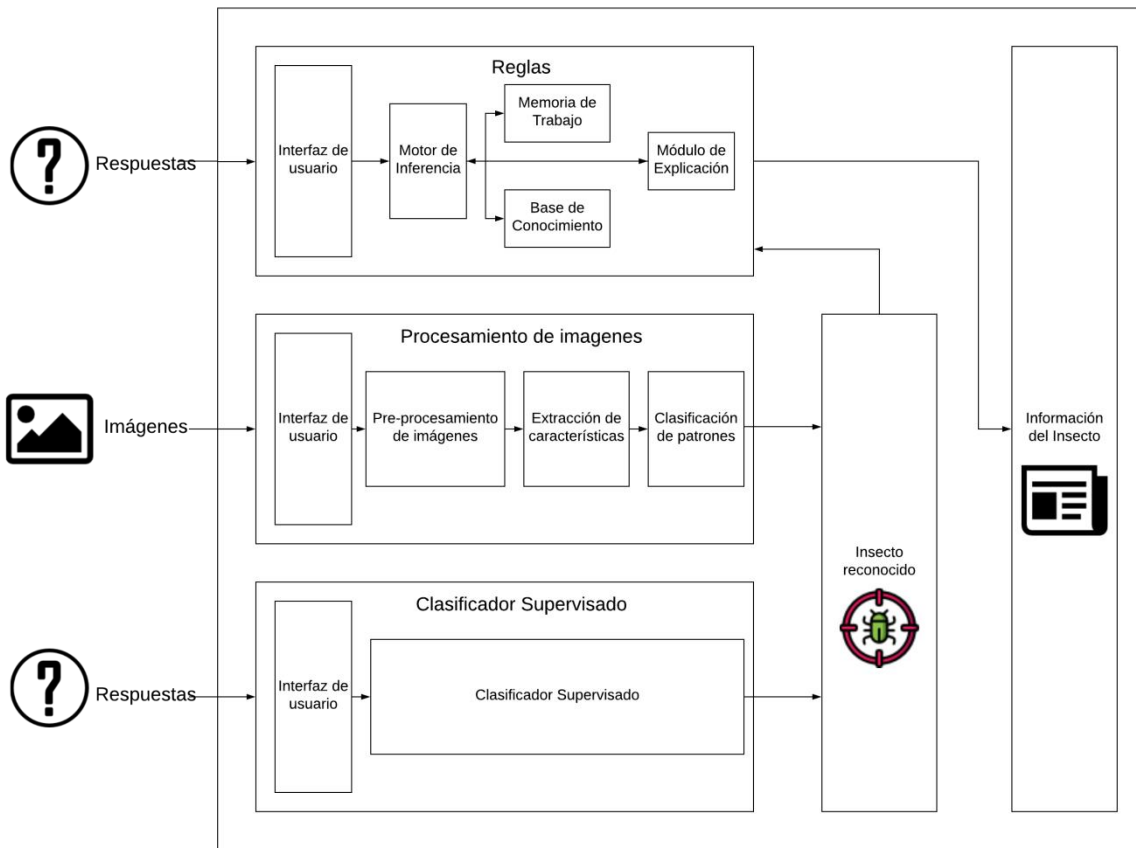
- Identificación de insectos de forma sencilla
- Visualización de características físicas importantes
- Identificación en campo
- Identificación en laboratorio
- Identificación remota del insecto

Para satisfacer los requerimientos mencionados se dividió el Sistema en 3 módulos principales, quedando conformado como sigue:

- Módulo de Reglas: Consta de reglas almacenadas en la base de conocimiento, alimentadas con los hechos que el usuario ingresa para inferir la especie de insecto descortezador.
- Módulo de Procesamiento de Imágenes: Partiendo de imágenes del insecto se realizará un preprocesamiento y una extracción de características para su posterior clasificación. La respuesta que se obtenga de la clasificación si es positiva, esto quiere decir que el insecto fue reconocido, entrará al motor de inferencia del Módulo por Reglas para obtener la información del insecto, en caso de que el insecto no fuera reconocido se iniciarán las preguntas del Módulo por Reglas.
- Módulo de Clasificador Supervisado: A través de preguntas realizadas al usuario se identificara al insecto con ayuda de un clasificador supervisado.

En la Figura 12 se puede apreciar el esquema de la arquitectura del Sistema propuesto.

Figura 12. Esquema de la arquitectura del Sistema



Fuente: Elaboración propia

10.2. Adquisición de conocimiento para el desarrollo del Sistema

Para la realización de este proyecto se contó con la participación de expertos y técnicos en entomología forestal de la Universidad Autónoma Chapingo ubicada en el Municipio de Texcoco, Estado de México.

Como primer acercamiento de adquisición de conocimiento, se realizó una revisión de bibliografía especializada (revistas, libros, claves dicotómicas, etc.) con el afán de familiarizarse con el dominio, previo a las sesiones con los expertos, logrando aprender conceptos que permitieron una comunicación efectiva desde el inicio del proyecto.

Desde la primera sesión con los expertos y técnicos, siempre se tuvo buena disposición para brindar la información y apoyo necesario. Sin embargo, se presentaron obstáculos no relacionados a la voluntad y entrega de los expertos, entre los que se pueden mencionar:

- Diferencia de criterios en las descripciones de los insectos.
- No disponer de los expertos y técnicos a tiempo completo, debido a sus responsabilidades.
- Rehacer modelos, debido a la inclusión o eliminación de variables.

Las sesiones de trabajo con los expertos tenían un tiempo de entre 2 a 4 horas, donde se realizaban entrevistas y discusiones, en donde se definieron los elementos importantes del dominio, además de servir para formular los instrumentos para recabar la información relevante.

10.2.1. Adquisición de conocimiento: Características insectos descortezadores

Se realizaron sesiones con los expertos y técnicos, donde se discutió respecto a las características que diferencian a cada insecto, partiendo de las encontradas en la bibliografía especializada, debido a que como la información provenía de diversas fuentes se debía homologar, además de tomar lo relevante de cada autor para complementar las descripción de cada insecto.

En la Tabla 7 se enlistan las características de cada género de insecto, debido a que aunque son descortezadores la forma de identificar ambos géneros difiere.

Se realizó una selección de variables basándose en la elección de los mejores atributos del problema, cabe agregar que se hicieron dos elecciones derivado que en este trabajo se busca poder realizar la identificación de las especies tanto en campo como en laboratorio, como ayuda a personas inexpertas como expertas en el área. Se consideró en mayor medida las variables (características) que utilizan los expertos en el área para diferenciar a las especies, como se observa en la Tabla 7 se realizó una distinción entre las variables de campo siendo las que se pueden observar sin mayor complicación en el insecto, imagen o fotografía, las variables de laboratorio que requieren de equipo sofisticado o especializado para poder visualizar las características con una mayor definición y precisión, y las variables que no se considerarán para realizar los análisis debido a que no aportan información relevante para identificar a las especies de insecto descortezador.

Por otra parte, se realizó una discretización de los valores de las características para cambiarlos a valores continuos. Cabe señalar que la mayoría de los valores de las características eran más una descripción que un valor numérico, porque es como lo utilizan los expertos para identificar a las especies de insectos, básicamente lo que se hizo fue crear una clave numérica de las descripciones por característica agrupando los valores semejantes.

Tabla 7. Características insectos descortezadores

Dendroctonus	Ips
Tamaño mínimo	Tamaño mínimo
Tamaño máximo	Tamaño máximo
Coloración cabeza	Coloración cabeza
Coloración cuerpo	Coloración cuerpo
Intensidad color	Intensidad color
Tubérculo frontal	Tubérculo frontal
Tubérculo epistomal	Tubérculo epistomal
Vestidura epistémica	Vestidura epistémica
Vestidura pronatal	Vestidura pronatal
Vestidura elitral	Vestidura elitral
Distribución	Distribución
Hospedantes	Hospedantes
Grumos de resina	Grumos de resina
Aserrín en la corteza	Aserrín en la corteza
Muerte del árbol	Muerte del árbol
Lugar de ataque	Lugar de ataque
Área de ataque	Área de ataque
Hongo manchadores	Hongos manchadores
Galerías	Galerías
Generaciones por año	Generaciones por año
Puntuaciones Pronoto	Puntuaciones Pronoto
Forma pronoto	Tubérculo frontal y epistomal unidos
Callo pronatal	Epistoma bífido en el centro
Mazo antenal	Suturas mazo antenal
Setas declive elitral	Espinas declive elitral
Inicio de ataque	Forma tercer espina
Diámetro a la altura de pecho de árboles atacados	Tercer espina capitada
Actividad	Espinas 1 y 2
Forma de ataque	Espinas 2 y 3
Ciclo de vida (días)	Vestidura frontal
Altitud mínima MSNM	
Altitud máxima MSNM	

Fuente: Elaboración propia.

Nota: Las variables de color gris oscuro representan las observadas en laboratorio, las de gris claro las que se pueden observar en campo y las de color blanco son variables que no se considerarán para realizar los análisis.

10.3. Interfaz de usuario del sistema: Menú Principal

En la Figura 13 se puede observar la pantalla principal del sistema, donde se presentan las distintas opciones para realizar la identificación de un insecto: por medio de reglas de inferencia (búsqueda libre, guiada y listado de especies) encerradas en el recuadro color rojo, a través de imágenes del insecto (búsqueda con visión) encerrado en el recuadro color amarillo y con técnicas de IA (búsqueda con técnicas de IA) encerrado de color verde.

Figura 13. Menú Principal del Sistema

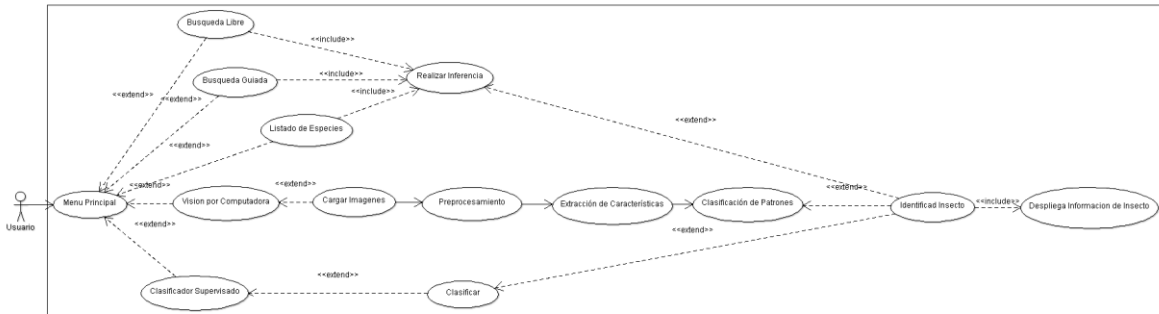


Fuente: Elaboración propia.

10.4. Diagramas de Caso de Uso

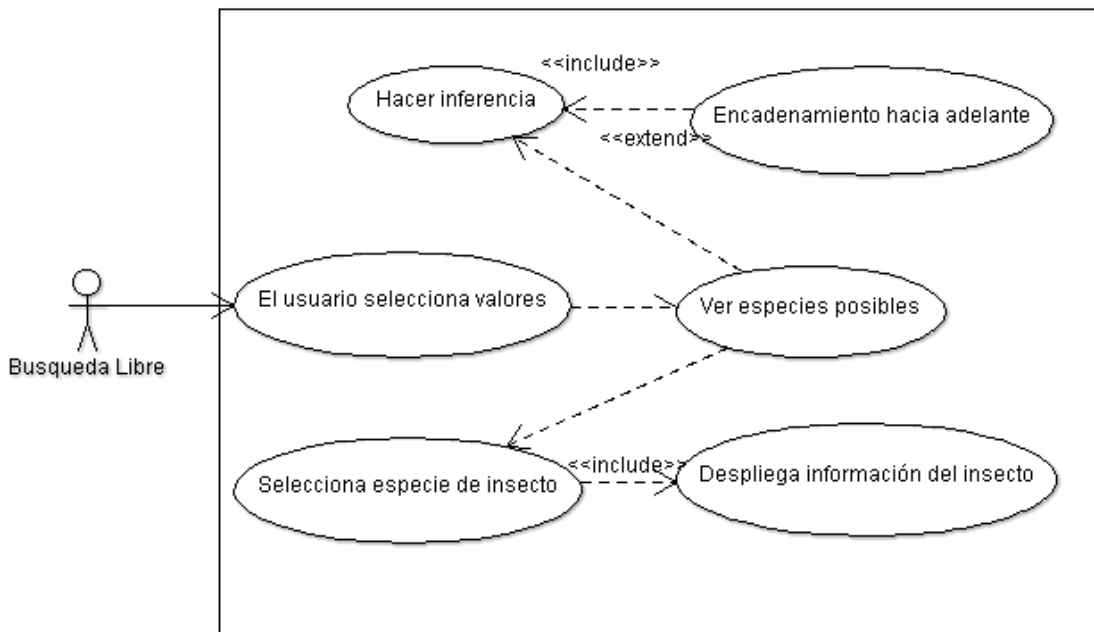
A continuación se describen los diagramas de Casos de Uso del Sistema.

Figura 14. Diagrama de Caso de Uso: Menú Principal



Fuente: Elaboración propia.

Figura 15. Diagrama de Caso de Uso: BUSLI



Fuente: Elaboración propia.

Tabla 8. Caso de Uso: Menú Principal

Nombre del Caso de Uso:	Menú Principal
No. de Caso de Uso:	1
Actores:	Usuario Final, Sistema
Descripción:	Muestra las opciones del sistema para identificar insectos Descortezadores.
Precondiciones:	<ol style="list-style-type: none"> 1. El sistema debe estar instalado en la computadora. 2. El sistema debe estar en la pantalla principal
Poscondiciones:	<ol style="list-style-type: none"> 1. El estado del sistema estará en espera de la selección de una nueva opción por parte del usuario.
Flujo Normal:	<ol style="list-style-type: none"> 1. El sistema despliega las opciones del menú mediante 5 botones: Búsqueda Libre (opción 1), Búsqueda Guiada (opción 2) y Listado de Especies (opción 3), identificación con Visión por Computadora (opción 4) y con Clasificadores Supervisados (opción 5). 2. El usuario selecciona la opción búsqueda libre y se ejecuta el caso de uso BUSLI. 3. Se regresa a 1. 4. El usuario selecciona la opción Búsqueda Guiada y se ejecuta el caso de uso BUSGUI. 5. Se regresa a 1. 6. El usuario selecciona la opción Listado de Especies y se ejecuta el caso de uso LISESP. 7. Se regresa a 1. 8. El usuario selecciona la opción identificación con Visión por Computadora y se ejecuta el caso de uso VISCOMP. 9. Se regresa a 1. 10. El usuario seleccionara la opción identificación con Clasificadores Supervisados se ejecuta el caso de uso CLASUP. 11. Se regresa a 1.
Flujos Alternativos:	<ol style="list-style-type: none"> 1.1 El usuario no elige ninguna opción y oprime el botón de cerrar 1.2 El usuario puede acceder a la opción del menú Descortezadores 1.3 El usuario puede acceder a la opción del menú Acerca de
Incluye:	<ol style="list-style-type: none"> 1. Caso de uso BUSLI 2. Caso de uso BUSGUI 3. Caso de uso LISESP 4. Caso de uso VISCOMP 5. Caso de uso CLASUP
Frecuencia de Uso:	Cada vez que se inicie el sistema.

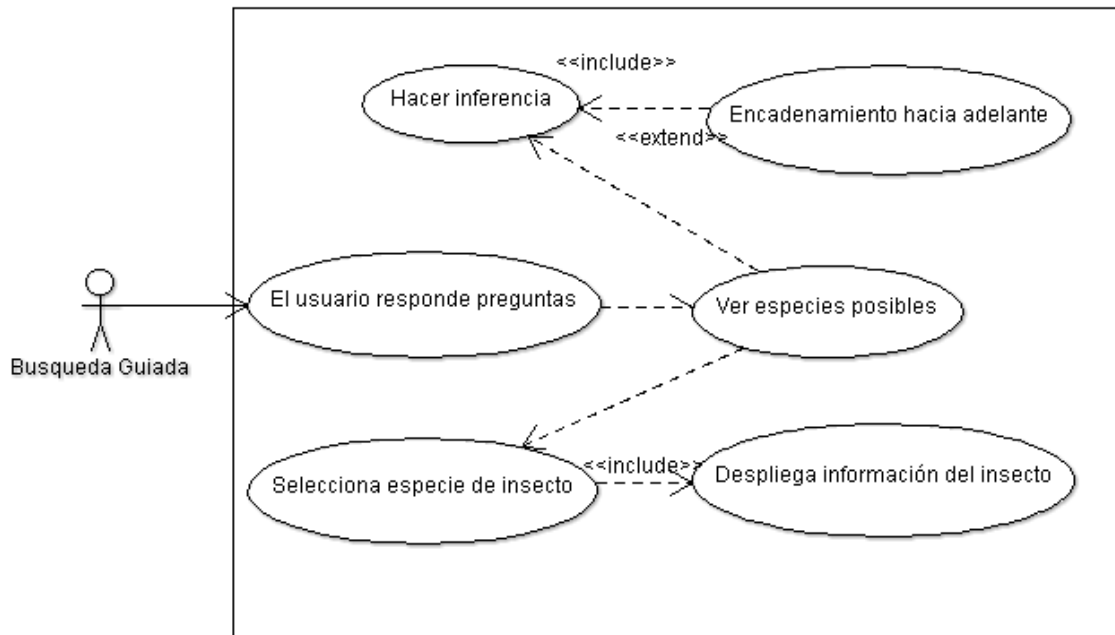
Fuente: Elaboración propia.

Tabla 9. Caso de Uso: BUSLI

Nombre del Caso de Uso:	BUSLI
No. de Caso de Uso:	2
Actores:	Búsqueda Libre
Descripción:	Identificar insectos Descortezadores a través de la selección de valores de las características por parte del usuario
Precondiciones:	El usuario debe encontrarse en la ventana de búsqueda libre.
Poscondiciones:	
Flujo Normal:	<ol style="list-style-type: none"> 1. El sistema mostrará una ventana donde deberá seleccionar los valores de las características de los insectos. 2. Después de seleccionar el valor de las características, dará clic en el botón ver especies posibles. 3. El sistema hará una inferencia con encadenamiento hacia adelante con los valores de las características seleccionadas por el usuario. 4. El sistema mostrará en una tabla las especies de insecto posibles. 5. El usuario podrá dar clic sobre la fila de la tabla con la especie de insecto, el sistema desplegará una ventana con la información del insecto.
Flujos Alternativos:	<ol style="list-style-type: none"> 1.1 El usuario puede acceder a la opción de glosario. 1.2 El usuario puede acceder a la opción de regresar al menú principal. 1.3 El usuario puede acceder a la opción del menú Acerca de. 1.4 El usuario puede acceder a la opción del menú Descortezadores.
Excepciones:	<ol style="list-style-type: none"> 1. Si se presenta error al momento de realizar la inferencia, el sistema informara al usuario y dará la opción de volver a iniciar la identificación.
Incluye:	
Frecuencia de Uso:	Cada vez que se seleccione la búsqueda libre.

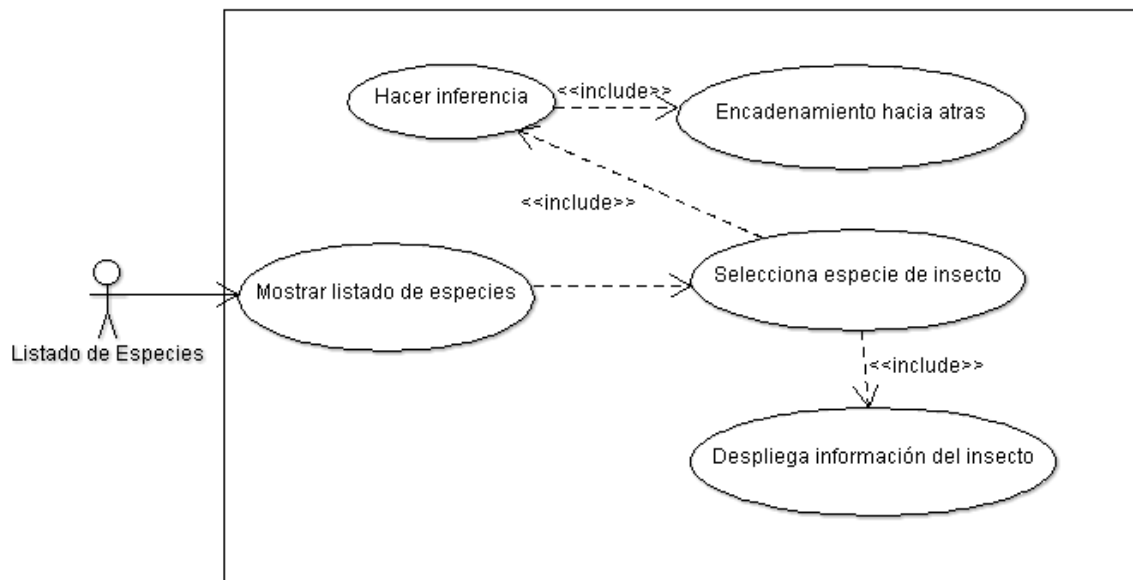
Fuente: Elaboración propia.

Figura 16. Diagrama de Caso de Uso: BUSGUI



Fuente: Elaboración propia.

Figura 17. Diagrama de Caso de Uso: LISESP



Fuente: Elaboración propia.

Tabla 10. Caso de Uso: BUSGUI

Nombre del Caso de Uso:	BUSGUI
No. de Caso de Uso:	3
Actores:	Búsqueda Guiada
Descripción:	Identificar insectos Descortezadores a través de la respuesta de preguntas realizadas al usuario sobre las características físicas del insecto
Precondiciones:	El usuario debe encontrarse en la ventana de búsqueda guiada
Poscondiciones:	
Flujo Normal:	<ol style="list-style-type: none"> 1. El sistema mostrará una ventana con las preguntas donde el usuario deberá responder. 2. Después de responder las preguntas el usuario dará clic en el botón ver especies posibles. 3. El sistema hará una inferencia con encadenamiento hacia adelante con las respuestas del usuario sobre las características del insecto. 4. El sistema mostrará en una tabla las especies de insecto posibles. 5. El usuario podrá dar clic sobre la fila de la tabla con la especie de insecto, el sistema desplegará una ventana con la información del insecto.
Flujos Alternativos:	<ol style="list-style-type: none"> 1.1 El usuario puede acceder a la opción de glosario. 1.2 El usuario puede acceder a la opción de regresar al menú principal. 1.3 El usuario puede acceder a la opción del menú Acerca de. 1.4 El usuario puede acceder a la opción del menú Descortezadores.
Excepciones:	<ol style="list-style-type: none"> 1. Si se presenta error al momento de realizar la inferencia, el sistema informara al usuario y dará la opción de volver a iniciar la identificación.
Incluye:	
Frecuencia de Uso:	Cada vez que se seleccione la búsqueda guiada.

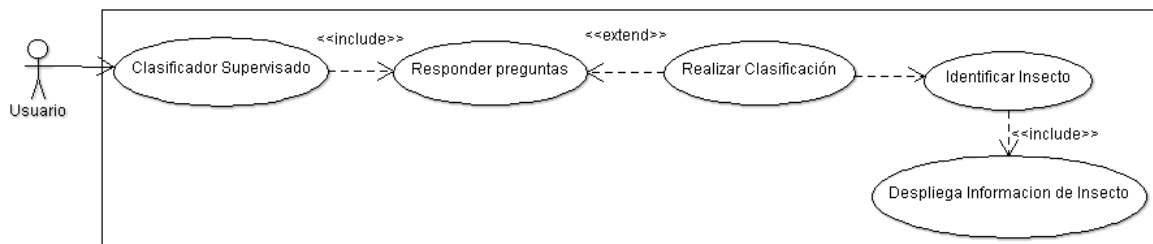
Fuente: Elaboración propia.

Tabla 11. Caso de Uso: LISESP

Nombre del Caso de Uso:	LISESP
No. de Caso de Uso:	4
Actores:	Listado de Especies
Descripción:	Se mostrará un listado con las especies que el sistema puede identificar, para que el usuario pueda ver la información.
Precondiciones:	El usuario debe encontrarse en la ventana de listado de especies
Poscondiciones:	
Flujo Normal:	<ol style="list-style-type: none"> 1. El sistema mostrará una ventana con el listado de especies. 2. El usuario deberá seleccionar una especie, dando clic sobre el nombre. 3. El sistema hará una inferencia con encadenamiento hacia atrás con la especie que el insecto selecciono. 4. El sistema desplegará una ventana con la información del insecto.
Flujos Alternativos:	<ol style="list-style-type: none"> 1.1 El usuario puede acceder a la opción de glosario. 1.2 El usuario puede acceder a la opción de regresar al menú principal. 1.3 El usuario puede acceder a la opción del menú Acerca de. 1.4 El usuario puede acceder a la opción del menú Descortezadores.
Excepciones:	<ol style="list-style-type: none"> 1. Si se presenta error al momento de realizar la inferencia, el sistema informara al usuario y dará la opción de volver a iniciar la identificación.
Incluye:	
Frecuencia de Uso:	Cada vez que se seleccione el listado de especies.

Fuente: Elaboración propia.

Figura 18. Diagrama de Caso de Uso: Clasificador Supervisado



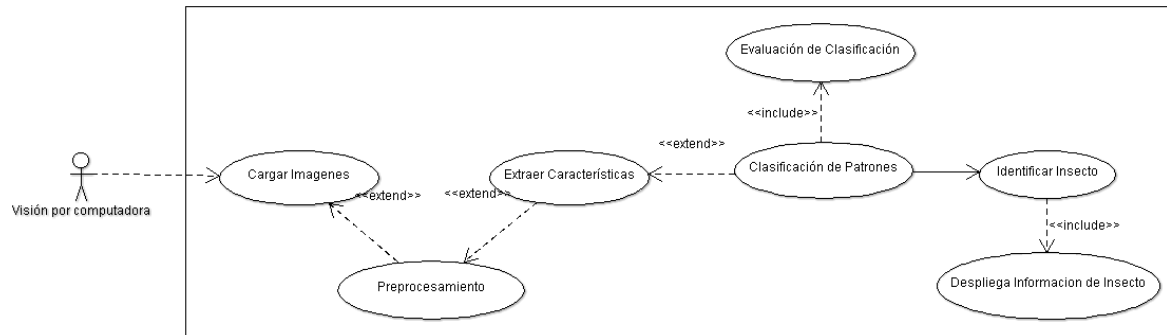
Fuente: Elaboración propia.

Tabla 12. Caso de Uso: Clasificador Supervisado

Nombre del Caso de Uso:	Clasificador Supervisado
No. de Caso de Uso:	5
Actores:	Usuario Final Sistema
Descripción:	Identificar insectos Descortezadores con clasificadores supervisados
Precondiciones:	El usuario debe encontrarse en la ventada de clasificador supervisado
Poscondiciones:	
Flujo Normal:	<ol style="list-style-type: none"> 1. El sistema desplegará una ventana con las preguntas que el usuario deberá responder para iniciar la identificación. 2. El usuario debe responder 6 preguntas que servirán para formar el vector que se enviará al clasificador. 3. El sistema realizará la clasificación. 4. Si el sistema reconoce el insecto, mostrará en pantalla que fue identificado. 5. El sistema desplegará la información del insecto en una ventana.
Flujos Alternativos:	<ol style="list-style-type: none"> 1.1 El usuario puede acceder a la opción de glosario. 1.2 El usuario puede acceder a la opción de regresar al menú principal. 1.3 El usuario puede acceder a la opción del menú Acerca de. 1.4 El usuario puede acceder a la opción del menú Descortezadores. 5.1. Si el insecto no es reconocido el sistema regresará al menú principal para que seleccione otra forma de identificar el insecto, caso de uso Menú Principal
Excepciones:	<ol style="list-style-type: none"> 1. Si se presenta error al momento de realizar la clasificación, el sistema informara al usuario y dará la opción de volver a iniciar la identificación.
Incluye:	<ol style="list-style-type: none"> 1. Caso de uso Menú Principal
Frecuencia de Uso:	Cada vez que se inicie la identificación con Clasificador Supervisado.

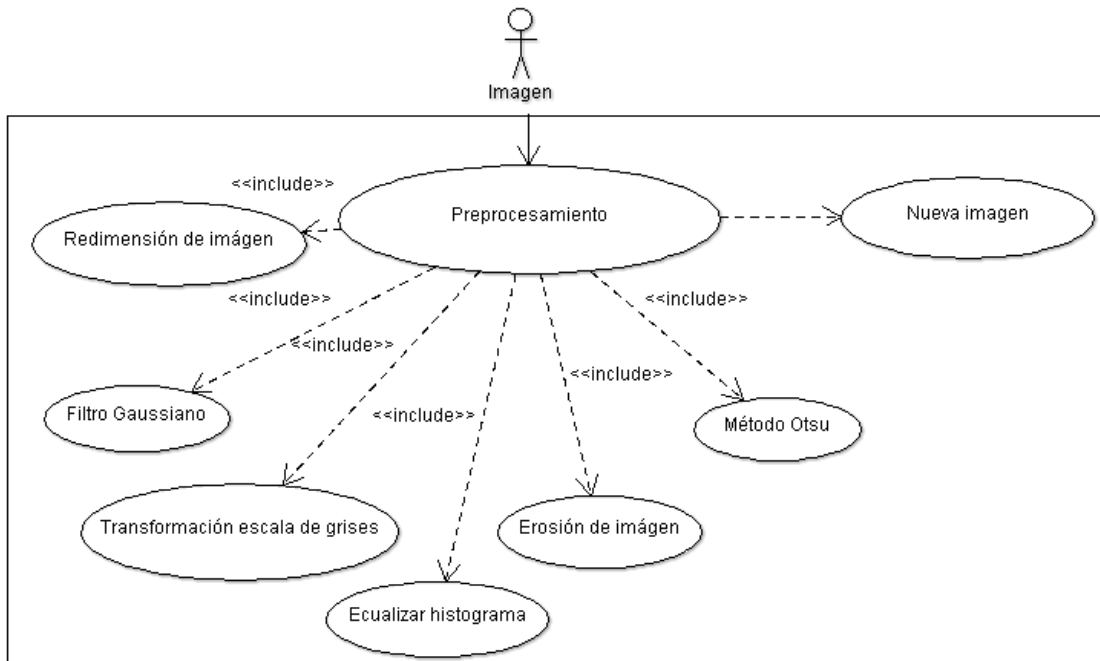
Fuente: Elaboración propia.

Figura 19. Diagrama de Caso de Uso: Visión por Computadora



Fuente: Elaboración propia.

Figura 20. Caso de Uso: Preprocesamiento



Fuente: Elaboración propia.

Tabla 13. Caso de Uso: Visión por Computadora

Nombre del Caso de Uso:	Visión por Computadora
No. De Caso de Uso:	6
Actores:	Usuario Final, Sistema
Descripción:	Identificar insectos Descortezadores con imágenes de los insectos previamente capturadas.
Precondiciones:	<ol style="list-style-type: none"> 1. El usuario debe contar con las imágenes del insecto de la cabeza, pronoto, cuerpo, lateral, declive elitral y élitro. 2. El usuario debe encontrarse en la ventana de visión por computadora.
Poscondiciones:	El sistema estará en espera de la selección de las imágenes por parte del usuario.
Flujo Normal:	<ol style="list-style-type: none"> 1. El sistema desplegará una ventana donde el usuario debe seleccionar 6 imágenes de diferentes partes del cuerpo del insecto donde se aprecien sus características, posteriormente dará clic en el botón Identificar. 2. Después de cargar las fotografías el usuario hará clic en el botón de iniciar análisis. 3. El sistema preparará la imagen con un pre procesamiento, caso de uso Preprocesamiento. 4. Con la imagen preprocesada se iniciará la obtención de características, caso de uso EXTCAR 5. Inicia la clasificación, caso de uso CLAS. 6. Si la imagen es clasificada correctamente el sistema desplegará la información del insecto.
Flujos Alternativos:	<ol style="list-style-type: none"> 1.1 El usuario puede acceder a la opción de glosario. 1.2 El usuario puede acceder a la opción de regresar al menú principal. 1.3 El usuario puede acceder a la opción del menú Acerca de. 1.4 El usuario puede acceder a la opción del menú Descortezadores. 6.1. Si el sistema no identifica el insecto en la imagen, propondrá al usuario iniciar la identificación con otra opción de identificación.
Excepciones:	<ol style="list-style-type: none"> 1. Si se presenta error al momento de realizar la carga de las imágenes, el sistema informara al usuario y dará la opción de volver a cargar las imágenes.
Incluye:	<ol style="list-style-type: none"> 1. Caso de uso Menú Principal 2. Caso de uso Preprocesamiento 3. Caso de uso CLAS 4. Caso de uso EXTCAR
Frecuencia de Uso:	Cada vez que se inicie la identificación con Visión por Computadora

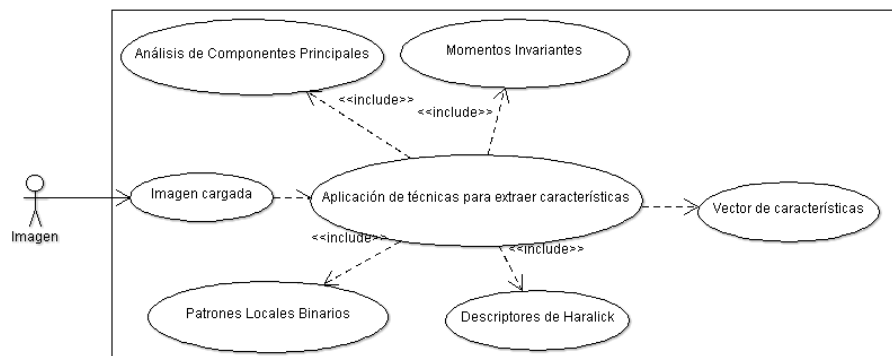
Fuente: Elaboración propia.

Tabla 14. Caso de Uso: Preprocesamiento

Nombre del Caso de Uso:	Preprocesamiento
No. De Caso de Uso:	7
Actores:	Imagen Sistema
Descripción:	El sistema iniciará un preprocesamiento de las imágenes cargadas o capturadas por el usuario, para prepararlas para obtener las características que servirán para iniciar la clasificación.
Precondiciones:	El usuario debe haber cargado las imágenes
Poscondiciones:	Se obtendrá un vector con las nuevas imágenes generadas por las técnicas del preprocesamiento
Flujo Normal:	<ol style="list-style-type: none"> 1. La imagen será escalada a 640x480 pixeles. 2. El sistema aplicará un filtro Gaussiano. 3. La imagen se transformará a escala de grises. 4. Se ecualizará el histograma de la imagen. 5. Se aplicará el Método Otsu para binarizar la imagen. 6. La imagen será erosionada. 7. Cuando el sistema termine de aplicar las técnicas generará un vector con las nuevas imágenes
Flujos Alternativos:	
Excepciones:	Si se presenta error, el proceso se terminará informando al usuario.
Incluye:	
Frecuencia de Uso:	Cada vez que se inicie la identificación con Visión por Computadora

Fuente: Elaboración propia.

Figura 21. Caso de Uso: EXTCAR



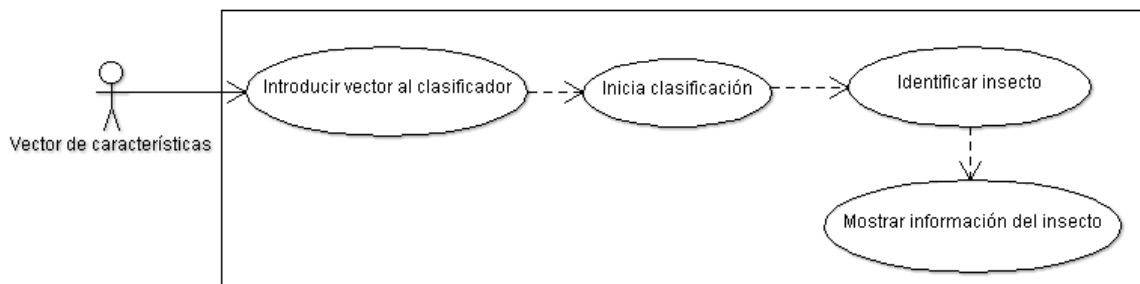
Fuente: Elaboración propia.

Tabla 15. Caso de Uso: EXTCAR

Nombre del Caso de Uso:	EXTCAR
No. De Caso de Uso:	8
Actores:	Imagen Sistema
Descripción:	El sistema iniciará la extracción de características a partir de las imágenes preprocesadas, para obtener el vector de características que servirá para iniciar la clasificación.
Precondiciones:	El sistema debe haber realizado el preprocesamiento a las imágenes cargadas por el usuario.
Poscondiciones:	Se obtendrá un vector con las características.
Flujo Normal:	<ol style="list-style-type: none"> 1. El sistema aplicará Momentos Invariantes. 2. El sistema aplicará Patrones Locales Binarios. 3. El sistema aplicará Descriptores de Haralick. 4. El sistema aplicará Análisis de Componentes Principales. 5. Cuando el sistema termine de aplicar las técnicas generará un vector con las características.
Flujos Alternativos:	
Excepciones:	Si se presenta error, el proceso se terminará informando al usuario.
Incluye:	
Frecuencia de Uso:	Cada vez que se inicie la identificación con Visión por Computadora

Fuente: Elaboración propia.

Figura 22. Caso de Uso: CLAS



Fuente: Elaboración propia.

Tabla 16. Caso de Uso: CLAS

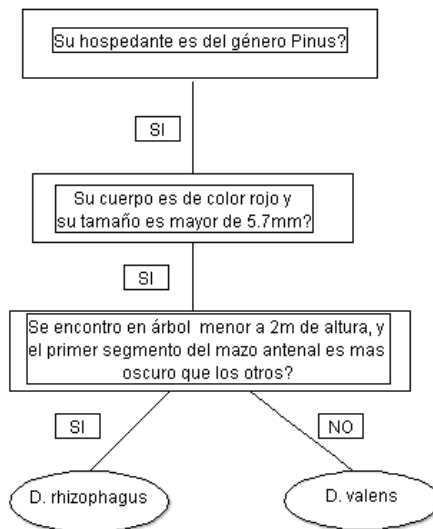
Nombre del Caso de Uso:	CLAS
No. De Caso de Uso:	9
Actores:	Vector de características Sistema
Descripción:	El sistema iniciará la clasificación a partir del vector de características.
Precondiciones:	El sistema debe haber realizado la extracción de características.
Poscondiciones:	Se obtendrá el nombre de la especie de insecto.
Flujo Normal:	<ol style="list-style-type: none"> 1. El sistema aplicará iniciará la clasificación del insecto. 2. Si el sistema logra identificar el insecto, mostrara la información del insecto en una ventana nueva. 3. Cuando el sistema termine la clasificación se obtendrá la especie de insecto descortezador.
Flujos Alternativos:	
Excepciones:	Si se presenta error, el proceso se terminará informando al usuario.
Incluye:	
Frecuencia de Uso:	Cada vez que se inicie la identificación con Visión por Computadora
Fecha:	12 de Septiembre 2018

Fuente: Elaboración propia.

11.5. Módulo por Reglas

A partir de la Tabla 7 se procedió a construir las reglas para identificar cada especie de insecto. Para la formulación de las reglas se realizaron dos árboles de decisiones: uno contemplando el orden de importancia de las características que los expertos en el área consideran, y otro generado a partir de las claves dicotómicas para identificar especies de insectos descortezadores del género *Dendroctonus* e *Ips*. En la Figura 23 se puede apreciar un ejemplo de una rama del árbol de decisión para dos especies de insecto descortezador del género *Dendroctonus*, la diferencia entre las dos especies son: la coloración de los segmento del mazo antenal y la altura del árbol atacado.

Figura 23. Ejemplo árbol de decisión para *D. valens* y *D. rhizophagus*



Fuente: Elaboración propia

Haciendo un seguimiento de ambos árboles se determinó que el hecho con las claves es el que da las bifurcaciones de la mejor manera porque el orden en que considera las características permite avanzar por las reglas para obtener la especie. Es importante comentar, que las características que los expertos indicaron como importantes, se encuentran en diferente orden que el árbol generado por las claves. El que se hizo con el orden de las características de los expertos en algunos casos quedaba con bifurcaciones sin respuesta. Cabe mencionar que como no se está trabajando con todas las especies de ambos géneros, quedaron caminos sin resultado en ambos árboles de decisión, dando por terminado esos caminos informando que no es una especie registrada. De los arboles antes mencionados se obtienen las siguientes reglas con IF anidados para identificar ambos géneros de insectos.

Ejemplo notación con IF anidados:

IF declive elitral no presenta espinas

IF color de cuerpo es rojo

IF coloración del primer segmento del mazo antenal es más oscuro que los otros

D. rhizophagus

ELSE

D. valens

Para una mejor comprensión, se cambió la notación de las reglas creando una por cada especie de insecto del tipo IF-THEN comparado con los IF anidados que se manejaban, con esto se logra evitar dejar caminos sin resultado (ver Anexo 1).

Como se presentó anteriormente, se emplearon reglas para representar el conocimiento debido a que los enunciados de tipo IF...THEN es la forma coloquial que los expertos utilizan. Las estructuras de las reglas como los valores de las características fueron plasmadas en Prolog.

En la Figura 24 se presenta las características plasmadas en Prolog a través de hechos que se utilizan para identificar a una de las especies del género *Dendroctonus* llamada adjuntus Blandford. Dentro de las características se encuentran: las especies de pino donde habitan; los daños que ocasiona en los arboles de pino; la parte del árbol que ataca; si es plaga primaria o secundaria y la cantidad de generaciones que presenta al año, por mencionar algunas.

Figura 24. Representación de conocimiento por medio de reglas en Prolog de la especie *Dendroctonus adjuntus* Blandford

```

%-----D. adjuntus_Blandford-----
% B habita en A
habita_en(maximinoi_H_E_Moore,adjuntus_Blandford).
habita_en(arizonica_Engelm,adjuntus_Blandford).
habita_en(leiophylla_Schl_y_Cham,adjuntus_Blandford).
habita_en(oocarpa_Shiede,adjuntus_Blandford).
habita_en(muntezumae_Lamb,adjuntus_Blandford).
habita_en(ayacahuite_Ehren,adjuntus_Blandford).
habita_en(teocote_Schl_y_Cham,adjuntus_Blandford).
habita_en(durangensis_Ehren,adjuntus_Blandford).
habita_en(pringlei_Shaw,adjuntus_Blandford).
habita_en(pseudostrobos_Lindl,adjuntus_Blandford).
% B se distribuye en A
se_distribuye_en(jalisco,adjuntus_Blandford).
se_distribuye_en(chihuahua,adjuntus_Blandford).
se_distribuye_en(durango,adjuntus_Blandford).
se_distribuye_en(michoacan,adjuntus_Blandford).
% B se desarrolla en una altitud de A
altitud(1300,3949,adjuntus_Blandford).
% B tiene un tamaño mínimo de A y máximo de C
tamaño(2.9,6.9,adjuntus_Blandford).
% B es de color A
color(negro,adjuntus_Blandford).
color(cafe_oscuro,adjuntus_Blandford).
% B tiene forma de cabeza A
forma_cabeza(convexa, adjuntus_Blandford).
% B tiene puntuaciones en pronoto
puntuaciones_pronoto(si,adjuntus_Blandford).
% B tiene setas en el declive elitral A
setas_declive_elitral(grandes_y_escasas,adjuntus_Blandford).
% B ocasiona A
daños(grumos_resina,adjuntus_Blandford).
daños(muerte_arbol,adjuntus_Blandford).
% B ataca A
ataca(tronco,adjuntus_Blandford).
% B es de tipo de plaga A
tipo_plaga(primaria,adjuntus_Blandford).
% B presenta generaciones por año A
generaciones(1,adjuntus_Blandford).

```

Fuente: Elaboración propia.

11.5.1. Validación de reglas

La validación de las reglas fue realizada por los expertos que participaron desde el inicio del proyecto. Para validar el conjunto de reglas, se les proporcionaron de forma textual las pertenecientes a cada especie de insecto, haciendo mención en si los valores de cada característica eran correctos. Por ejemplo, la regla:

IF

declive elitral presenta espinas AND
5 espinas en el declive elitral AND
tercer espina es pequeña y con gancho AND
epistoma presenta línea de tuberculos AND
tamaño entre 4 y 4.7 mm

THEN

I. lecontei

Se traduce de forma textual a:

SI

- El insecto presenta espinas en el declive elitral Y
- La cantidad de espinas es 5 Y
- El tamaño de la tercer espina es pequeña y con forma de gancho Y
- Presenta línea de tuberculos en el epistoma Y
- El tamaño del insecto se encuentra entre 4 y 4.7 mm

Entonces

- El insecto es *Ips lecontei*

11.5.2. Validación de interfaz de usuario

Se realizaron dos prototipos para determinar la mejor forma de interacción entre el usuario y el módulo por reglas. La interfaz fue implementada en Java. Se optó por utilizar este lenguaje debido a la flexibilidad que ofrece, lo cual permitió desarrollar una interfaz amigable de fácil entendimiento. Cabe mencionar que se evaluó la posibilidad de emplear Prolog, si bien es cierto este lenguaje permite la posibilidad de hacer interfaces gráficas, el inconveniente radica en que las interfaces se manejan mediante programación lógica, la cual difiere radicalmente del estilo de programación imperativo con el cual se está familiarizado. No obstante, se hicieron prototipos de interfaz básicos, pero dado que no brinda la misma riqueza en interfaces gráficas, se prefirió descartarlo.

El primer prototipo se desarrolló para comprobar el funcionamiento de las reglas hechas en prolog, por lo que no contaba con una interfaz gráfica amigable para el usuario. El segundo prototipo estaba enfocado en la utilización del sistema por parte del usuario por lo que se realizó la interfaz gráfica sencilla y amigable.

11.5.3. Primer prototipo

El primer prototipo tenía la posibilidad de brindar al usuario una interfaz donde el SE preguntaba al usuario las características físicas del insecto para que se describiera cada uno de ellas a partir de un conjunto de posibles valores. Por ejemplo:

SE: ¿El insecto presenta espinas en el declive elitral?

Usuario: SI

SE: ¿Cuántas espinas presenta?

Usuario: 5

SE: ¿Qué coloración presenta el insecto?

Usuario: Negro

Aunque se asemeja a un dialogo entre humanos, la desventaja fue que de esta forma la interacción era un poco tediosa, además de que algunos usuarios no reconocían de que parte del insecto se refería la pregunta, porque el nivel de capacitación era distinta entre usuarios, y la intención era desarrollar un sistema que pueda ser utilizado por cualquier persona. Partiendo de lo anterior se diseñó un segundo prototipo y versión final, que se describe a continuación.

11.5.4. Segundo prototipo

En este prototipo se presentaron tres formas de interacción:

1. Búsqueda Libre: en una misma ventana se preguntaban diversas características para consultar que especie de insecto cumplía con ellas.
2. Búsqueda Guiada: a través de preguntas sobre las características que presenta el insecto acompañado de una imagen que refería la parte del insecto a evaluar se iba consultando las especies de insecto.
3. Listado de Especies: en una ventana se enlistan todas las especies de insecto, donde el usuario al momento de seleccionar puede ver la descripción del insecto.

Con este prototipo se logró una mejor interacción por parte de los usuarios dando una mayor idea de lo preguntado.

11.5.5. Descripción de la Interfaz de Usuario

En la Figura 13 se muestra la interfaz principal del sistema, las opciones de búsqueda del Módulo por Reglas son las que se encuentran encerradas en un rectángulo rojo.

La Figura 25 muestra la interfaz de la búsqueda libre, donde el usuario debe seleccionar los valores de las características, posteriormente dar clic en el botón ver especies posibles mostrando los resultados en una tabla de la misma ventana, al dar clic en alguna fila de la tabla se mostrara la información del insecto.

Figura 25. Interfaz de búsqueda libre

The screenshot shows the 'Busqueda Libre' window in the SEIID 1.0 application. The window title is 'SEIID 1.0' and the menu bar includes 'Inicio', 'Glosario', 'Descortezadores', and 'Acerca de'. The main area is titled 'Busqueda Libre' and contains several search filters:

- Distribución:** A dropdown menu with 'jalisco' selected.
- Género:** A dropdown menu with 'Seleccione uno' selected.
- Pinos que ataca:** A dropdown menu with 'Seleccione uno' selected.
- Color:** A dropdown menu with 'Seleccione uno' selected.
- Altitud:** A slider ranging from 0 to 4000, with the current value set to 0.
- Tamaño:** A slider ranging from 2 to 10, with the current value set to 2.

Below the filters is a table with the following columns: 'Fotografías', 'Nombre', 'Género', and 'Tipo plaga'. The table contains two rows of results:

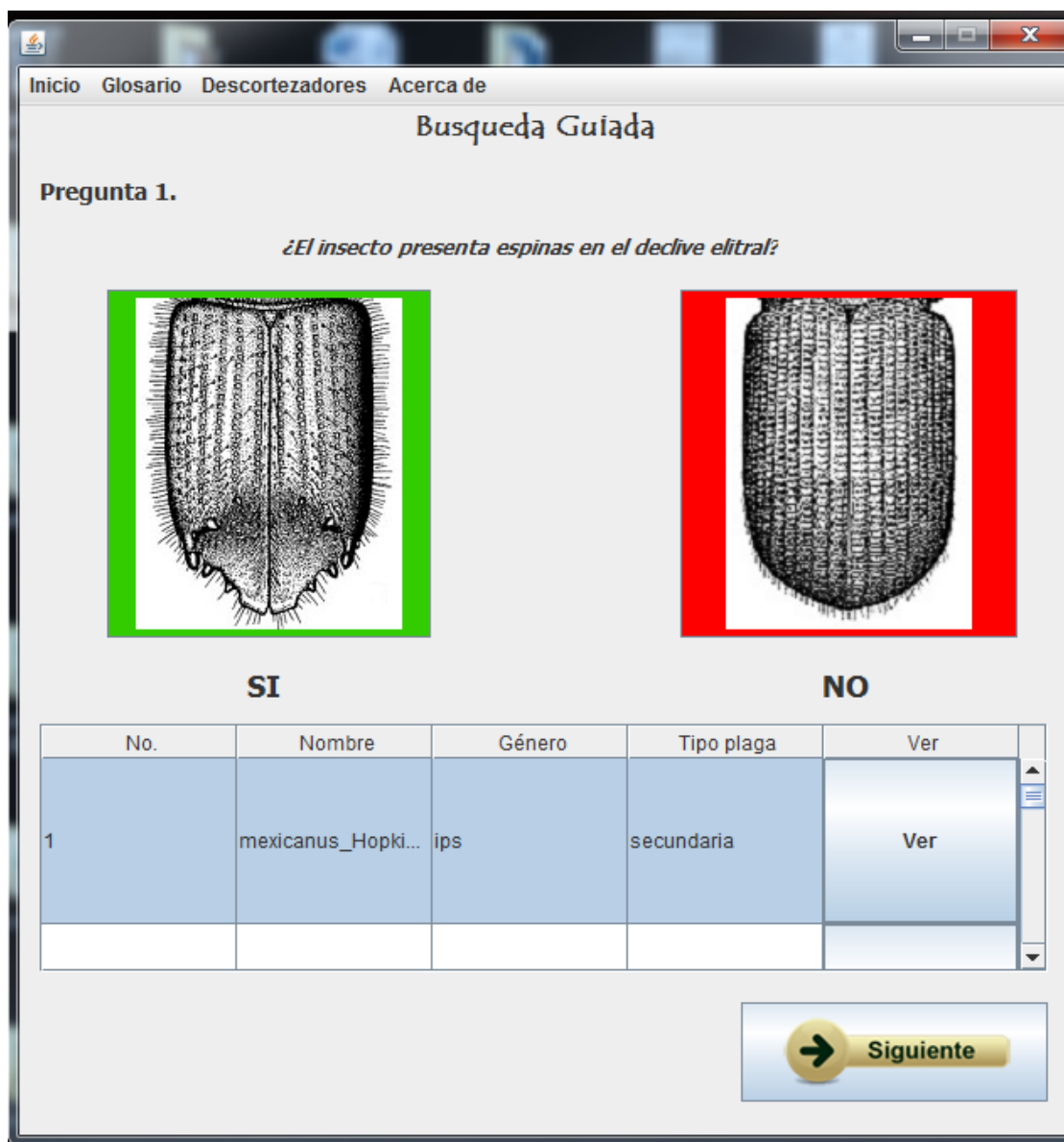
Fotografías	Nombre	Género	Tipo plaga
javax.swing.JLabel[,0,0,...	adjuntus_Blandford	dendroctonus	primaria
javax.swing.JLabel[,0,0,...	adjuntus_Blandford	dendroctonus	primaria

At the bottom of the window is a button labeled 'Ver especies posibles'.

Fuente: Elaboración propia.

En la Figura 26 tenemos la búsqueda guiada, dependiendo la pregunta se muestra una imagen correspondiente a la parte del insecto que se está evaluando, al seleccionar alguna de las opciones se muestran las especies que presenta la opción elegida, el botón siguiente muestra la siguiente pregunta de la búsqueda guiada.

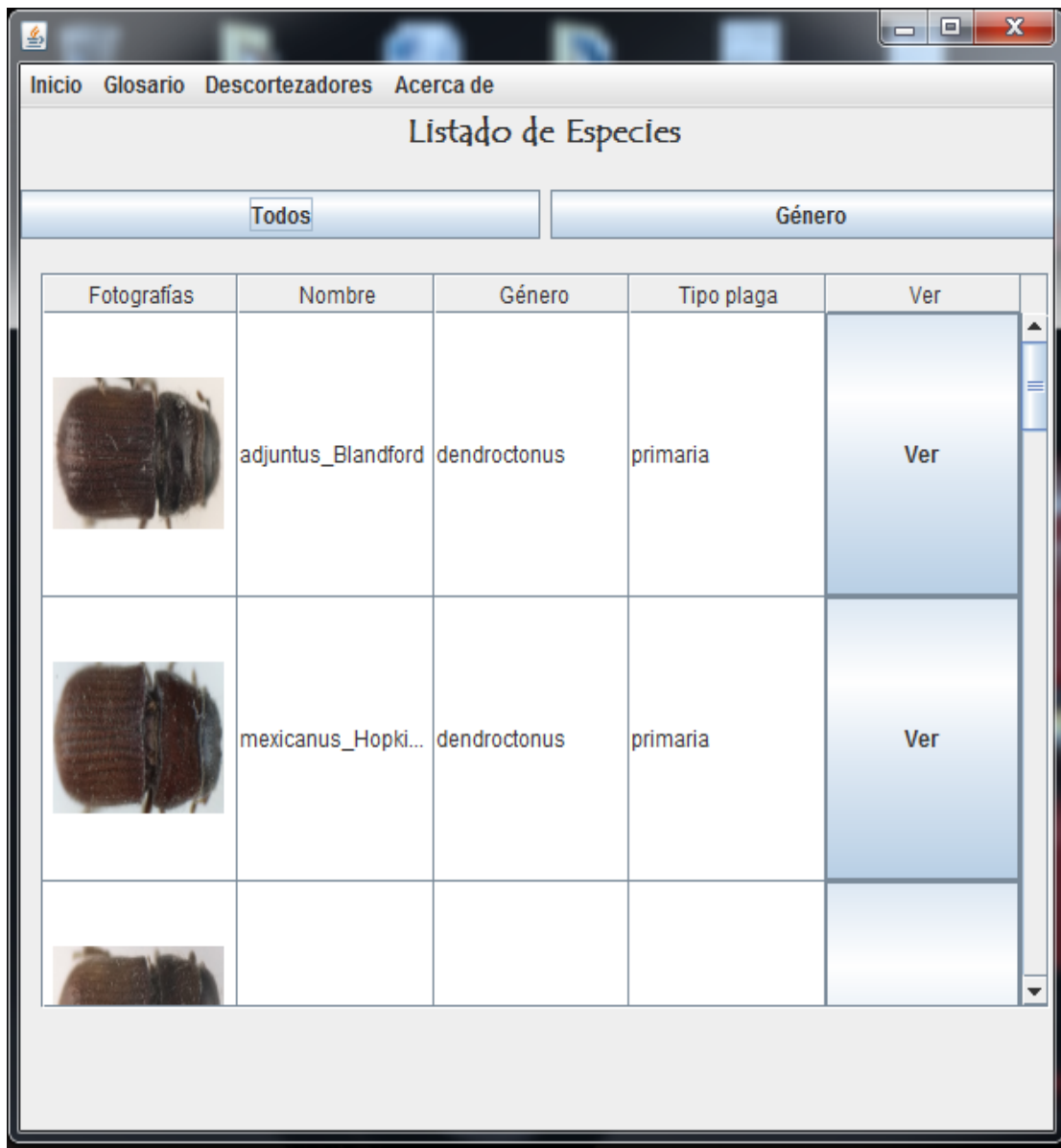
Figura 26. Interfaz de búsqueda guiada



Fuente: Elaboración propia.

En la Figura 27 se tiene la interfaz de listado de especies, se muestra todas las especies de insecto que se tienen en la base de conocimiento, al seleccionar alguna al igual que las otras interfaces carga la información del insecto.

Figura 27. Interfaz listado de especies



Fuente: Elaboración propia.

11.5.6. Resultados

La implementación de la interface se desarrolló con el lenguaje Java, debido a la flexibilidad y disponibilidad de componentes para realizar la conexión con Prolog, lenguaje utilizado para plasmar las reglas, lo que permitió crear un sistema amigable y de fácil uso por cualquier tipo de usuario. Otro punto a favor de la utilización de ambos programas fue que son libres.

Cabe agregar que también se evaluó la posibilidad de utilizar Visual Prolog, este lenguaje proporciona la posibilidad de hacer interfaces gráficas con programación lógica, se hicieron prototipos básicos de interfaces, pero debido a que no brinda la misma riqueza en interfaces gráficas por lo que fue descartado.

Se diseñó la base de conocimiento desde cero que almacena las reglas para identificar las 17 especies de insecto descortezador que atacan árboles de pino.

La interfaz proporciona al usuario tres formas (búsqueda libre, búsqueda guiada y listado de especies) de identificar los insectos.

La búsqueda guiada va preguntando al usuario con ayuda de ilustraciones de la parte evaluada para un mayor entendimiento de lo solicitado, orientado para usuarios inexpertos. Por otro lado la búsqueda libre es orientada para usuarios con mayor conocimiento sobre el tema donde se coloca las variables para ir moviendo los valores de forma manual mostrando las coincidencias. Ambas búsquedas utilizan reglas con encadenamiento hacia adelante para inferir las especies de insecto descortezador.

El listado de especies utiliza el encadenamiento hacia atrás, a partir de la especie de insecto infiere sus características y muestra la información.

11.5.7. Evaluación del Sistema Experto

Se procedió a simular situaciones basándose en registros históricos procedentes de identificaciones realizadas por los expertos. Para ello se empleó la información siguiente:

- Base de datos de insectos descortezadores identificados en diferentes partes de la República Mexicana, emitida por el laboratorio de entomología del departamento de Ciencias Forestales de la Universidad Autónoma Chapingo.

La base de datos con la que se cuenta registra datos de localidad, municipio y estado en donde se encontró el insecto, altitud, especie de árbol atacado, etc., pero no incluye características físicas del insecto debido a que los expertos las conocen y no son relevantes para almacenarlas en la base de datos. Por lo que conociendo la especie que se trate en cada registro los expertos pueden contestar lo solicitado en cuanto a las características físicas que debe tener cada especie.

De la base de datos se seleccionaron 30 situaciones diferentes, tomando información como: estado donde se encontró el insecto, altitud, especie de árbol atacado. Para que el usuario pudiera contestar la información sobre las características físicas del insecto se utilizaron imágenes correspondientes a cada especie. Los resultados

obtenidos con el Sistema Experto fueron contrastados con los que se tenían en la base de datos, tras lo cual se concluye que el 92% de los casos fueron idénticos.

11.5.8. Discusión

Los resultados obtenidos se encuentran por debajo de los reportados en diversos trabajos como el sistema que Shiguihara y Kepler (2009), trabaja con síntomas ocasionados en las plantas por plagas y enfermedades, el sistema propuesto no podría hacerlo debido a que todos los insectos descortezadores pueden ocasionar los mismos síntomas, o en su defecto se puede encontrar varias especies atacando un árbol al mismo tiempo lo que causaría confusión si se trabajase con síntomas que presentan los árboles atacados.

Por otro lado Nascimento et al (2016) con su aplicación para dispositivos Android alcanza un porcentaje de precisión del 98.44%, en nuestro caso la precisión fue similar y como trabajo futuro es desarrollar una aplicación para dispositivos móviles para que se pueda llevar a campo.

Uno de los sistemas que se asemeja al desarrollado en este trabajo es el de Kaloudisa et al (2005) donde considera especies de insectos que atacan árboles de pino, sin considerar insectos descortezadores, aunque no muestran el porcentaje de precisión, mencionan que se obtuvieron buenos resultados probándolo con diferentes grupos de personas desde expertos en el área hasta usuarios sin experiencia.

Ochoa y Kemper (2009) plantea un sistema experto que integra variables en tiempo real obtenidas por una estación Agro Climática, recomienda tratamientos indicando producto, dosis y época de aplicación, además de dar la opción de generar una bitácora para dar seguimiento. Como futura actualización del sistema que se desarrolló en este trabajo sería agregar variables climáticas para guardarlas a la hora que se realice la identificación para dar al usuario estadísticas adicionales que le puedan servir.

11.6. Módulo de Visión por Computadora

En este apartado, se describen los materiales y métodos empleados en el procesamiento de imágenes, con objeto de dar una solución a la problemática expuesta.

Materiales de trabajo:

- Base de datos conformada por 144 imágenes de insectos descortezadores.
- Laptop con procesador Intel Core i5 y 4 GB de memoria RAM.

- Matlab v2017a.

11.6.1. Adquisición de imágenes

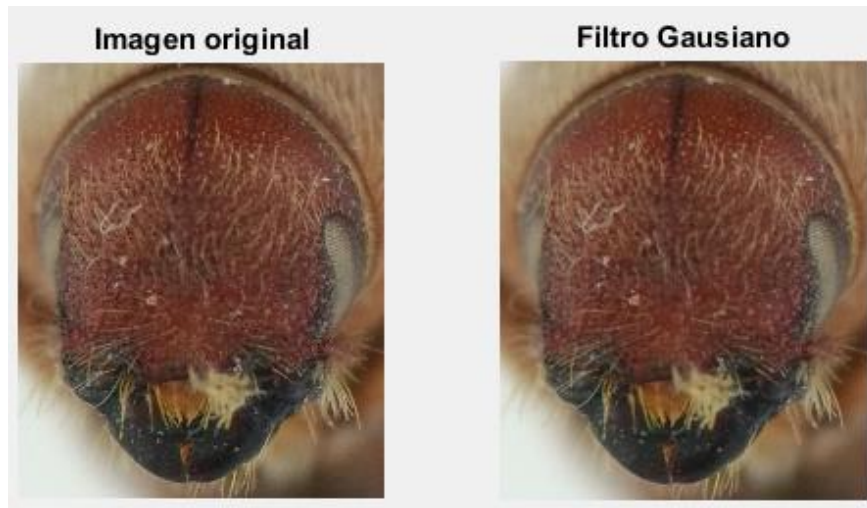
Para iniciar el experimento de visión por computadora se utilizó la base de datos de imágenes de las 4 especies del género *Dendroctonus*, esto debido a que son las especies de descortezadores que provocan mayores daños y de las que se tenía mayor cantidad especímenes para fotografiar.

11.6.2. Creación de base de datos

Se realizó un recorte manual de las áreas de interés (partes del insecto) de las 144 imágenes destinadas tanto para entrenar como para evaluar el sistema, además con ayuda de la función de Matlab `imresize()`, se escaló cada imagen a 640 x 480 píxeles para no tener inconvenientes cuando se iniciaran las operaciones dentro del preprocesamiento.

Se aplicó un filtro gaussiano (Figura 28) a cada imagen RGB escalada con la finalidad de eliminar ruido cuando se capturo la fotografía, tratando de emborronar los bordes lo menos posible, utilizando una máscara 3x3 y la función `imfilter()` de Matlab.

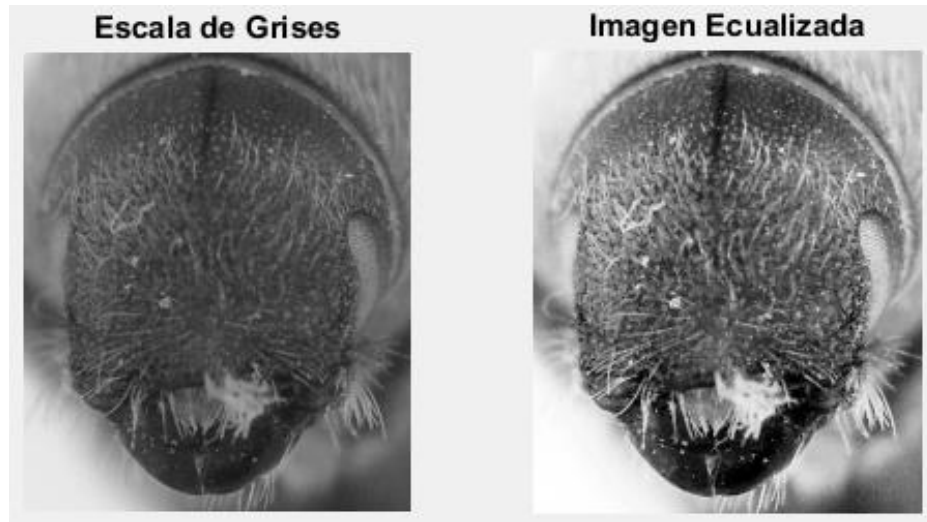
Figura 28. Aplicación de Filtro Gaussiano



Fuente: Elaboración propia.

Posteriormente se transforma la imagen RGB con filtro Gaussiano a escala de grises para poder realizar la ecualización del histograma como se observa en la Figura 29, con las funciones `rgb2gray()` y `histeq()` respectivamente.

Figura 29. Ecuación de imagen a escala de grises



Fuente: Elaboración propia.

Subsiguientemente a partir de la imagen ecualizada se aplicó el método Otsu para binarizar la imagen (Figura 30), utilizando un algoritmo implementado en Matlab.

Figura 30. Binarización de imagen ecualizada



Fuente: Elaboración propia.

Para lograr una mejor segmentación del insecto, se aplicó la función de erosión con la finalidad de rellenar espacios en blanco estilizando la forma del insecto como se observa en la Figura 31.

Figura 31. Erosión de imagen



Fuente: Elaboración propia.

Como resultado del proceso de preprocesamiento se obtiene una imagen segmentada con la que se aplicaran técnicas de extracción de características, explicadas en el apartado siguiente.

11.6.3. Extracción de características

Después de segmentar la imagen se procedió a extraer características para su posterior clasificación.

- Momentos Invariantes
- Patrones binarios locales (LBP)
- Descriptores de Haralick

Con el objetivo de reducir el costo computacional debido a la dimensionalidad de los vectores de características se aplicó el ACP. La cantidad de componentes se seleccionó a partir de la suma de los valores propios que acumule el 95% del total de la varianza.

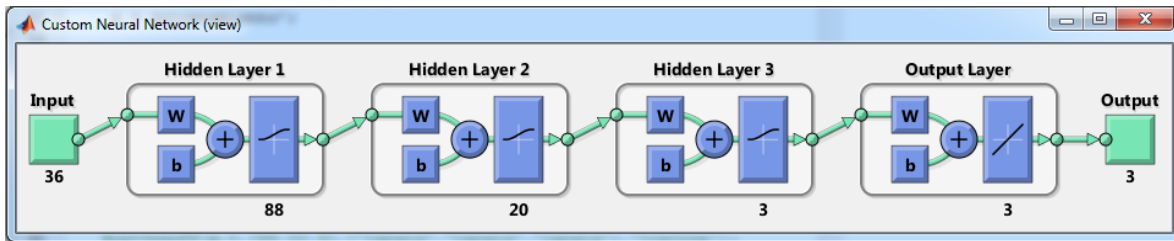
11.6.4. Vector de Características

El vector resultante de la extracción de características está conformado por: 7 momentos invariantes, 255 patrones binarios locales y 14 descriptores de Haralick; obteniendo una matriz de vectores de 276 por 144. Utilizando ACP se redujo la dimensión de la matriz de vectores a 36 características por 144 imágenes.

11.6.5. Clasificación de patrones

Se utilizó una RNFF, ver Figura 32, para la clasificación de patrones, con 3 capas cuya función de activación fue sigmoideal en sus tres capas, la cantidad de neuronas en la primer capa fue de 88 como método de entrenamiento se usó el método trainrp.

Figura 32. Arquitectura Red Neuronal

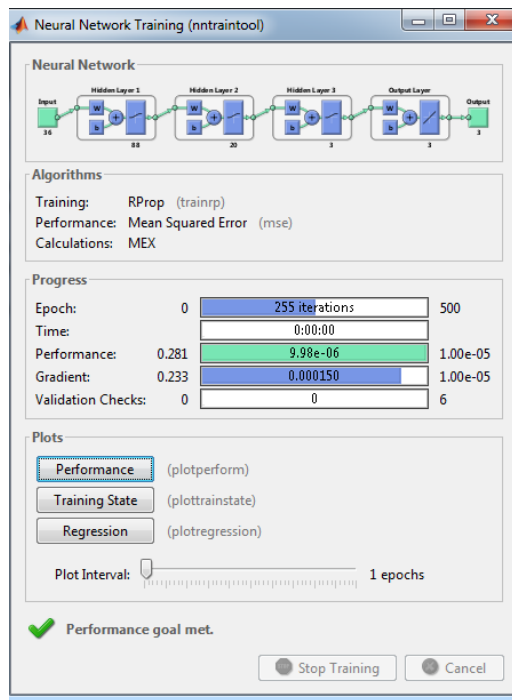


Fuente: Elaboración propia.

11.6.6. Resultados del Experimento

Después de realizar 10 corridas con diferentes conjuntos de datos obtenidos aleatoriamente, se obtuvo el mejor modelo quedando entrenado en 255 épocas, como se observa en la Figura 33.

Figura 33. Entrenamiento de la Red



Fuente: Elaboración propia.

Generalmente en el funcionamiento de la red al momento del entrenamiento el error se reduce con el paso de las épocas de entrenamiento, pero puede empezar a aumentar cuando la red comienza a sobre ajustar los datos de entrenamiento. En este caso en la época 255 se obtuvo el mejor rendimiento del entrenamiento con un valor de $9.9846e^{-6}$.

11.6.7. Evaluación del clasificador

Para llevar a cabo la evaluación del clasificador se utilizó el porcentaje de precisión global, calculado a partir de la matriz de confusión, ver Tabla 17. El cálculo se realizó con la función provista por matlab para la tarea, `plotconfusion(target,output)`, dando como resultado un 86.2% de precisión.

Tabla 17. Matriz de confusión del clasificador

Clase	1	2	3	Totales
1	7	0	0	7
2	0	7	0	7
3	0	4	11	15
Totales	7	11	11	29

Fuente: Elaboración propia.

11.6.8. Descripción de la interfaz de usuario

La opción para acceder al módulo dentro del Menú Principal se encuentra localizada en la parte inferior izquierda como se aprecia en la Figura 13, encerrada en un cuadro color amarillo. Al seleccionar la opción de Búsqueda con Visión, nos abre una nueva ventana (Figura 34) donde se tiene que cargar las fotografías de las partes del insecto, para su posterior preprocesamiento y clasificación, que se lleva a cabo al momento de dar clic en el botón identificar.

Figura 34. Búsqueda con Visión Artificial



Fuente: Elaboración propia.

11.6.9. Discusión de los Resultados del Experimento

El porcentaje de precisión alcanzado en este experimento está por debajo de los trabajos realizados por diversos autores que trabajan en la identificación de insectos utilizando procesamiento de imágenes, como lo demuestran: Ebrahimi et al. (2017) aplicando SVM se alcanzó una precisión de menos del 2.5% de error en la identificación de parásitos de invernadero; Wang et. al. (2012) alcanzando un 94% de precisión con redes neuronales identificando insectos de acuerdo al orden al que pertenecen, considerando solo 8 órdenes de insecto y un 98% de exactitud con SVM.

Para el caso de la red neuronal se descartó el orden Coleoptera que es el orden al que pertenecen los insectos de este experimento; Qing et. al. (2012) utilizando SVM alcanzaron una tasa de precisión del 97.5%, identificando cuatro especies del género Lepidoptera; Hongtao et. al. (2011) utiliza una SVM para identificar 15 especies de insectos que atacan grano almacenado utilizando 7 características con una precisión del 94.67%; Al-Saqer & Hassan (2011) aplica una red neuronal feed-forward para identificar gorgojos y otros insectos que atacan la palma roja obteniendo una precisión del 99 y 93% respectivamente; Leow et al. (2015) al igual que los autores anteriores utiliza una red neuronal feed-forward para identificar cepodos con una precisión de 93.1%; Do, Harp, & Norris (1999) también utiliza una red neuronal para identificar arañas, obteniendo un 81% de precisión.

El porcentaje de precisión se debe en mayoría por el conjunto de imágenes y la cantidad de especímenes por especie que se tenía a disposición para el experimento, aunque se trató de controlar la iluminación y posición del insecto al tomar la fotografía, utilizando un estereoscopio para enfocar, el tamaño diminuto y la semejanza entre especies del mismo género fueron un factor que hizo que la observación de las características en las imágenes fuera complicada, además de que algunos insectos debido al manejo en laboratorio y campo al momento de realizar la colecta se encontraban maltratados faltándoles algunas partes del mismo (antenas, patas, etc.) lo que perjudica a la hora de entrenar el modelo para el reconocimiento del insecto. También cabe agregar que en este experimento no se diferenció el sexo del insecto lo que puede haber representado un problema, porque en algunas especies no comparten ciertas características o pueden variar entre sexos.

Aunque el sistema está pensado que pueda ser utilizado en campo, la visión por computadora se complicaría debido a las condiciones para capturar las imágenes del insecto, Martineau et. al. (2017) menciona que para el reconocimiento de insectos con procesamiento de imágenes, es mejor utilizar imágenes capturadas en laboratorio bajo condiciones controladas. Las imágenes obtenidas en laboratorio son la mejor opción para realizar el reconocimiento, debido a que se tiene un mayor control de las condiciones para tomar la fotografía y el posicionamiento del insecto se realiza de forma manual con pose libre, debido a que en campo es más difícil lograr las condiciones adecuadas.

11.7. Módulo de Inteligencia Artificial (IA)

Este módulo del sistema está pensado para que los usuarios a través de un cuestionario con preguntas de 6 características, logren identificar a 5 especies del género *Dendroctonus* por medio de la observación de las características físicas del

insecto que se pueden observar en campo con ayuda de una lupa de mano o a través de una imagen del mismo.

11.7.1. Adquisición de imágenes y observaciones

Para la obtención de las imágenes de las especies de insectos; *adjuntus*, *brevicomis*, *frontalis*, *mexicanus* y *valens* del género *Dendroctonus*, se capturaron como se especifica en el apartado de **Materiales y métodos**.

Por otro lado, haciendo una revisión de bibliografía especializada (Cibrian Tovar D. , Méndez Montiel, Campos Bolaños, Yates III, & Flores Lara, 1995) (Wood, 1982) y con las claves dicotómicas que identifican insectos Descortezadores del género *Dendroctonus*, se obtuvieron un total de 27 partes de los insectos que caracteriza a cada especie, listados en el apartado de Adquisición de conocimiento: Características insectos descortezadores

Considerando que la identificación de especies es con características a nivel de campo, se eliminaron rasgos que no son visibles a simple vista o con una lupa. También con la intención de quedarse con menos características, se aplicó la técnica de análisis de conglomerados (Benítez & Díez, 2005) para determinar niveles de separabilidad entre cada especie, con ello se redujeron a 6 características, pero no fue suficiente para una separación total de especies, siempre hubo dos especies que no se diferenciaban. Estos resultados indicaban que se debía usar otra técnica que lograra la separación de todas las especies. Cualquier otra técnica a usarse requería de datos observados, para ello se planteó un cuestionario que se pudiera aplicar a personas, como si estuvieran en campo.

Para plantear y determinar los ítems que conformaron el cuestionario se consideraron las características coloración del insecto, presencia de tubérculo frontal y epistomal, tamaño de las setas en el declive elitral, coloración del mazo antenal y forma del pronoto, es de destacar que se tuvo la ayuda de los expertos entomólogos y una lupa de mano con aumento de 20x, ya que algunas especies son muy pequeñas. Antes de aplicar el cuestionario se evaluó con juicio de expertos para determinar que las preguntas sobre las características estuvieran planteadas correctamente y que las opciones de respuesta sirvieran para el propósito del trabajo.

Las imágenes fueron segmentadas (Figura 35) de forma manual para mantener la parte del insecto que se quería evaluar y que fueron incluidas en el cuestionario.

Figura 35. Segmentación de la imagen



a) Imagen original



b) Imagen segmentada

Fuente: Elaboración propia.

El cuestionario se aplicó a 140 personas obteniendo un total de 700 registros con los cuales se realizó un análisis de los datos para verificar las respuestas correctas, observando que pocos fueron los casos donde la cantidad de aciertos se alcanzaron en su totalidad. Con base en lo anterior se decidió depurar los datos considerando los registros que tenían una tasa del 50% de preguntas contestadas correctamente, dejando un total de 562 registros.

11.7.2. Técnicas usadas para la identificación de especies de insectos

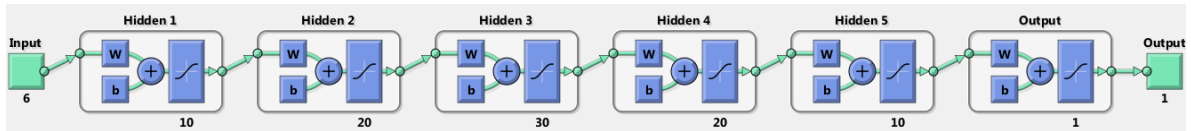
En el experimento se probaron las técnicas Red Neuronal Feed-Fordward Backpropagation, Red Bayesiana y Máquinas de Soporte Vectorial.

11.7.2.1. Red Neuronal Feed-Fordward Backpropagation

Se utilizó una Red Neuronal Feed-Forward conformada por 5 capas ocultas con 10, 20, 30, 20, 10 neuronas respectivamente; la capa de entrada con 6 neuronas que corresponden a cada una de las características evaluadas; la capa de salida solamente tiene 1 neurona que pertenece al valor numérico de cada clase; como función de transferencia en cada capa se utilizó la tangente hiperbólica sigmodea $a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$ como se puede apreciar en la Figura 36. Los parámetros establecidos para el entrenamiento del modelo fueron: 0.1 como tasa de aprendizaje, 1.0 como

alfa, 1000 épocas como la cantidad máxima, 0.00001 como valor mínimo de gradiente de rendimiento y 0.9 como constante de momento.

Figura 36. Estructura Red Neuronal Feed-Forward Backpropagation



Fuente: Elaboración propia.

11.7.2.2. Máquinas de Soporte Vectorial

Se optó por usar una clasificación uno contra uno, convirtiéndose en una clasificación binaria clásica. Se hizo de esta forma debido a que si se manejaba una clasificación de uno contra todos, al momento de generar los conjuntos de datos, la proporción de observaciones se sesgaba hacia la clase negativa, por lo que la identificación se inclinaba a respuestas erróneas, dando una tasa de precisión baja, esto debido a que la cantidad de observaciones por clase era de 1 contra 4. La técnica uno a uno la utilizaron Hongtao et. al. (2011) mostrando buenos resultados, en este trabajo se usó la función de Base Radial en el kernel y la configuración predeterminada de Matlab para la creación de la SVM: centrado automático de los puntos de datos en su valor medio, valor de C igual a 1, límite de cache del kernel igual a 5000, el método de Optimización Mínima Secuencial para la separación en el hiperplano, [1 -1] como parámetros del Perceptron Multicapa, orden del Kernel polinomial fue 3, factor de escala RBF igual a 1, el porcentaje de variables permitidas para violar las condiciones de Karush-Kuhn-Tucker fue de 0. Se obtuvieron un total de 20 clasificadores binarios, cuatro por cada clase.

11.7.2.3. Clasificador Bayesiano

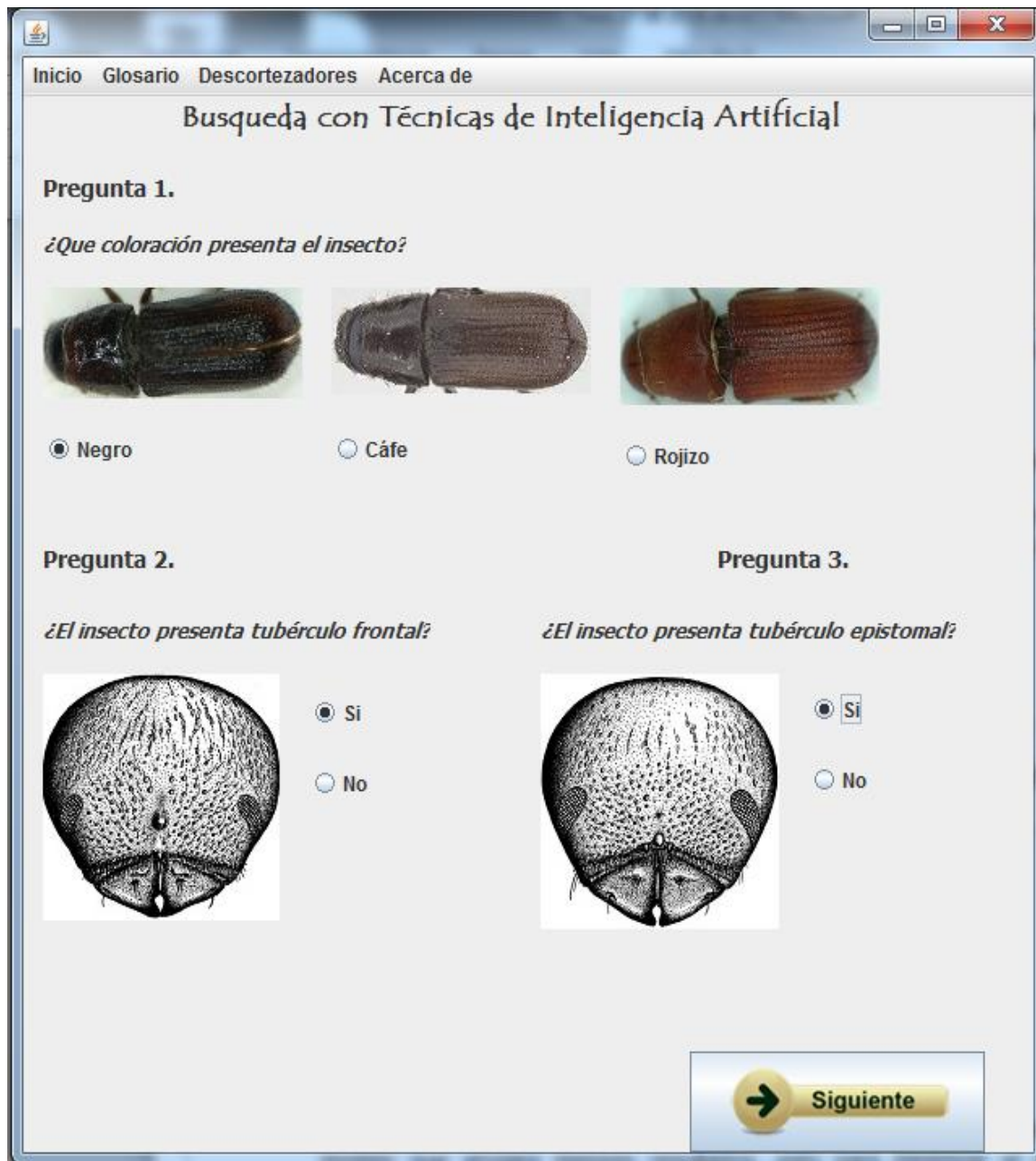
Para llevar a cabo el experimento se implementó el algoritmo de un CB Simple en MATLAB, donde se asume que los atributos son independientes dada la clase; cada atributo X_i es condicionalmente independiente de los demás atributos dada la clase: $P(X_i | X_j, C) = P(X_i | C), \forall j \neq i$.

11.7.3. Descripción de la interfaz de usuario del Módulo de Inteligencia Artificial

La opción dentro del Menú Principal para iniciar el módulo de Inteligencia Artificial, es la que se encuentra encerrada en un rectángulo verde, ver Figura 13. Para el caso de la interfaz de las preguntas (

Figura 37, Figura 38, Figura 39, Figura 40) que sirven para evaluar las características del insecto formando con las respuestas el vector que se envía al clasificador, se decidió colocar imágenes de las partes del insecto evaluadas para un mayor entendimiento por parte de los usuarios al momento de responder.

Figura 37. Interfaz Módulo IA: Pregunta 1 a 3



Fuente: Elaboración propia.

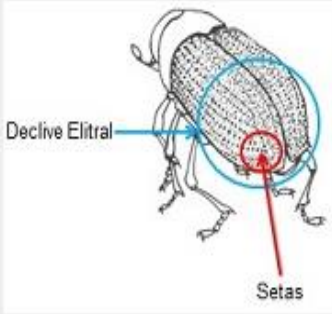
Figura 38. Interfaz Módulo IA: Pregunta 4

Inicio Glosario Descortezadores Acerca de

Busqueda con Técnicas de Inteligencia Artificial

Pregunta 4.

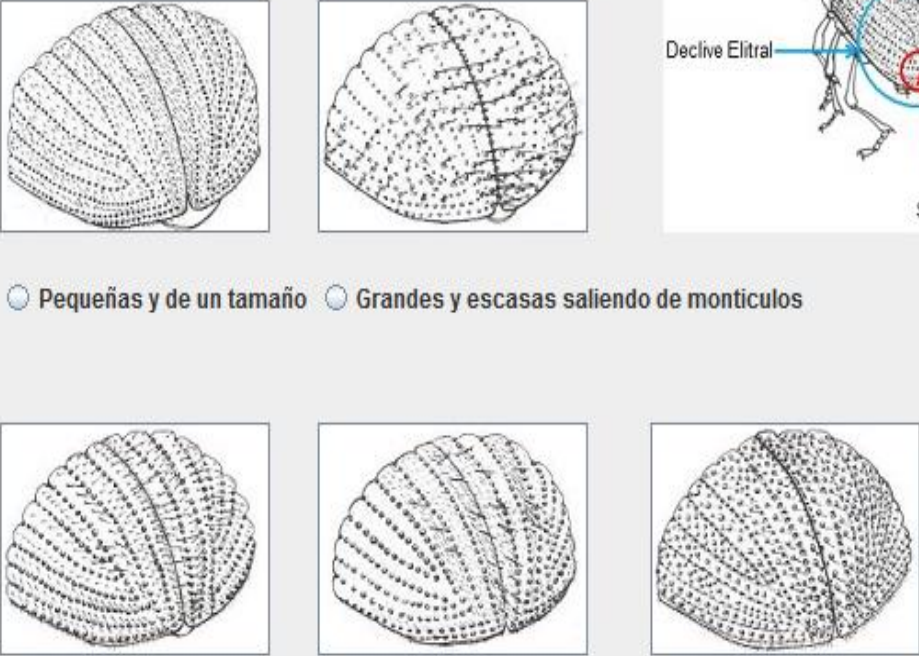
¿De que tamaño son las setas del declive elitral?




Declive Elitral

Setas

Pequeñas y de un tamaño Grandes y escasas saliendo de montículos

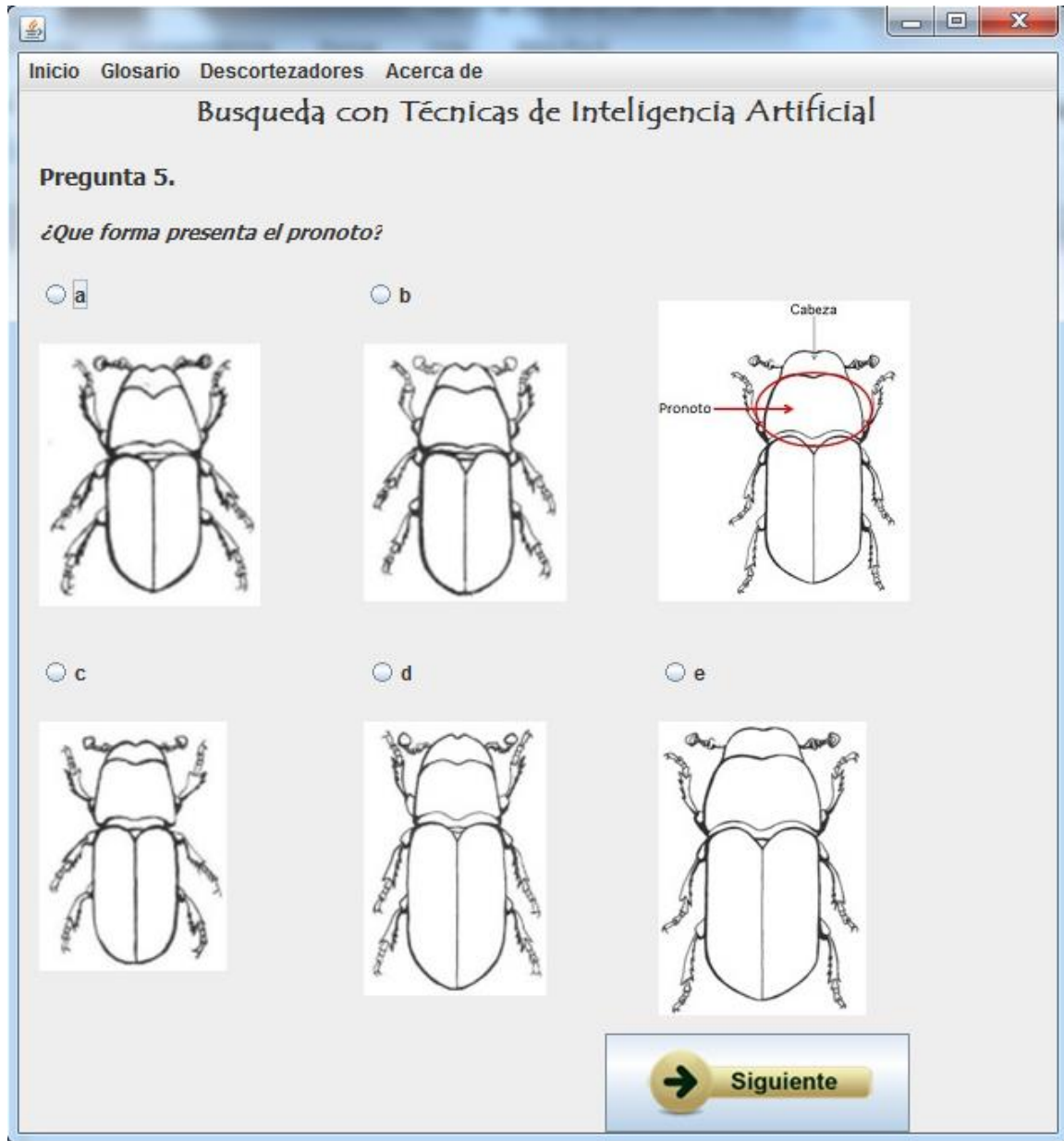


Más de dos tamaños Dos tamaños Abundantes y no saliendo de montículos

 Siguiente

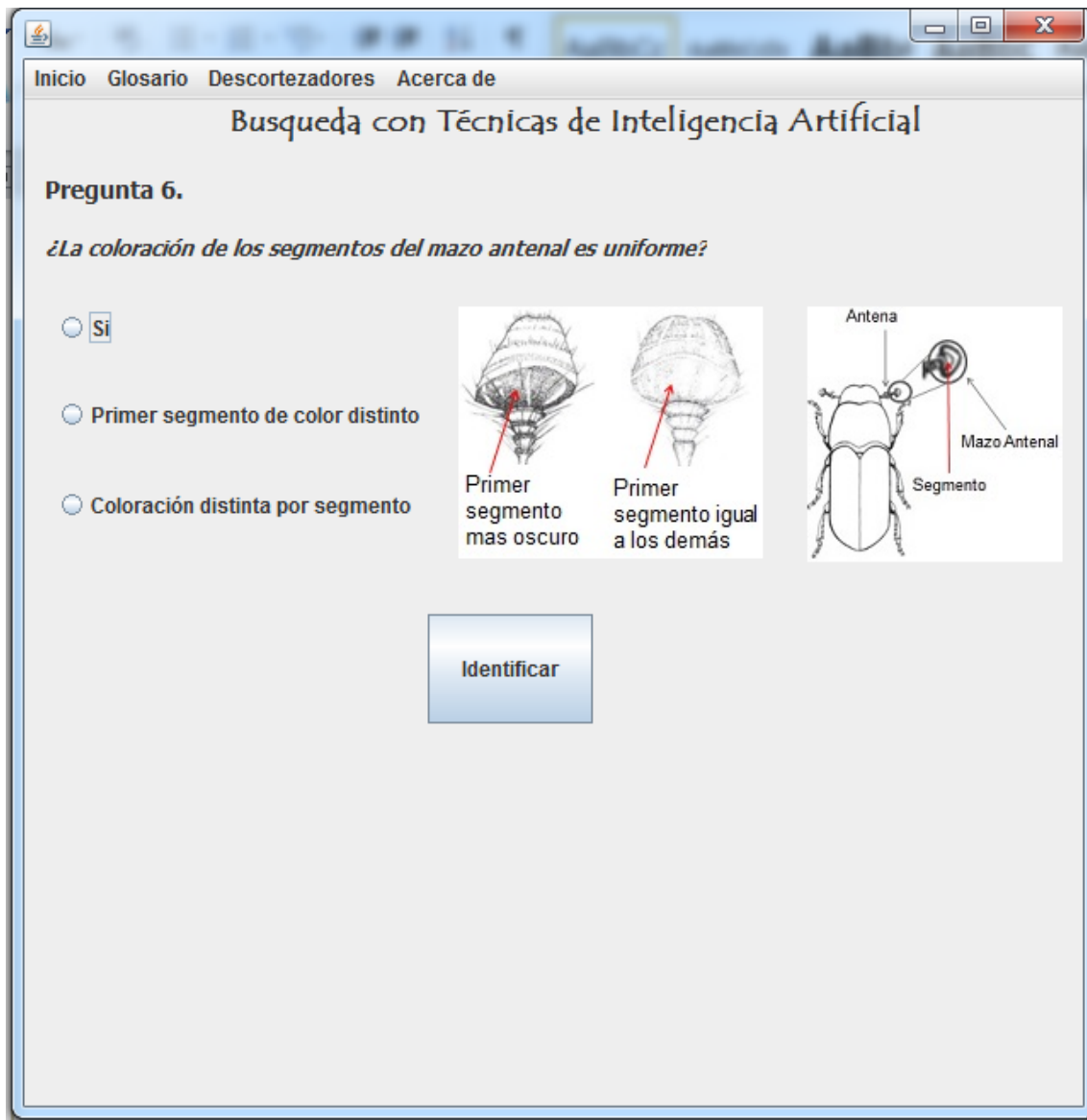
Fuente: Elaboración propia.

Figura 39. Interfaz Módulo IA: Pregunta 5



Fuente: Elaboración propia.

Figura 40. Interfaz Módulo IA: Pregunta 6



Fuente: Elaboración propia.

11.7.4. Resultados del Experimento y Evaluación de las Técnicas

Con cada técnica (RNFF, CB y SVM) se realizaron 10 pruebas para obtener el modelo que arrojara mejores resultados, esto para minimizar el error que se genera por la aleatoriedad de los datos de entrenamiento. En cada una de ellas se hicieron 20 repeticiones con diferentes conjuntos de datos generados aleatoriamente considerando 6 características, con una proporción del 80% de datos de entrenamiento y 20% de datos de prueba, como se aprecia en la Tabla 18.

Tabla 18. Registros por especie

Clase	Especie	Registros entrenamiento (80%)	Registros prueba (20%)	Total registros
1	D. adjunctus	60	15	75
2	D. brevicomis	94	24	118
3	D. frontalis	75	19	94
4	D. mexicanus	109	27	136
5	D. valens	111	28	139
Total		449	113	562

Fuente: Elaboración propia.

Tabla 19. Matriz de confusión RNFF

Clase	1	2	3	4	5	Totales
1	8	1	1	0	1	11
2	5	21	3	1	0	30
3	2	1	15	4	0	22
4	0	1	0	22	3	26
5	0	0	0	0	24	24
Totales	15	24	19	27	28	113

Fuente: Elaboración propia.

Tabla 20. Matriz de Confusión CB

Clase	1	2	3	4	5	Totales
1	7	0	0	1	1	9
2	1	22	2	1	0	26
3	3	1	16	3	0	23
4	4	1	1	22	0	28
5	0	0	0	0	27	27
Totales	15	24	19	27	28	113

Fuente: Elaboración propia.

Debido a que en la SVM se obtiene un modelo que contempla un conjunto de 20 clasificadores binarios, la matriz de confusión solo contiene dos clases que son el promedio de los resultados obtenidos por cada clasificador binario.

Tabla 21. Matriz de Confusión SVM

Clase	Correcta	Incorrecta	Totales
Correcta	55	1	56
Incorrecta	2	55	57
Totales	57	56	113

Fuente: Elaboración propia.

La precisión general y el índice kappa fueron obtenidos calculando el promedio por cada clasificador binario de ambas, dando un total de 97.3498% y 0.9620 respectivamente. La Tabla 21 es un promedio de las 20 matrices de confusión generadas por cada clasificador binario.

Tabla 22. Precisiones por clase

Clase Técnica	Clase 1	Clase 2	Clase 3	Clase 4	Clase 5	% Global
RNFF	53.3334	87.5000	78.9474	81.4815	85.7143	77.3953
CB	46.6667	91.6667	84.2105	81.4815	96.4286	80.0908
SVM	97.6237	95.7256	95.4336	97.9663	100.0000	97.3498

Fuente: Elaboración propia.

Para todas las técnicas se generó la matriz de confusión, con la finalidad de calcular el porcentaje de precisión general (Tabla 22) y el índice kappa (Tabla 23) con los que los modelos fueron evaluados.

Tabla 23. Índice kappa por técnica

Técnica	Kappa
RNFF	0.7425
RB	0.7866
SVM	0.9620

Fuente: Elaboración propia.

11.7.5. Discusión de los Resultados del Experimento

Se decidió partir de cuestionarios sobre seis características físicas de los insectos, que pueden ser observadas en el campo con una lupa, esto con la intención que personas que no tienen experiencia en la identificación de insectos, puedan aportar las observaciones de las seis características y con ello hacer la identificación. Es de destacar, que algunas especies miden solo 2 mm, por ello se justifica el uso de la lupa para observar las características de los insectos y evitar el inconveniente de la similitud entre especies de insectos, sin embargo, como se puede presentar fallas en las observaciones en la práctica se utilizaron técnicas de inteligencia artificial para considerar el error que puedan ocasionar estas inconsistencias.

El modelo de RNFF con mayor precisión fue de 77.3953% con un valor de índice kappa de 0.7425 mostrando que el grado de acuerdo entre las observaciones que hicieron los usuarios y las predicciones de la RNFF para el conjunto de datos de prueba es aceptable, pero no relevante en este experimento. Observando la matriz de confusión (Tabla 19) el modelo presentó problemas en la clasificación de insectos para la Clase 1 y 3 debido a que la apariencia física de ambas especies es similar, esto mismo ocasionó desconcierto a los usuarios cuando respondieron el cuestionario; además de que solamente 2 de 6 características evaluadas permiten diferenciar las clases comentadas. Con esta técnica de identificación, otros autores han reportado mejores resultados, Al-Saqer & Hassan (2011) reporta 93.5% precisión en la identificación del gorgojo de palmera roja; Leow et. al. (2015) obtuvo un 93.13% para identificar copépodos; Wang et. al. (2012) logra una precisión de 93% para identificar especies de insecto de acuerdo al orden al que pertenecen; Do, Harp, & Norris (1999) alcanzaron el 81% de precisión identificando especies de arañas. Las bajas precisiones obtenidas en el experimento con RNFF pudieron haber sido inducidas por las observaciones erróneas que realizaron los encuestados y no se tuvieron mejores datos para que la RNFF tuviera un mejor conjunto de datos de entrenamiento.

El desempeño del modelo CB tuvo una precisión general promedio de 80.0908% de las clases clasificadas correctamente y el índice kappa de 0.7866, al igual que la RNFF el índice kappa reveló que el grado de acuerdo entre las observaciones que hicieron los usuarios y las predicciones del clasificador para el conjunto de datos de prueba fue aceptable pero no para los objetivos de este experimento. Al igual que la RNFF el CB presentó problemas para identificar de forma correcta la clase 1, con un total 46.6% de aciertos, para las otras clases los porcentajes de clasificación superaron el 81%, ver Tabla 20, las bajas precisiones de la red bayesiana se deben principalmente, al igual que en RNFF a las fallas de los encuestados ya que las probabilidades de tener una aserción fueron muy bajas.

En la identificación de insectos con la técnica SVM para el reconocimiento de las 5 especies de insecto, se tenía una multi clasificación pero se optó usar una clasificación uno contra uno, convirtiéndose en una clasificación binaria clásica. Se hizo de esta forma debido a que si se manejaba una clasificación de uno contra todos, al momento de generar los conjuntos de datos la proporción de observaciones se sesgaba hacia la clase negativa, por lo que la identificación se inclinaba a respuestas erróneas, dando una tasa de precisión baja, esto debido a que la cantidad de observaciones por clase era de 1 contra 4. La técnica uno a uno la utilizo Hongtao et. al. (2011) mostrando buenos resultados, para este trabajo se usó la función de Base Radial en el kernel y la configuración predeterminada de Matlab para la creación de la SVM, se obtuvieron un total de 20 clasificadores binarios, cuatro por cada clase.

Con la técnica SVM se obtuvo una clasificación global de 97.3498% y un índice Kappa de 0.9620. Estos valores indican que la SVM supero sustancialmente las identificaciones que se hicieron con RNFF y CB, además que en todas las clases, la identificaciones individuales tienen un reconocimiento de más del 95%, comparándolas con las de RNFF y CB que la clase 1 no superaron el 54%. Se considera que la SVM fue el modelo con mejores resultados, porque, aunque no todas las clases tengan la misma cantidad de observaciones, permite tener una mejor proporción de las observaciones para realizar el entrenamiento. La precisión obtenida de 97.3498%, está dentro del rango reportado en diversos trabajos relacionados; Wang et. al. (2012) obtuvo una precisión del 95% utilizando 6 características para clasificar insectos de acuerdo al orden al que pertenecen; Qing et. al. (2012) identifico 4 especies del género Lepidoptera que atacan cultivos de arroz con una precisión del 97.5%; Diello et. al. (2016) hace un reconocimiento de enfermedades de soja con una precisión mayor del 98%; Hongtao et. al. (2011) identifica 15 especies de insectos que atacan grano almacenado utilizando 7 características con una precisión del 94.67%; por su parte Ebrahimi et. al. (2017) logró identificar parásitos en invernaderos incorporando técnicas de procesamiento de imágenes con SVM con menos del 2.5% de error.

En las tres técnicas la clase 5 no mostró mayor problema de clasificación, esto se debe a que el insecto presenta una coloración rojiza que lo diferencia de las demás especies, caso contrario para la clase 1 la RNFF y el CB presentaron conflicto en clasificarla debido a que el insecto es similar a la clase 2 y 3.

12. Discusión General

Las tres pruebas realizadas en este trabajo arrojan resultados por arriba del 85% de precisión. Pero para efectos de este trabajo el procesamiento de imágenes por ahora

queda descartado como una opción de reconocimiento de los insectos descortezadores debido a que fue el resultado más bajo con un 86.2% de precisión utilizando una RNFF, cuando trabajos similares reportan porcentajes arriba del 95%. El conjunto de imágenes que se obtuvo se podría probar con otras técnicas de clasificación para observar si el porcentaje de clasificación incrementa.

Por otra parte la aplicación de un cuestionario con pocas preguntas de fácil entendimiento para los usuarios, ayudó a generar un conjunto de datos con las respuestas dadas por los usuarios. A partir del conjunto de datos se entrenó una RNFF, un CB y una SVM, de los tres clasificadores el que arrojó mayor porcentaje de precisión fue la SVM con un 97.34%, este porcentaje fue alcanzado por el entrenamiento uno contra uno que realizó la SVM permitiendo crear una mayor separabilidad entre especies.

La clasificación a través de reglas de propagación con encadenamiento hacia adelante permite que la identificación de los insectos sea más interactiva ya que se creó una interfaz en Java para que el usuario pudiera contestar una serie de preguntas que alimentara la memoria de trabajo, con las cuales se logrará realizar la inferencia de las especies de insecto descortezador que cubrían los valores de las características seleccionadas por los usuarios.

13. Conclusiones

La identificación de insectos descortezadores se puede llevar a cabo con técnicas de inteligencia artificial. En el caso del procesamiento de imágenes, se debe mejorar el conjunto de imágenes utilizando equipo especializado para capturar los insectos, debido a que como se observó en este trabajo el tamaño que presentan dificultó obtener buenas fotografías para lograr un buen entrenamiento.

Para lograr una mejor identificación de los insectos descortezadores es recomendable realizarla en laboratorio debido a que se pueden controlar mejor las condiciones para observarlos y a su vez se contar con el equipo adecuado, comparado cuando se realiza en las inspecciones de campo aunque si se realizara en este tipo de escenarios se podría agilizar el tiempo de identificación.

14. Trabajos futuros

Para la visión por computadora:

- Captura de imágenes

- Utilizar una cámara fotográfica para microfotografía, o en su defecto utilizar un estereoscopio con mayor acercamiento para lograr enfocar las características más pequeñas del insecto.
- Mejorar la iluminación hacia el insecto.
- Contar con más especímenes por insecto que se encuentren completos.
- Obtener fotografías de otras posiciones del insecto.
- Diseñar una base que permita posicionar los insectos
- Mejorar la segmentación del insecto, estableciendo áreas de interés.
- Aplicar otras técnicas para mejorar las imágenes de los insectos.

Para el sistema experto:

- Reglas
 - Colocar una opción que permita agregar, editar y eliminar reglas.
 - Agregar otros géneros de insectos descortezadores.
- Interface
 - Mejorar la forma en que se presenta la información del insecto.
 - Crear una interfaz interactiva con animación.
- Funcionamiento
 - Agregar una opción que muestre los tratamientos.
 - Crear base de datos para guardar registro de cada identificación, con el fin de poder desplegar un mapa información de cada insecto, considerando datos como coordenadas, datos climáticos, etc.

15. Bibliografía

Al-Saqer, S. M., & Hassan, G. M. (2011). Artificial neural networks based red palm weevil (*Rynchophorus ferrugineus* Olivier) recognition system. *American Journal of Agricultural and Biological Sciences*, 6(3), 356-364.

Alvarado, J., & Fernández, J. (2012). Análisis de textura en imágenes a escala de grises, utilizando patrones locales binarios (LBP). *ENGI Revista Electrónica De La Facultad De Ingeniería*.

Amardeep, R. (2017). Training Feed Forward Neural Network with Backpropagation Algorithm. *International Journal of Engineering and Computer Science*, 6(1), 19860-19866.

Badaro, S., Ibañez, L. J., & Agüero, M. J. (2013). Sistemas Expertos: Fundamentos, Metodología y Aplicaciones. *Revista Ciencia y Tecnología*(13), 349-364.

- Blanco, E., Bonet, J., & Eizaguirre, M. (2009). Uso de imágenes satélite Landsat para la detección de rodales *Pinus nigra* Arn. y *Pinus sylvestris* L. afectados por escolítidos. *Investigación Agraria: Sistemas y Recursos Forestales*, 18(3), 264-275.
- Bula, H., Aramendiz, H., Salas, D., Vergara, W., & Villadiego, A. (2012). SISTEMA EXPERTO PARA EL DIAGNÓSTICO DE PLAGAS Y ENFERMEDADES EN LOS CULTIVOS DE BERENJENA (*Solanum melongena* L.) EN LA REGIÓN CARIBE DE COLOMBIA. *Revista Ingeniería e Innovación*, 1(1).
- Cáceres, C., Amaya, D., & Ramos, O. (2015). Procesamiento de imágenes para reconocimiento de daños causados por plagas en el cultivo de *Begonia semperflorens* (flor de azúcar). *Acta Agronomica*, 64(3), 273-279.
- Castellanos Bolaños, J. F., Martínez Ruiz, E. O., Cárdenas Gómez, M., & Cubas González, R. (2013). *Guía metodológica para combatir plagas de descortezadores de pino en el sur de México*. Oaxaca: Códigos Editorial.
- Castillo, E., Cobo, Á., Gutiérrez, J. M., & Pruneda González, R. E. (1999). *Introducción a las Redes Funcionales con Aplicaciones*. Madrid: Paraninfo.
- Castrillón, O. D., Giraldo, J. A., & Sarache, W. A. (2008). Sistema de Clasificación Bayesiano basado en Múltiples Clases Omar. *Sistemas, Cibernética e Informática*, 5(1), 25-28.
- Cibrian Tovar, D., Méndez Montiel, T., Campos Bolaños, R., Yates III, H., & Flores Lara, J. (1995). Insectos que se alimentan de floema y corteza. En *Insectos Forestales de México* (págs. 266-353). México: Universidad Autónoma Chapingo.
- Cibrian Tovar, D., Méndez Montiel, T., Campos Bolaños, R., Yates III, H., & Flores Lara, J. (1995). Insectos que se alimentan de FLoema y Corteza. En *Insectos Forestales de México* (págs. 266-353). México: Universidad Autónoma Chapingo.
- Ciesla, W. M., Cadahía, D., & Robredo, F. (1976). La detección de plagas de insectos y enfermedades forestales. *Bol. Serv. Plagas*, 37-53.
- Comisión Nacional Forestal. (2013). *Programa Nacional de Sanidad Forestal 2013-2018*.
- CONAFOR. (2013). *Programa Nacional Forestal 2013-2018*.
- Congalton, R. G., & Green, K. (2009). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. New York: CRC Press.

- Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 273-297.
- Cristianini, N., & Schölkopf, B. (2002). Support Vector Machines and Kernel Methods. *AI Magazine*, 23(3), 31-42.
- Cuevas, E., Zaldívar, D., & Pérez, M. (2010). *Procesamiento digital de imágenes con MATLAB y Simulink*. México: Alfaomega.
- Deschamps, P. (2016). *Nota Informativa 46: Plagas y enfermedades forestales en México: urgen políticas públicas para fomentar la acción de los dueños y habitantes de bosques y selvas*. CCMSS.
- Diello, R., Nunes, D., Margarido, J., Sanches, W., Rodrigues, J., Brandoli, B., y otros. (2016). Local descriptors for soybean disease recognition. *Computers and Electronics in Agriculture*(125), 48-55.
- Diello, R., Nunes, D., Margarido, J., Sanches, W., Rodrigues, J., Brandoli, B., y otros. (2016). Local descriptors for soybean disease recognition. *Computers and Electronics in Agriculture*, 48-55.
- Do, M., Harp, J., & Norris, K. (1999). A test of a pattern recognition system for identification of spiders. *Bulletin of Entomological Research*(89), 217-224.
- Durán, E., & Poloni, A. (2015). Escarabajos descortezadores: diversidad y saneamiento en bosques de Oaxaca. *Biodiversitas*, 7-12.
- Durkin, J. (1994). *Expert systems: desing and development*. New York: Macmillan.
- Ebrahimi, M. A., Khoshtaghaza, M. H., Minaei, S., & Jamshidi, B. (2017). Vision-based pest detection based on SVM classification method. *Computers and Electronics in Agriculture*(137), 52-58.
- Ebrahimi, M. A., Khoshtaghaza, M. H., Minaei, S., & Jamshidi, B. (2017). Vision-based pest detection based on SVM classification method. *Computer and Electronics in Agriculture*, 137, 52-58.
- Freeman, J., & Skapura, D. (1993). *Redes Neuronales: Algoritmos, Aplicaciones y Técnicas de Programación*. Paris: Addison-Wesley.
- Friedman, N., Geiger, D., & Goldszmidt, M. (1997). Bayesian Network Classifiers. *Machine Learning*, 29, 131-163.

- García, O., & Alegre, E. (s.f.). Capítulo 7. Descripción de textura en imágenes utilizando Local Binary Pattern (LBP). En *Conceptos y métodos en Visión por Computador* (págs. 115-130).
- Gernandt, D. S., & Pérez De La Rosa, J. A. (2014). Biodiversidad de Pinophyta (coníferas) en México Biodiversity. *Revista Mexicana de Biodiversidad*, 85, 126-133.
- Gonzalez, R. C., & Woods, R. E. (2002). *Digital Image Processing*. New Jersey: Pearson.
- Guerra, A. (2011). *Universidad Veracruzana*. Recuperado el 04 de 09 de 2017, de <https://www.uv.mx/aguerra/documents/2011mpi-02.pdf>
- Hagan, M. T., Demuth, H. B., & Beale, M. H. (1995). *Neural Network Desing*. Boston Massachusetts: PWS.
- Haralick, R. M., Shanmugan, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6), 610-621.
- Harwood, D., Ojala, T., Pietikäinen, M., Kelman, S., & Davis, S. (1993). *Texture classification by center-symmetric auto-correlation, using Kullback discrimination of distributions*. University of Maryland, Computer Vision Laboratory, Center for Automation Research, Maryland.
- Hilera, J. R., & Martinez, V. J. (2000). *Redes Neuronales Artificiales*. Madrid, España: Alfaomega-Rama.
- Hongtao, Z., Shuping, Y., & Yuxia, H. (2011). SVM Classifier of Stored-Grain Insects Based on Grid Search. *CSE 2011, Part II, CCIS 202*, 436-441.
- Hongtao, Z., Shuping, Y., Yuxia, H., & Power, E. (2011). SVM Classifier of Stores-Grain Insects Based on Grid Search. En Z. M., & T. H., *Advances in Computer Science and Education Applications. Communications in Computer and Information Science* (Vol. 202, págs. 436-441). Berlin, Heidelberg: Springer.
- Hu, M. K. (1962). Visual Pattern Recognition by Moment Invariants. *IRE Transactions on Information Theory*, 179-187.
- Kaloudisa, S., Anastopoulos, D., Yialouris, C. P., Lorentzos, N. A., & Sideridis, A. B. (2005). Insect identification expert system for forest protection. *Expert Systems and Applications*, 28, 445-452.

- Larios, N., Deng, H., Zhang, W., Sarpola, M., Yuen, J., Paasch, R., y otros. (2008). Automated insect identification through concatenated histograms of local appearance features: feature vector generation and region detection for deformable objects. *Machine Vision and Applications*, 19, 105-123.
- Leow, L. K., Chew, L.-L., Chong, V. C., & Dhillon, S. K. (2015). Automated identification of copepods using digital image processing and artificial neural network. *BMC Bioinformatics*, 16(Suppl 18), S4.
- Magallón, C. (2013). *Sanidad Forestal*. Guadalajara: Impre-Jal., S.A. de C.V.
- Martineau, M., Conte, D., Raveaux, R., Arnault, I., Munier, D., & Venturini, G. (2017). A survey on image-based insect classification. *Pattern Recognition journal*, 65, 273-284.
- Mirov, N. T. (1967). *The genus Pinus*. New York: The Ronald Press.
- Morales, E. (2017). *Instituto Nacional de Astrofísica, Óptica y Electrónica*. Recuperado el 04 de 09 de 2017, de <https://ccc.inaoep.mx/~emorales/Cursos/Prolog/curso.pdf>
- Nascimento, D., Anunciação, R., Arnhold, A., Ferraz, A., Dos Sandos, A., & Cola, J. (2016). Expert system for identification of economically important insect pests in commercial teak plantations. *Computers and Electronics in Agriculture*.
- Ochoa, L., & Kemper, N. (2009). Diseño y desarrollo de un sistema inteligente para el diagnóstico y monitoreo de plagas y enfermedades en frutas y hortalizas. *Octava Conferencia Iberoamericana en Sistemas, Cibernética e Informática [and] 6to Simposium Iberoamericano en Educación, Cibernética e Informática*. Orlando, Florida, EE. UU.
- Otsu, N. (January de 1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-9(1), 62-66.
- Pacheco, A. (s.f.). *Instituto Tecnológico de Chihuahua*. Recuperado el 03 de 09 de 2017, de <http://www.depi.itch.edu.mx/apacheco/ai/repconoc.htm>
- Penagos, D. (2004). *Universidad Nacional de Colombia*. Recuperado el 04 de 09 de 2017, de http://dis.unal.edu.co/~fgonza/courses/2004-I/AI/PRACTICAS_DE_PROLOG.html
- Qing, Y. A., Qing-jie, L. I., Guang-qiang, D., Bao-jun, Y., & Hong-ming, C. (2012). An Insect Imaging System to Automate Rice Light-Trap Pest Identification YAO. *Journal of Integrative Agriculture*, 11(6), 978-985.

- Qing, Y. A., Qing-jie, L. I., Guang-qiang, D., Bao-jun, Y., Hong-ming, C., Jun, L. V., y otros. (2012). An Insect Imaging System to Automate Rice Light-Trap Pest Identification. *Journal of Integrative Agriculture*, 11(6), 978-985.
- Raunch, W. (1989). *Aplicaciones de la inteligencia artificial en la actividad empresarial, la ciencia y la industria, (Fundamentos y Aplicaciones)*. Madrid: Díaz de Santos.
- Salinas, Y., Vargas, C., Zuñiga, G., Victor, J., Ager, A., & Hayes, J. (2010). *ATLAS DE DI STRI BUCI ÓN GEOGRÁFI CA DE LOS DESCORTEZADORES DEL GÉNERO Dendroctonus (Curculionidae: Scolytinae) EN MÉXICO*.
- Sánchez, A. (2008). Una visión actual de la diversidad y distribución de los pinos de México. *Madera y Bosques*, 14(1), 107-120.
- SEMARNAT. (28 de 04 de 2014). *Diario Oficial de la Federación*. Recuperado el 17 de 10 de 2016, de Programa Nacional Forestal 2014-2018: http://www.dof.gob.mx/nota_detalle.php?codigo=5342498&fecha=28/04/2014
- SEMARNAT. (2015). *Anuario Estadístico de la Producción Forestal 2014*. Coyoacán, México D.F.
- SEMARNAT. (2016). *Anuario Estadístico de Producción Forestal 2015*. Coyoacán, México D.F.
- SEMARNAT. (2017). *Anuario Estadístico de Producción Forestal 2016*. Coyoacán, México D.F.
- Shiguihara, P., & Valverde, J. (2009). SEDFE: Un Sistema Experto para el Diagnóstico Fitosanitario del Espárrago usando Redes Bayesianas. *Ciencia y Tecnología de la Universidad de Palermo*.
- Shin, B., Russell, J., & Klette, R. (2012). Feature Extraction and Classification for Insect Footprint Recognition. *CIARP 2012, LNCS 7441*, 196-203.
- Sossa, H., Vázquez, R., & Barrón, R. (2005). Reconocimiento y localización de instrumental médico usando análisis automatizado de imágenes. *Ingeniería Biomédica*, 26, 75-85.
- Sucar, L. (s.f.). *Clasificadores Bayesianos: de Datos a Conceptos*. Tonantzintla, Puebla, México.
- Sucar, L. (s.f.). *Redes Bayesianas*. Tonantzintla, Puebla, México.

- Torres, L. (s.f.). *Universidad Nacional de Colombia*. Recuperado el 04 de 09 de 2017, de <http://disi.unal.edu.co/~lctorress/iartificial/IA0005l.pdf>
- Wang, J., Lin, C., Ji, L., & Liang, A. (2012). A new automatic identification system of insect images at the order level. *Knowledge-Based Systems*, 33, 102-110.
- Wang, J., Lin, C., Ji, L., & Liang, A. (2012). Knowledge-Based Systems A new automatic identification system of insect images at the order level. *Knowledge-Based Systems*(33), 102-110.
- Wen, C., & Guyer, D. (2012). Image-based orchard insect automated identification and classification method. *Computers and Electronics in Agriculture*, 89, 110-115.
- Wood, S. (1982). *The Bark and Ambrosia Beetles of North and Central America (Coleoptera: Scolytidae), a Taxonomic Monograph*. Utah: The Great Basin Naturalist.
- Xie, C., Zhang, J., Li, R., Li, J., Hong, P., Xia, J., y otros. (2015). Automatic classification for field crop insects via multiple-task sparse representation and multiple-kernel learning. *Computers and Electronics in Agriculture*, 119, 123-132.

16. Anexos

Anexo 1

Notación con una regla por especie de insecto:

//Género Dendroctonus

IF

declive elitral no presenta espinas AND

color de cuerpo es rojo AND

coloración del primer segmento del mazo antenal es más oscuro que los otros

THEN

D. rhizophagus

IF

declive elitral no presenta espinas AND

color de cuerpo es rojo AND
coloracion del mazo antenal es del mismo color

THEN

D. valens

IF

declive elitral no presenta espinas AND
color de cuerpo es de cafe oscuro a negro AND
setas de declive elitral son de un mismo tamaño

THEN

D. brevicomis

IF

declive elitral no presenta espinas AND
color de cuerpo es de cafe oscuro a negro AND
setas de declive elitral son de dos tamaños

THEN

D. frontalis

IF

declive elitral no presenta espinas AND
color de cuerpo es de cafe oscuro a negro AND
setas de declive elitral son de mas de dos tamaños

THEN

D. mexicanus

IF

declive elitral no presenta espinas AND
color de cuerpo es de cafe oscuro a negro AND
setas del declive elitral son grandes y escasos AND
setas del declive elitral salen de montículos

THEN

D. adjunctus

IF

declive elitral no presenta espinas AND
color de cuerpo es de café oscuro a negro AND
mitad basal del pronoto es mas ancha AND
mitad basal del pronoto casi con lados paralelos AND
sin presencia de callo pronatal AND
sin presencia de tubérculo frontal

THEN

D. parallelocolis

IF

declive elitral no presenta espinas AND
color de cuerpo es de café oscuro a negro AND
mitad basal del pronoto es más ancha AND
mitad basal del pronoto casi con lados paralelos AND
presencia de callo pronotal AND
presencia de tubérculo frontal

THEN

D. approximatus

//Género Ips

IF

declive elitral presenta espinas AND
3 espinas en el declive elitral AND
tamaño menor de 5mm AND
suturas del mazo antenal fuertemente arqueadas AND
tercer espina estrecha y capitada

THEN

I. mexicanus

IF

declive elitral presenta espinas AND
4 espinas en el declive elitral AND
tercer espina capitada AND
tuberculo frontal presente AND
tuberculo epistomal presente AND
tubérculo epistomal unido al tubérculo frontal AND
suturas del mazo antenal fuertemente anguladas AND
tamaño entre 4.6 y 5.7 mm

THEN

I. integer

IF

declive elitral presenta espinas AND
4 espinas en el declive elitral AND
tercer espina capitada AND
tubérculo frontal presente AND
tubérculo epistomal presente AND

tubérculo epistomal unido al tubérculo frontal AND
suturas del mazo antenal fuertemente anguladas AND
puntuaciones posteriores a la cima del pronoto son más pequeñas en relación
con las poslaterales del mismo AND
tamaño entre 3.3 y 4.3 mm

THEN

I. pini

IF

declive elitral presenta espinas AND
4 espinas en el declive elitral AND
tercer espina cilíndrica AND
tubérculo frontal presente AND
tubérculo epistomal presente AND
tubérculo epistomal unido al tubérculo frontal AND
tamaño entre 2.9 y 3.4 mm

THEN

I. bonanseai

IF

declive elitral presenta espinas AND
5 espinas en el declive elitral AND
segunda espina separada de la tercer espina AND
tubérculo frontal grande presente AND
sin tubérculo epistomal presente AND
tamaño entre 2.9 y 4.7 mm

THEN

I. grandicollis

IF

declive elitral presenta espinas AND
5 espinas en el declive elitral AND
segunda espina y tercer espina unidas en la parte basal AND
tamaño entre 2.9 y 3.6 mm

THEN

I. cribricollis

IF

declive elitral presenta espinas AND
5 espinas en el declive elitral AND
tercer espina es pequeña y con gancho AND

epistoma presenta linea de tuberculos AND
tamaño entre 4 y 4.7 mm

THEN

I. lecontei

IF

declive elitral presenta espinas AND
5 espinas en el declive elitral AND
distancia entre la primer y segunda espina es aprox. Igual a la distancia de la
segunda y tercer espina AND
tamaño entre 3.5 y 4.2 mm

THEN

I. confusus

IF

declive elitral presenta espinas AND
6 espinas en el declive elitral AND
tubérculo frontal se encuentra separado del margen epistomal por una
distancia menor a dos veces el grosor del tubérculo frontal AND
tamaño entre 3.5 y 5.9 mm

THEN

I. calligraphus