

UNIVERSIDAD CATÓLICA SANTO TORIBIO DE MOGROVEJO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN



**SOFTWARE PARA IDENTIFICAR Y LOCALIZAR DIVERSAS
VARIETADES DE PLANTAS DE QUINUA EN EL DISTRITO DE
QUEROCOTO - CAJAMARCA**

**TESIS PARA OPTAR EL TÍTULO DE
INGENIERO DE SISTEMAS Y COMPUTACIÓN**

AUTOR

CRISTIAN ABELARDO VALLEJOS FONSECA

ASESOR

MSc. Ing. LUIS AUGUSTO ZUÑE BISPO

Chiclayo, 2019

DEDICATORIA

A Dios, por haberme permitido lograr este objetivo, además de su infinita bondad y amor, *a* mis padres por haberme apoyado en todo momento, por sus consejos, sus valores, por la motivación constante que me ha permitido ser una persona de bien, mostrándome para salir adelante ante toda dificultad que se me presente, y sobre todo por su amor incondicional que me dan, y *a mis amigos que me apoyaron en todo momento, hasta ver cumplir mí objetivo.*

AGRADECIMIENTOS

Agradecer a mi asesor Luis Zuñe Bispo por tenerme mucha paciencia y dedicación durante el término de mi tesis, a la empresa FSDG por haberme apoyado constantemente para realizar mi tesis.

RESUMEN

El proyecto desarrolló un software para la identificación y localización de manera rápida de variedades de plantas de quinua Utilizando datos y la inteligencia artificial. Se halló patrones para la identificación de quinua, que permita determinar la similitud con otras de la misma especie, de forma rápida y confiable. El software realizará una búsqueda en los repositorios de datos de la comunidad científica, para luego analizarlos en busca de patrones que los identifique como parte de las diversas variedades de quinua.

Esta investigación apoyó a la comunidad científica ligada a la investigación en agricultura a encontrar variedades de plantas de quinua con características que les permita mejorar y conservar el recurso genético de *Chenopodium Quinoa*.

La comunidad científica que investiga en el campo biológico, cada día se enfrenta a mayores retos que implican el manejo de enormes volúmenes de datos que crecen de manera exponencial en tamaño y complejidad, es así que uno de los problemas que presenta la empresa es de no tener un resultado exacto en el momento de identificar una determinada planta de quinua, hasta hoy se realiza de forma manual por los especialistas, requiriendo tiempo y esfuerzo para su correcta identificación y localización entre sus variedades. Es por ello que se desarrolló un software por medio de un algoritmo de reconocimiento de imágenes que ayudará a los investigadores que requieren procesar grandes volúmenes de información en el menor tiempo posible con la mayor fiabilidad.

El proyecto pretende obtener un software que utilizando datos y la inteligencia artificial, pueda realizar el análisis automático, aportando información a los investigadores que requieren procesar grandes volúmenes de información en el menor tiempo posible con la mayor fiabilidad.

Con base en esto, el presente trabajo consiste en automatizar el proceso de identificación y localización de las plantas de quinua, para la localización se tuvo el apoyo del GPS, y para el proceso de identificación se basó en imágenes fotográficas, utilizando reconocimiento de patrones, teniendo como objetivo general la construcción de un software para la identificación y localización de las diversas variedades de la planta de quinua, utilizando procesamiento de imágenes para lograrlo

Palabras Clave: Quinua, Inteligencia Artificial, Sistema, Software, Tiempo

ABSTRACT

The project will develop software for the rapid identification and localization of varieties of Quinoa plants.

Using data and artificial intelligence, it is tried to find patterns for the identification of quinoa, which allows to determine the similarity with others of the same species, in a fast and reliable way.

The software will search the data repositories of the scientific community, then analyze them for patterns that identify them as part of the various varieties of quinoa.

This research will support the scientific community linked to agricultural research to find varieties of quinoa plants with characteristics that allow them to improve and conserve the genetic resource of *Chenopodium Quinoa*.

The scientific community that researches in the biological field, every day faces greater challenges that involve the handling of huge volumes of data that grow exponentially in size and complexity, so one of the problems presented by the company is not have an exact result when identifying a specific quinoa plant, until today it is done manually by the specialists, requiring time and effort for its correct identification and location among its varieties. That is why software was developed by means of an image recognition algorithm that will help researchers who need to process large volumes of information in the shortest time possible with the highest reliability.

Based on this, the present work consists of automating the process of identification and location of the quinoa plants, for the location was supported by the GPS, and for the identification process was based on photographic images, using pattern recognition, having like general aim the construction of a software for the identification and location of the diverse varieties of the plant of quinoa, using processing of images to obtain it.

KEYWORDS: Quinoa, Artificial intelligence, System, Software, Time

CONTENIDO

RESUMEN

INTRODUCCIÓN

CAPÍTULO I MARCO TEÓRICO 14

2.1 Antecedentes14

2.2 Bases Teórico Científicas17

2.2.1 Metodologías para el desarrollo de sistemas inteligentes expertos17

2.2.2 Inteligencia Artificial19

2.2.3 Sistema Experto20

2.2.4 Visión Artificial21

2.2.5 Histograma32

2.2.6 Reconocimiento de patrones33

2.2.7 Redes Neuronales34

2.2.8 Base de Datos.....36

2.2.9 Lenguaje de Programación37

2.2.10 Cultivo de la Quinoa.....38

2.2.11 Marco Conceptual39

CAPÍTULO II METODOLOGÍA 40

3.1 Tipo y nivel de investigación40

3.1.1 Tipo de investigación40

3.1.2 Nivel de investigación40

3.2 Diseño de investigación40

3.3 Población, muestra y muestreo41

3.3.1 Población.....41

3.3.2 Muestra y Muestreo41

3.4 Criterios de selección42

3.5 Operacionalización de variables42

3.5.1 Variables42

3.5.2 Indicadores43

3.6 Técnicas e instrumentos de recolección de datos43

3.7 Procedimientos43

3.8 Plan de procesamiento y análisis de datos44

3.9 Matriz de consistencia.....47

3.10 Consideraciones Éticas48

CAPÍTULO III RESULTADOS	49
CAPÍTULO IV DISCUSIÓN	89
CONCLUSIONES	93
RECOMENDACIONES	94
REFERENCIAS BIBLIOGRÁFICAS	95
ANEXOS	98
ANEXO N° 1.....	98
ANEXO N° 2.....	99
ANEXO N° 3.....	102

ÍNDICE DE FIGURAS

FIGURA 1: IMÁGENES CON DIFERENTES NIVELES DE GRIS	22
FIGURA 2: PROCESO FILTRADO	23
FIGURA 3: FILTRO PASA-BAJOS: (A) EN FRECUENCIA, (B) EN EL DOMINIO ESPACIAL.....	24
FIGURA 4: MÁSCARA PARA FILTRO GAUSSIAO DE 3X3.....	25
FIGURA 5: FILTRO PASA-ALTO: (A) EN FRECUENCIA, (B) EN EL DOMINIO ESPACIAL	26
FIGURA 6: MÁSCARA DE 3X3 PARA UN FILTRO PASA-ALTO SIMPLE	26
FIGURA 7: MÁSCARA DE 3X3 PARA UN FILTRO CON ÉNFASIS EN LAS ALTAS FRECUENCIAS.....	27
FIGURA 8: LA BINARIZACIÓN EN ESCENAS SENCILLAS MUY CONTRASTADAS COMO LAS OBTENIDAS A CONTRALUZ.....	30
FIGURA 9: UNA MONTAÑA (A) Y SUS 64 CONTENEDORES DEL HISTOGRAMA NIVEL GRIS (B).....	32
FIGURA 10: ESTRUCTURA DE UNA RED NEURONAL MULTICAPA, CON UNA CAPA DE ENTRADA, UNA CAPA DE OCULTA Y UNA CAPA DE SALIDA.....	35
FIGURA 11: BASE DE DATOS	37
FIGURA 12: FASES DE LA METODOLOGÍA DE DESARROLLO DE SISTEMAS INTELIGENTES	45
FIGURA 13: DISEÑO DE ENTRADA DEL SOFTWARE	61
FIGURA 14: INICIO DE SESIÓN PARA LOS USUARIOS.....	62
FIGURA 15: MENÚ PRINCIPAL.....	62
FIGURA 16: MANTENIMIENTO	63
FIGURA 17: MANTENIMIENTO - ABRIR IMAGEN	63
FIGURA 18: MANTENIMIENTO - FILTROS.....	64
FIGURA 19: MANTENIMIENTO - HISTOGRAMA	64
FIGURA 20: MANTENIMIENTO - SEGMENTACIÓN.....	65
FIGURA 21: TRANSACCIONES	65
FIGURA 22: TRANSACCIONES - EXTRACCIÓN DE CARACTERÍSTICAS	66
FIGURA 23: TRANSACCIONES - RECONOCIMIENTO DE IMAGEN	66
FIGURA 24: DEMOSTRACIÓN - INICIO SESIÓN	67
FIGURA 25: ADQUISICIÓN DE IMÁGENES	68
FIGURA 26: IMAGEN FILTRO BLUR.....	69
FIGURA 27: IMAGEN FILTRO GAUSSIAN BLUR.....	70
FIGURA 28: IMAGEN FILTRO BILATERAL	71
FIGURA 29: IMAGEN MEDIAN BLUR.....	72
FIGURA 30: IMAGEN HISTOGRAMA.....	72
FIGURA 31: IMAGEN SEGMENTACIÓN.....	73
FIGURA 32: IMAGEN DE EXTRACCIÓN DE CARACTERÍSTICAS.....	74
FIGURA 33: EXTRACCIÓN DE VECTOR DE CARACTERÍSTICAS GEOMÉTRICAS BÁSICAS...	74
FIGURA 34: RED NEURONAL BACK PROPAGATION	77
FIGURA 35: TABLA PATRONES_RECONOCER	78
FIGURA 36: TABLA DE PATRONES RECONOCER.....	79
FIGURA 37: PORCENTAJE DE PLANTAS DE QUINUA LOCALIZADAS.	83
FIGURA 38: PORCENTAJE DE PLANTAS DE QUINUA LOCALIZADAS.	84
FIGURA 39: PORCENTAJE DE PLANTAS DE QUINUA IDENTIFICADAS.....	87
FIGURA 40: PORCENTAJE DE PLANTAS DE QUINUA IDENTIFICADAS.....	88
FIGURA 41: DEMOSTRACIÓN DE GPS DE UNA IMAGEN.....	90
FIGURA 42: TABLA PLANTA	90

ÍNDICE DE TABLAS

TABLA 1: TABLA DE DESCRIPTOR.....	34
TABLA 2: CARACTERÍSTICAS DE MYSQL	37
TABLA 3: TIPIFICACIÓN DE LA POBLACIÓN	41
TABLA 4: INDICADORES.....	43
TABLA 5: MÉTODOS Y TÉCNICAS DE RECOLECCIÓN DE DATOS.	43
TABLA 6: MATRIZ DE CONSISTENCIA	47
TABLA 7: ASUNTOS DE VIABILIDAD DEL PROBLEMA	50
TABLA 8: ASUNTOS DE VIABILIDAD DEL EXPERTO	50
TABLA 9: ASUNTO DE VIABILIDAD DEL DESPLIEGUE	51
TABLA 10: VIABILIDAD DEL PROYECTO	51
TABLA 11: TABLA DE COSTO.....	52
TABLA 12: TABLA DE BENEFICIO	52
TABLA 13: TABLA DE INDICADORES PARA ANÁLISIS	56
TABLA 14: RECOLECTAR CONOCIMIENTO ADICIONAL.....	56
TABLA 15: HISTORIA USUARIO - ABRIR IMAGEN	57
TABLA 16: HISTORIA USUARIO - FILTROS	57
TABLA 17: HISTORIA USUARIO - HISTOGRAMA DE LA IMAGEN	58
TABLA 18: HISTORIA USUARIO – SEGMENTACIÓN	58
TABLA 19: HISTORIA USUARIO - EXTRAER CARACTERÍSTICAS.....	59
TABLA 20: HISTORIA USUARIO - RECONOCIMIENTO	59
TABLA 21: HISTORIA USUARIO - INICIO SESIÓN.....	60
TABLA 22: HISTORIA USUARIO - CERRAR SESIÓN	60
TABLA 23: DISMINUIR LOS ERRORES DE COMPARACIÓN.....	76
TABLA 24: REDUCIR EL TIEMPO DE IDENTIFICACIÓN	77
TABLA 25: PLANTAS IDENTIFICADAS	78
TABLA 26: COMPOSICIÓN DE LA MUESTRA DE LAS PLANTAS DE QUINUA	80
TABLA 27: PRUEBA T PARA DOS MUESTRAS SUPONIENDO VARIANZAS IGUALES.....	81
TABLA 28: PRUEBA T PARA DOS MUESTRAS SUPONIENDO VARIANZA RELACIONADAS.	82
TABLA 29: PORCENTAJE DE LA LOCALIZACIÓN DE LAS PLANTAS DE QUINUA.	83
TABLA 30: PRUEBA T PARA DOS MUESTRAS SUPONIENDO VARIANZAS IGUALES. IDENTIFICACIÓN DE LAS DIVERSAS VARIEDADES DE PLANTAS DE QUINUA.....	85
TABLA 31: PRUEBA T PARA DOS MUESTRAS SUPONIENDO VARIANZAS RELACIONADAS.	86
TABLA 32: PORCENTAJE DE LA IDENTIFICACIÓN DE LAS PLANTAS DE QUINUA.....	87
TABLA 33: OBJETIVOS	89

INTRODUCCIÓN

La quinua es una planta que tiene su origen en la región andina de Sudamérica, esta planta cuyo cultivo primordial es para suplir las necesidades alimenticias de los seres humanos, se utiliza como componente en la dieta de los astronautas, y presenta un importante aporte de proteínas e hidratos de carbono [1], además de tener importantes aplicaciones en la medicina como elemento base de bandas cicatrizantes, y actualmente en investigaciones recientes como material para la construcción de nano robots biodegradables. Cada una de las variedades de la planta de quinua, tiene propiedades específicas, investigaciones caracterizan las propiedades físico-químicas, y estructurales de cada variedad, con la finalidad de no solo lograr la explotación de la misma como un producto para la industria alimentaria, sino también en el área de biomateriales [1].

La empresa Washington Native Plant Society, presenta el problema de la clasificación de hojas, la cual es una tarea que siempre ha estado presente en la labor diaria de los botánicos, debido al gran volumen de familias y clases que existen en el ecosistema y a las nuevas especies que van apareciendo. En las últimas décadas, se han desarrollado disciplinas que necesitan de esta tarea apoyada con la visión artificial para poder lograr este objetivo.

La comunidad científica que investiga en el campo de la biología, en el afán de encontrar respuestas a estudios para identificar diferentes variedades de plantas de quinua, cada día se enfrenta a mayores retos que implican el manejo de enormes volúmenes de datos que crecen de manera exponencial, debido a los avances tecnológicos que permiten hacer cálculos más precisos. Afortunadamente, el desarrollo tecnológico tanto en el ámbito de la electrónica como el desarrollo de software y las telecomunicaciones han permitido un avance significativo en las técnicas para el procesamiento y análisis inteligente de los datos, beneficiando los estudios científicos que permiten conocer mejor las estructuras de los organismos vivos [2]. El proceso de identificación de una planta puede tomar de 1 mes a 3 meses aproximadamente ocupando 3 investigadores para esta tarea.

Uno de los problemas es identificar y localizar diferentes variedades de plantas de quinua, de acuerdo a sus similitudes y diferencias entre sus variedades, este procedimiento, hasta hoy se realiza de forma manual por los especialistas, requiriendo tiempo y esfuerzo para una correcta identificación y localización entre sus variedades, las cuales son aproximadamente 16 variedades de plantas de quinua. Esta amplia variedad de plantas es una limitante para realizar las exploraciones y análisis en las amplias bases de datos de variedades de quinua estudiadas en el mundo.

Dado que hoy existen bases de datos con información de variedades estudiadas por investigadores de todo el mundo, que no están siendo aprovechadas de manera óptima con el propósito de preservar, mejorar y aprovechar el germoplasma de la quinua. [3]

Estas variedades son analizadas de acuerdo a la ubicación geográfica generando tiempo de búsqueda de ellas, según su naturaleza, encontrándose entre ellas diversas variedades, las que serán sometidas a manipulación microbiológica con ayuda de la tecnología. Por otro lado se pueden presentar problemas como el error de información, debido a que al identificar y comparar cada especie de quinua pueda suceder que no sea exacta. [4]

La empresa es una asociación dedicada a ejecutar proyectos de desarrollo para el Distrito de Querocoto, con la finalidad de financiar programas de carácter social ejecutados con parte de los recursos obtenidos a través de procesos de promoción de la inversión privada. En este proyecto se establece dos tipos de trabajadores, uno son los técnicos de campo, la cual cumplen la función de recolectar diversas variedades de plantas de quinua en el campo, utilizando un marcado para dicha ubicación. Por otro lado, se encuentra los expertos, que son los encargados de hacer el proceso de identificación de las diversas variedades de plantas de quinua.

El problema seleccionado para el presente trabajo son las deficiencias en la identificación y localización de las variedades de plantas de quinua, en cuanto a los expertos son el excesivo volumen de información a procesar para identificar las variedades plantas de quinua; la pocas horas de capacitaciones para la identificación de la planta de quinua; los errores de comparación de las características de una planta al momento de identificar, y por contar con expertos y la escasa experiencia en cuanto a la identificación, habiendo cometido errores en el momento de la identificación de las

variedades de plantas de quinua, que alguna veces han puesto en problemas a la empresa con sus clientes por el envío equivocado de plantas. En cuanto a los trabajadores de campo se tiene un mal procedimiento de marcado de la ubicación de una determinada variedad de la planta de quinua, lo que trae como consecuencia que exista poca certeza en la ubicación de las variedades de la planta quinua después de haber sido localizadas, algunas veces ya no pueden ser encontradas.

De lo expresado anteriormente ¿De qué manera se pueda contribuir con el proceso de identificación y localización en las diversas variedades de plantas de quinua?, se propone que con un Software para la identificación y localización de las diversas variedades de la planta de quinua se reducirán las deficiencias de identificación y localización de las diversas variedades de quinua.

Se plantea como objetivo general la construcción de un software para la identificación y localización de las diversas variedades de la planta de quinua, con los siguientes objetivos específicos:

- Mejorar el nivel de certeza de los técnicos de campo para localizar las diversas variedades de la planta de Quinua.
- Incrementar el nivel de eficiencia de los expertos para identificar las diversas variedades de la planta de Quinua.

El presente trabajo se justifica en un impacto social el cual beneficiaría a las personas en su alimentación como por ejemplo el programa social Qali Warma, entre otros; ya que contiene propiedades medicinales, la quinua al poder emplearse como un cereal, es especialmente beneficiosa en la dieta de personas celíacas, ya que no contiene gluten. Asimismo, por su alto contenido en fibra y su mayor aporte proteico respecto a los cereales, la quinua tiene un bajo índice glucémico, lo que la vuelve ideal para personas con diabetes o que desean adelgazar comiendo sano.

Sin embargo, en el contexto tecnológico con nuestro software queremos beneficiar en este caso a la empresa donde se va a realizar una mayor exactitud en el momento de identificar las diversas variedades de plantas de quinua, y a su vez poder localizarlas con mayor rapidez, con el apoyo de la inteligencia artificial aportando a los científicos

expertos que requieran procesar grandes volúmenes de información en el menor tiempo y con la posible mayor fiabilidad.

Por otro lado, en lo económico el uso de este software permitirá a los investigadores a tener una herramienta complementaria para la identificación y localización de variedades de plantas de quinua, así poder ser exportadas a diversos países del mundo, sin ser engañados con plantas similares y así no poder ocasionar problemas económicos en la empresa. De esta manera contribuirá a aumentar los ingresos económicos a dichos investigadores para la identificación rápida y fiable de las variedades de plantas de quinua.

CAPÍTULO I

MARCO TEÓRICO

2.1 Antecedentes

Según [5] Este trabajo describe el proceso de extracción de patrones característicos de imágenes, mediante la ayuda de redes neuronales artificiales. La información de la Red Neuronal junto con datos adicionales de las imágenes, serán almacenados en una base de datos y consumidos por un servicio web. Un teléfono móvil con sistema operativo Android consumirá la información almacenada en el servicio web. Posteriormente al realizar una captura de imagen con la cámara del teléfono, este procesará la imagen y junto con los datos consumidos por el servicio web será capaz identificar de qué imagen se trata. Para el tratamiento de las imágenes se utilizarán librerías OpenCV, tanto en el servidor como en el teléfono móvil.

Del trabajo [6] Este libro tiene el objetivo de mejorar el aspecto de las imágenes y hacer más evidentes en ellas ciertos detalles que se desean hacer notar, mediante el procesamiento e interpretación de imágenes, captadas con distintos tipos de sensores, fundamentalmente cámaras, y utilizando la presentación en el computador. El desarrollo de este tipo de sistemas requiere una combinación de etapas de bajo nivel, para mejorar la calidad de las imágenes capturadas en otras etapas, como el reconocimiento de patrones y de la interpretación de imágenes para reconocer los elementos presentes en una escena. Las aplicaciones de esta área son, en general, trascendentales para resolver problemas concretos, con la aplicación de diversas técnicas matemáticas, estadísticas y de inteligencia artificial.

El trabajo de [7] Esta tesis de posgrado tiene la finalidad para el reconocimiento de tipos de plantas a partir de comparaciones de imágenes de hojas digitalizadas, mostrando como resultado el porcentaje de parecido entre la hoja buscada y las almacenadas en la base de datos e información adicional de la especie, relacionada con la hoja, como su nombre científico, una breve descripción, y fotos complementarias tales como: vista delantera y trasera de la hoja, fruto y/o corteza del árbol.

Según [8] En este trabajo se propone el uso de descriptores locales codificados con el método de Vectores de Fisher para resolver el problema de clasificación de hojas de plantas usando imágenes. Se demuestra que con este enfoque se obtienen resultados de estado del arte en conjuntos de datos públicos, sin la necesidad de recurrir a la creación de nuevos descriptores visuales específicos para el problema, ni la necesidad de realizar ningún pre procesamiento de las imágenes a analizar. El problema de clasificación de especies de plantas es un tema al cual se le está prestando una creciente atención en los últimos años. Algunas de las razones son la gran cantidad de especies en peligro de extinción y las altas tasas de deforestación debida al corrimiento de la frontera agropecuaria y a una mala planificación urbana.

Según [9] Este trabajo plantea el desarrollo de un sistema de visión artificial por capturas y en tiempo real, como primera etapa de un macroproyecto para la inspección visual automática de cultivos de frijol, el cual emplea métodos heurísticos para la diferenciación entre los elementos de la imagen. El sistema consta principalmente de cuatro etapas las cuales son: Ambiente controlado, adquisición, procesamiento de imágenes y reconocimiento de patrones. La integración del sistema se encuentra basado en el uso del software MATLAB. Este interactúa con conecta una cámara web para la obtención de imágenes en la etapa de adquisición, con la alternativa de importar imágenes guardadas en el sistema. En MATLAB también se realizan las otras dos etapas, siendo el procesamiento de imágenes el verdadero plus del proyecto, pues integra las funciones más importantes para el procesamiento de imágenes en una interfaz gráfica, integrando en ella un modo automático y un modo manual.

Según [10] Este trabajo plantea el desarrollo de un sistema de visión artificial por capturas y en tiempo real, como primera etapa de un macroproyecto para la inspección visual automática de cultivos de frijol, el cual emplea métodos heurísticos para la diferenciación entre los elementos de la imagen. Este presente investigación se plantea el desarrollo de un sistema de visión artificial por capturas y en tiempo real, como primera etapa de un macroproyecto para la inspección visual automática de cultivos de frijol, el cual emplea métodos heurísticos para la diferenciación entre los elementos de la imagen, Este trabajo plantea el desarrollo de un sistema de visión artificial por capturas y en tiempo real, como primera etapa de un macroproyecto para

la inspección visual automática de cultivos de frijol, el cual emplea métodos heurísticos para la diferenciación entre los elementos de la imagen.

Según [11] Este presente trabajo se ha realizado el análisis micrográfico de hojas de *Smallanthus sonchifolius*, que se comercializan para la elaboración de té con fines medicinales. Se aplicó la técnica de disociado leve con solución acuosa de NaOH al 5% a 100 °C durante 5 min. Los caracteres de valor diagnóstico están representados por dos tipos diferentes de pelos tectores 3-10-celulares, unos cónicos, rígidos y de base ancha y otros flexuosos y con base no ensanchada; los pelos glandulares sésiles, de cabeza pluricelular biseriada, constituyen una referencia complementaria. Estos elementos son útiles para efectuar control de calidad botánico en muestras de droga vegetal muy fragmentada, como se presenta el material en el mercado. Con este objetivo se ha llevado a cabo el presente trabajo en el que se define el patrón de identificación micrográfico de las hojas de “yacón”, cuya comercialización, tanto desde una producción artesanal como industrial, implica una profunda fragmentación de la materia prima.

Según [12] El objetivo del presente trabajo fue identificar especies de nematodos del género *Meloidogyne* asociadas a plantas ornamentales de altura. El estudio se realizó en el cantón de San Isidro de Heredia, Costa Rica, en un vivero comercial durante el periodo 2011-2012, en diez especies de plantas. Para la identificación se utilizaron mediciones morfométricas de la longitud del estilete, tamaño de cola y región hialina de los juveniles en segundo estado de desarrollo. La técnica de Reacción de Cadena de la Polimerasa, mediante Polimorfismos de Longitud de Fragmentos de Restricción (PCR-RFLP), comprende la técnica convencional de PCR junto con el uso de enzimas de restricción posterior a la amplificación. De forma general, esta técnica amplifica un fragmento específico del genoma del individuo y utiliza enzimas de restricción para generar cortes en los fragmentos amplificados; los tamaños generados por estos cortes se comparan con aquellos reportados en la literatura y resulta un sistema altamente efectivo para el proceso de identificación.

Según [13] El objetivo de este trabajo es determinar los resultados de aplicar el sistema al reconocimiento de un conjunto de rostros seleccionados. Aunque para nosotros reconocer caras no supone apenas esfuerzo, el problema es muy difícil de

resolver desde un punto de vista computacional. Cada mínimo cambio de luz ambiental, expresión de la cara, inclinación, cabello, etc. supone una imagen completamente distinta. Naturalmente tratándose de una entidad como la policía en algunos casos es posible adquirir una imagen de acuerdo a las condiciones del software obligando al individuo se ubique en una habitación con buena iluminación o colocado en una posición frente a la cámara que favorezca al proceso de reconocimiento. Es por ello que se va aplicar una técnica diferente (Hopfield) a la solución del problema de identificación de rostros, aprovechando la propiedad de memoria asociativa que posee, la cual parece imitar de mejor manera el comportamiento del cerebro humano, y así poder servir de base para otras investigaciones.

2.2 Bases Teórico Científicas

2.2.1 Metodologías para el desarrollo de sistemas inteligentes expertos

2.2.1.1 Metodología CommonKADS

KADS es una metodología completa para el desarrollo de SBCs. Ha sido el resultado de un trabajo que ha durado aproximadamente una década dentro de dos proyectos ESPRIT. En sus inicios, KADS se centraba en el problema del cuello de botella que se suponía la adquisición de conocimiento, para posteriormente convertirse en una metodología completa para el desarrollo de SBCs. En la actualidad, CommonKADS, nombre que recibe la evaluación de KADS, cubre la gestión del proyecto, el análisis organizacional y los aspectos relativos a la Ingeniería de Software y Conocimientos relacionados con el desarrollo de SBCs. [14]

2.2.1.2 Metodología Scrum

Es un proceso en el que se aplican de manera regular un conjunto de buenas prácticas para trabajar colaborativamente, en equipo, y obtener el mejor resultado posible de un proyecto. Estas prácticas se apoyan unas a otras y su selección tiene origen en un estudio de la manera de trabajar de equipos altamente productivos.

En Scrum se realizan entregas parciales y regulares del producto final, priorizadas por el beneficio que aportan al receptor del proyecto. Por ello, Scrum está especialmente indicado para proyectos en entornos complejos, donde se necesita obtener resultados pronto, donde los requisitos son cambiantes o poco definidos, donde la innovación, la competitividad, la flexibilidad y la productividad son fundamentales. [15]

2.2.1.3 Metodología de Buchanan

La característica más importante de esta metodología es la constante relación entre el ingeniero de conocimiento y el experto del área. Se destacan seis etapas fundamentales [16]:

- Identificación: se identifican los participantes y roles, los recursos, fuentes de conocimiento. Además, se establecen las facilidades computacionales y presupuestos que permitan lograr los objetivos y métodos que también se plantearán en esta fase.
- Conceptualización: Se analizarán los conceptos vertidos por el experto del área. Los conceptos deben de ser tomados en cuenta con sumo interés, pues el experto del área es quien conoce en detalle los fundamentos particulares del tema a investigar.
- Formalización: Se identifican los conceptos relevantes e importantes. El resultado de formalizar el diagrama de información conceptual es una especificación parcial para construir un prototipo de la base de conocimiento.
- Implementación: Se formaliza el conocimiento obtenido del experto y se elige la organización, el lenguaje y el ambiente de programación.
- Testeo: Se observa el comportamiento del prototipo, el funcionamiento de la base de conocimiento y la estructura de las inferencias, verificándose la performance del sistema.
- Revisión del prototipo: Se reformulan los conceptos y se rediseña y refina el prototipo.

2.2.1.4 Metodología de ingeniería del conocimiento de Jhon Durkin

Es importante culminar esta parte mencionando los dos propósitos que tiene la IA y que resume claramente John Durkin: “el propósito de la IA es hacer a las computadoras más útiles para los humanos. Esto puede lograrse produciendo programas de computadora que asistan a los humanos en sus decisiones, búsqueda inteligente de información o simplemente interfaces de comunicación el lenguaje natural. Un segundo propósito de la IA, pero de igual importancia que el anterior, es mejorar el entendimiento de la inteligencia humana. Construir una computadora inteligente requiere que entendamos como los humanos capturan, organizan y usan el conocimiento durante la solución de sus problemas”. [16]

Se ha realizado la comparación de las metodologías mencionadas para el desarrollo de los sistemas expertos se ha determinado que se usará, será la metodología de Jhon Durkin, ya que no solo se caracteriza por tener una interacción con el especialista para poder recolectar, analizar y procesar la información para armar una base de conocimiento, sino que además una de las fases que presenta la más resaltante es la documentación.

2.2.2 Inteligencia Artificial

Dada la identificación del propio concepto de “inteligencia”, prácticamente existe una definición de inteligencia artificial por cada autor que escribe sobre el tema. Tal vez una de las definiciones que se puede considerar más ajustada a la realidad en la Encyclopedia Of Artificial Intelligence.

La IA es un campo de la ciencia y la ingeniería que se ocupa de la comprensión, desde el punto de vista informático, de lo que denomina comúnmente comportamiento inteligente. También se ocupa de la creación de artefactos que exhiben este comportamiento.

El cualquier caso, desde distintas perspectivas, en el seno de la IA como ciencia y tecnología se han ido acumulando conocimientos sobre como emular las diversas capacidades del ser humano para exhibir comportamientos inteligentes y se han desarrollado sistemas cada vez más perfeccionados que reproducen

parcialmente dichas capacidades. Así, el estudio de sensores y mecanismos de interacción de información y aplicaciones prácticas. También la robótica, como especialidad relacionada con máquinas móviles capaces de interactuar inteligentemente con su entorno y manipular objetos, que suele considerarse parcialmente en el ámbito de la IA. [17]

2.2.3 Sistema Experto

Se puede definir como aquel programa de ordenador que contiene la erudición de un especialista humano versado en un determinado campo de aplicación. En este sentido, los expertos escasean y su contratación supone una gran inversión económica, por lo que se intenta construir un sistema de forma que los conocimientos del experto se representen en una forma que el ordenador pueda procesar. Esto es, un modelo computarizado de las capacidades de razonamiento y habilidades en resolución de problemas del especialista humano.

Dado que un sistema experto pretende emular la actividad de los expertos humanos, debería ser capaz, en principio, de:

- Resolver el problema que se les plantea de la misma manera que el experto humano.
- Trabajar con datos incompletos o información insegura (como hace el experto humano en ocasiones).
- Explicar el resultado obtenido.
- Aprender conocimientos nuevos sobre la marcha.
- Reestructurar los conocimientos de que dispone en función de datos nuevos.
- Saltarse las normas, cuando se llega a la conclusión de que éstas no son aplicables a nuestro caso concreto.

Los sistemas expertos tienen notables ventajas sobre los expertos humanos. Por una parte, el conocimiento contenido en los sistemas expertos es más fácil de documentar y de transferir que el de los expertos humanos. Por otra, dicho conocimiento es remanente, es decir, permanece tras la desaparición de los expertos, por lo que constituye lo que se ha dado en llamar la memoria institucional del organismo o empresa que lo ha desarrollado. Asimismo, resulta

fácilmente transportable dando lugar a una especie de “experto ubicuo” que, por expertos resultan, a la larga, más económicos para cualquier organización que los expertos humanos. [17]

2.2.4 Visión Artificial

La Visión Artificial describe la deducción automática de la estructura y propiedades de un mundo tridimensional posiblemente dinámico, bien a partir de una o varias imágenes bidimensionales del mundo. Las imágenes pueden ser monocromáticas (de niveles de gris) o colores, pueden provenir de una o varias cámaras e incluso cada cámara puede estar estacionaria o móvil.

Las estructuras y propiedades del mundo tridimensional que queremos deducir en visión artificial incluyen no sólo sus propiedades geométricas, sino también sus propiedades materiales

2.2.4.1 Pre-procesamiento de imágenes digitales

En la etapa de pre procesamiento se pretende corregir defectos que puedan producirse durante la captura y digitalización de las imágenes las placas radiográficas. Si bien se ha tomado las precauciones para obtener una muestra lo más limpia posible.

2.2.4.2 Segmentación de imágenes

Segmentación es un término genérico que designa una serie de transformaciones que, normalmente a partir de la imagen procesada, la divide en partes, cada una de las cuales muestra una fuerte correlación con algún objeto o región.

El primer paso del análisis de imágenes consiste generalmente en segmentar una imagen. La segmentación subdivide una imagen en sus partes constituyentes u objetos. El nivel al que se lleva a cabo esta subdivisión depende del problema a resolver. La segmentación deberá degenerarse cuando los objetos de interés de una aplicación hayan sido aislados.

El resultado de la segmentación suele ser el punto de partida de otro proceso de interpretación de más alto nivel. Detección de Discontinuidad:

La detección de los tres tipos básicos de discontinuidades de una imagen digital: puntos, líneas y bordes.

Este procedimiento implica calcular la suma de los productos de los coeficientes por los niveles de gris contenidos en la región encerrada por la máscara. Esto es la respuesta de la máscara en un punto cualquiera de la imagen

$$R = w_1z_1 + w_2z_2 + \dots + w_xz_x$$

$$= \sum w_i z_i \quad i=1$$

Donde Z_i es el nivel de gris asociado con el coeficiente de la máscara de w_i . Cuando la máscara está centrada en un píxel del límite, la respuesta se calcula utilizando el entorno parcial apropiado. [18]

2.2.4.3 Detención de borde formula básica

Un borde es la frontera entre dos regiones con propiedades de niveles de gris relativamente distintas. La figura 1(a) muestra una imagen de una banda clara sobre un fondo oscuro, el perfil (1(c)) del nivel de gris a lo largo de una línea de exploración horizontal de la imagen, y la primera (1(e)) y segunda(1(g)) derivada del perfil.

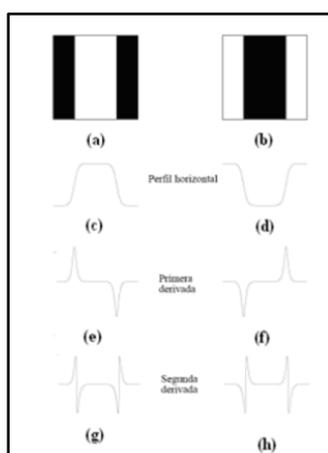


Figura 1: Imágenes con diferentes niveles de gris
Fuente: Gonzales, R. C., & Woods, R. E.

La figura 1 muestra que la primera derivada del perfil de nivel de gris es positiva en el borde de entrada de la transición y negativa en el borde de salida, como era de esperar, cero en las zonas de nivel de gris constante. La derivada segunda es positiva en la parte de la transición

asociada con el lado oscuro del borde, negativa en la parte de la transición asociada con el lado claro y cero en las zonas de nivel de gris constante. Por lo tanto, el módulo de la derivada primera se puede utilizar para detectar la presencia de un borde en una imagen, y el signo de la segunda derivada se puede utilizar para determinar si un píxel borde está situado en el lado oscuro o claro del mismo.

Hay que observar que la segunda derivada tiene un paso por cero en el punto medio de una transición del nivel de gris. [18]

2.2.4.3.1 Filtro

El filtrar una imagen (f) consiste en aplicar una transformación (T) para obtener una nueva imagen (g) de forma que ciertas características son acentuadas o disminuidas:

$g(x, y) = T[f(x, y)]$ Podemos considerar que la señal (imagen) pasa a través de una caja o sistema (filtro) cuya salida es la imagen filtrada.

(Ver fig. 2) De acuerdo a la teoría de sistemas, al pasar una señal por un sistema lineal, la salida es la convolución de la transformación del sistema (función de transferencia) con la señal de entrada: [20]

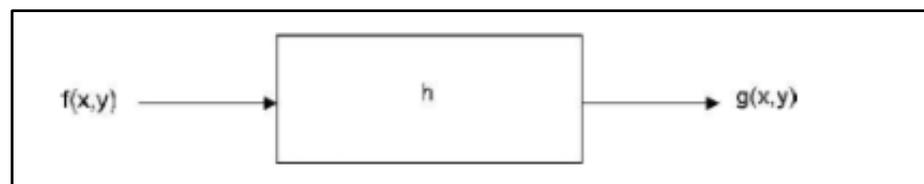


Figura 2: Proceso Filtrado

Fuente: Santillán Elías García

$$g(x; y) = h(x; y) * f(x; y)$$

Por el teorema de la convolución, esto corresponde a la multiplicación en el dominio de la frecuencia:

$$G(u; v) = H(u; v)F(u; v)$$

Por esto, podemos pensar en dos formas básicas de filtrar una imagen, realizarlo en el dominio espacial, que implica una convolución, o en el dominio de la frecuencia, que implica solo multiplicación, pero dos transformaciones de Fourier (de espacio a frecuencia y viceversa). Ambos tipos de filtros han sido ampliamente estudiados y a continuación veremos sólo una introducción general y su aplicación en imágenes.

Filtros de suavizamiento: El objetivo de los filtros de suavizamiento es eliminar ruido o detalles pequeños que no sean de interés. Esto corresponde a un filtro pasa bajo en el dominio de la frecuencia, es decir que se eliminan o reducen las altas frecuencias. En la figura 3 se muestra la respuesta de un filtro pasa-bajo en frecuencia (en una dimensión) y la correspondiente respuesta que debe tener en el dominio espacial.

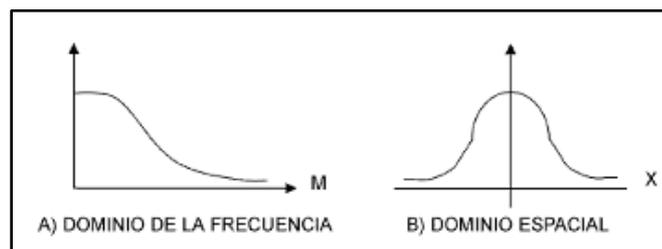


Figura 3: Filtro pasa-bajos: (a) en frecuencia, (b) en el dominio espacial
Fuente: Iván, García

Existen varios tipos de filtros para suavizamiento, los más comunes son: Promedio o media aritmética: Obtiene el promedio de los píxeles vecinos ($w = 1$); es decir, todos los valores de la máscara son 1.

Mediana: Substituye el valor del píxel central por el de la mediana de los valores contenidos en el vecindario. Gaussiano: Aproximación a una distribución gaussiana en dos dimensiones.

Considerando una media igual a cero, la función de transformación de un filtro tipo gaussiano es: [7]

$$T(x, y) = e^{-[(x^2+y^2)/2\pi\sigma^2]}$$

Donde es la desviación estándar. Para un máscara de 3x3 los valores de un filtro gaussiano “típico” se muestran en la figura 4. La cantidad de “suavizamiento” que realiza el filtro gaussiano se puede controlar variando la desviación estándar y el tamaño de la máscara.

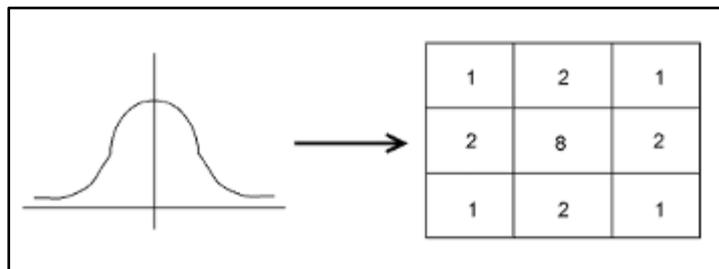


Figura 4: Máscara para filtro gaussiano de 3x3

Fuente: Sucar, Enrique

El filtro Gaussiano, en general, da mejores resultados que un simple promedio o media y se argumenta que la vista humana hace un filtrado de este tipo. El filtro Gaussiano “normal” o isotrópico tiene la desventaja de suavizar las orillas o discontinuidades, generando que se emborronen.

Lo anterior genera problemas en las posteriores etapas de visión. El algoritmo de mediana es particularmente efectivo en imágenes con poco ruido. Su efectividad decrece drásticamente en imágenes ruidosas.

Filtros de asentamiento: El objetivo de los filtros de asentamiento es intensificar los detalles y cambios bruscos de intensidad mientras atenúa las bajas frecuencias. El resultado es un asentamiento de las orillas (edge sharpening). Se conocen como filtros de pasa-alto porque dejan pasar las altas frecuencias y eliminan las bajas frecuencias, en forma inversa al filtro pasa-bajo. En la figura 5 se

muestra cómo se reducen las bajas frecuencias y se mantienen las altas.
[21]

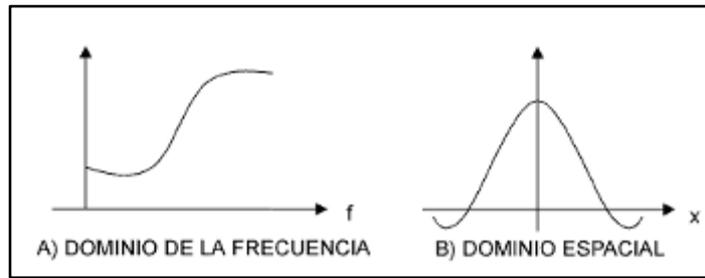


Figura 5: Filtro pasa-alto: (a) en frecuencia, (b) en el dominio espacial
Fuente: Sucar, Enrique

Existen también varias formas de implementar este tipo de filtrado. Una forma típica de un filtro pasa-alto es una máscara del tipo de la figura 6. Para este filtro la suma de los pesos es cero, de forma que una región de intensidad constante resultara en un valor 0. Nótese que a diferencia del filtro de suavizamiento los pesos de los vecinos son negativos, este efecto substractivo genera la acentuación de los cambios de intensidad.

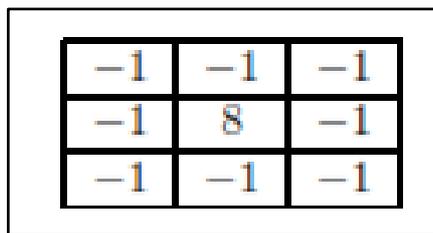


Figura 6: Máscara de 3x3 para un filtro pasa-alto simple
Fuente: Sucar, Enrique

Filtro para énfasis de altas frecuencias: El filtrado de asentamiento o pasa altos presenta sólo las discontinuidades, atenuando fuertemente las bajas frecuencias y haciendo que “desaparezcan” las regiones homogéneas. Un tipo de filtro que aún acentuando las altas frecuencias preserva las bajas es el filtro “énfasis de altas frecuencias” (high boost). Para obtener una imagen con énfasis de altas frecuencias (EA), se puede considerar que se multiplica la imagen original por una constante A, esta constante debe ser mayor que uno para que acentúe.
[21]

EA = (A) original PB

Eso es equivalente a la siguiente expresión:

EA = (A - 1) original + PA

En la práctica no es necesario hacer exactamente esta operación, sino se implementa haciendo la celda central del filtro pasa-alto:

w = 9A 1

Como se ilustra en la figura 7:

-1	-1	-1
-1	9A - 1	-1
-1	-1	-1

Figura 7: Máscara de 3x3 para un filtro con énfasis en las altas frecuencias.

Fuente: Sucar, Enrique

2.2.4.3.2 Gaussiano o Laplaciano

El Laplaciano de filtro gaussiano encuentra bordes mediante la búsqueda de cruces por cero después de la filtrar la imagen de entrada con un filtro Laplaciano de un filtro gaussiano. Considere la función gaussiana.

$$h(r) = -e^{\frac{-r^2}{2\sigma^2}}$$

Donde $r^2 = x^2 + y^2$ y es la desviación estándar. Esto es una función de suavizado que, de convolución con una imagen, se desenfoca. El grado de desenfoco se determina por el valor de σ . El Laplaciano de esta función (la segunda derivada con respecto a r) es σ :

$$\nabla^2 h(r) = -\left[\frac{r^2 - \sigma^2}{\sigma^4}\right]e^{\frac{-r^2}{2\sigma^2}}$$

Por obvias razones, esta función es llamada “Laplacian of a Gaussian” (LoG). Porque la segunda derivada es una operación lineal. Convolucionar una imagen con una $\nabla^2 h(r)$ es lo mismo que la convolución de la imagen con la función de alisado primero y luego calcular el Laplaciano del resultado. [19]

2.2.4.3.3 Blur

La función Blur, es el caso más simple. Cada píxel de la salida es la media simple de todos los píxeles en una ventana alrededor del píxel correspondiente en la entrada. Blur apoya 1-4 canales de imagen y funciona en imágenes de 8 bits o imágenes de punto flotante de 32 bits.

No todos los operadores de suavizado actuar sobre la misma clase de imágenes. Blur es esencialmente la misma que el Blur simple, excepto que no hay división lleva a cabo para crear un promedio. Por lo tanto, las imágenes de origen y de destino deben tener diferente precisión numérica de modo que la operación de desenfoque no dará lugar a un desbordamiento [22]

2.2.4.3.4 Gaussian Blur

Esta función tiene en cuenta el peso de los píxeles más cercanos que los alejados. La función realizada por este algoritmo de suavizado es la siguiente:

$$G(x, y) = \frac{1}{2 * \pi * \sigma^2} * e^{\frac{-x^2+y^2}{2*\sigma^2}}$$

Donde la “x” es la distancia desde el origen (punto del píxel a tratar) en el eje horizontal, e “y” es la distancia desde el origen en el eje vertical y “ σ ” es la desviación estándar de la distribución Gaussiana. Cuando se aplica en dos dimensiones, esta fórmula produce una superficie cuyos contornos son círculos concéntricos con la distribución Gaussiana desde el punto central. [22]

2.2.4.3.5 Filtro Bilateral

También se define como un promedio ponderado de pixeles adyacentes, de una manera muy similar a la convolución gaussiana. La diferencia es que el filtro bilateral tiene en cuenta la diferencia de valor con los vecinos para conservar bordes mientras el suavizado. La idea fundamental del filtro bilateral es que un píxel a píxel otra influencia, no sólo debería ocupar un lugar cercano, pero también tienen un valor similar. El filtro bilateral, denotada por BF [I], está definida por: [23]

$$BF[I]_{\mathbf{p}} = \frac{1}{W_{\mathbf{p}}} \sum_{\mathbf{q} \in \mathcal{S}} G_{\sigma_s}(\|\mathbf{p} - \mathbf{q}\|) G_{\sigma_r}(|I_{\mathbf{p}} - I_{\mathbf{q}}|) I_{\mathbf{q}},$$

Donde el factor de normalización garantiza que los pesos del píxel sumen 1.0:

$$W_{\mathbf{p}} = \sum_{\mathbf{q} \in \mathcal{S}} G_{\sigma_s}(\|\mathbf{p} - \mathbf{q}\|) G_{\sigma_r}(|I_{\mathbf{p}} - I_{\mathbf{q}}|).$$

2.2.4.3.6 Median Blur

En esta función se reemplaza cada píxel por la mediana o el "Promedio" de píxel (frente a la media del valor de píxel) en una plaza barrio alrededor del píxel central. Median Blur trabajará en un solo canal o tres canales o cuatro canales de imágenes de 8 bits, pero no se puede hacer en su lugar. Median blur puede ser sensible a las imágenes ruidosas, especialmente las imágenes con grandes puntos atípicos aislados (a veces llamado "ruido de disparo"). Las grandes diferencias en incluso un pequeño número de puntos pueden causar un movimiento notable en el valor medio. La mediana de filtrado es capaz de ignorar los valores extremos mediante la selección de los puntos medios. [24]

2.2.4.3.7 Binarización

Las imágenes binarias siempre se obtienen a partir de imágenes de niveles de gris. En la actualidad no existen cámaras comerciales que proporcionen imágenes binarias. El proceso de conversión de una imagen de nivel de gris a una imagen formada solo por dos valores o etiquetas (0 para el negro y 1 para el blanco) se conoce como binarización.

Para introducir el procedimiento de binarización consideremos la imagen que aparece que ha sido obtenida iluminando la pieza a contraluz. El histograma de esta imagen presenta dos poblaciones claramente diferenciadas: una de píxeles con muy altos niveles de gris pertenecientes al fondo retroiluminado y otra de píxeles con bajos valores correspondientes al objeto, ya que este no recibe luz en la cara que muestra a la cámara.

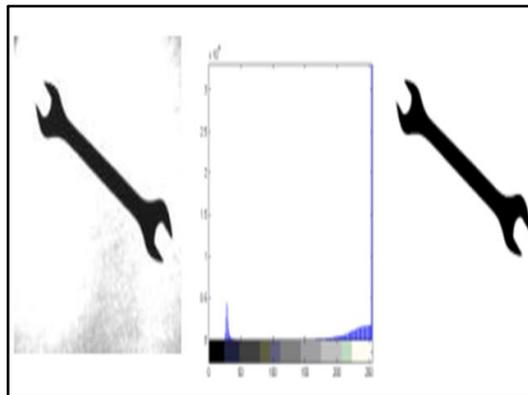


Figura 8: La binarización en escenas sencillas muy contrastadas como las obtenidas a contraluz.

Fuente: f.m.trespaderne, e. de la fuente

El algoritmo de binarización con umbral fijo, presupone que los objetos de interés aparecerán con niveles de gris más claros que el fondo. En caso contrario, una sencilla operación de inversión podría llevarse a cabo como paso siguiente a la binarización.

```

→ I:   Imagen original de niveles de gris de tamaño (filas,columnas)
  umbral: Umbral de binarización
← B:   Imagen binaria
BINARIZACIONFIJA(I,umbral)
  Inicializar B=0           %Se inicializan a 0 todos los valores de la matriz
  Mientras se recorre I(x,y) x=1...filas, y=1...columnas
    SI I(x,y) > umbral      %Presuponemos fondo oscuro
      B(x,y) = 1
    Fin
  Fin
  Devolver B

```

Algoritmo 1: Binarización con umbral fijo

Fuente: f.m.trespaderne, e. de la fuente

Binarización automática: Cuando se prevee que la iluminación va a estar expuesta a variaciones, no se puede trabajar con un umbral fijo. En este caso, lo apropiado es hallar para cada imagen, utilizando su histograma, el valor más idóneo como umbral.

El objetivo de esta técnica será calcular automáticamente, para cada imagen de niveles de gris que aparezca en el proceso, el umbral de binarización óptimo. Así, se pueden evitar los problemas de utilizar un umbral fijo, tales como el desplazamiento del histograma debido a cambios en la iluminación. El pseudocódigo para la binarización automática se muestra en el algoritmo 2. Este algoritmo puede optimizarse de forma sencilla, pero la implementación presentada muestra más claramente la esencia del algoritmo. [24]

```

→ I:   Imagen original de G niveles de gris de tamaño (filas,columnas)
← B:   Imagen binaria
BINARIZACIONAUTOMATICA(I)
  hist=HISTOGRAMA(I)
  Mientras se recorre hist(k), k=1... G
    q1=Σi=1k hist(i)           %número de píxeles en 1º grupo
    q2=Σi=k+1G hist(i)        %número de píxeles en 2º grupo
    var1=VARIANZA(hist(1),..., hist(k)) %varianza 1º grupo
    var2=VARIANZA(hist(k+1),..., hist(G)) %varianza 2º grupo
    suma=q1*var1+q2*var2      %suma ponderada varianzas
    SI k=1                    %inicializa variables en 1ª iteración
      menor_suma=suma
      umbral_optimo=1
    en caso contrario
      SI suma < menor_suma    %actualiza umbral si óptimo
        menor_suma=suma
        umbral_optimo=k
      Fin
    Fin
  Fin
  B=BINARIZACIONFIJA(I,umbral_optimo)
  Devolver B

```

Algoritmo 2: Binarización con selección automática del umbral

Fuente: f.m.trespaderne, e. de la fuente

2.2.5 Histograma

Es un hecho que el histograma de una imagen de intensidad se encuentra dentro de un rango de datos limitada. Esas imágenes tienen generalmente plano negro o blanco y el fondo. Figura 9 muestra un ejemplo de imagen cuya distribución de intensidad es blanco o negro. Desde la figura 9 (a) podemos ver que una porción muy grande de píxeles cuya intensidad se basa en el rango [0,50] o [180,255]. Una porción muy pequeña de píxeles reside en el rango de [50,180]. Esto hace que algunos detalles de la imagen apenas visibles, por ejemplo, los árboles en las montañas en la imagen mostrada en la figura 9. Este problema puede resolverse mediante una técnica de histograma de estiramiento llamado ecualización de histograma. [25]

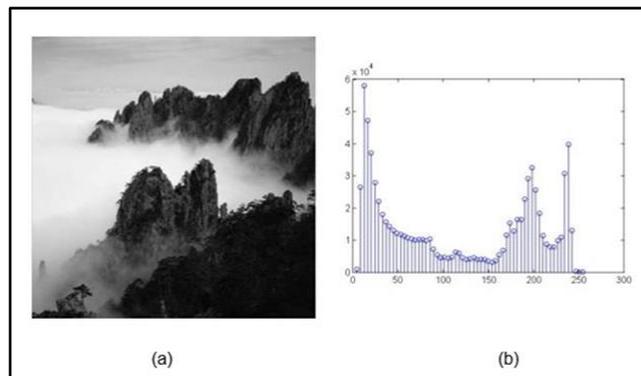


Figura 9: Una montaña (a) y sus 64 contenedores del histograma nivel gris (b)
Fuente: González, R.C., Wintz, P

2.2.5.1 Ecualización del histograma

Ecualización del histograma se realiza para mejora de la imagen, es una técnica utilizada para la obtención de histograma uniforme. Para cualquier r en el intervalo $[0, 1]$, la transformación es:

Cuando, s es el nivel producido por cada valor de píxel r en la imagen original. La función de densidad de probabilidad de nivel de gris transformada es:

$$(s) = \left[\Pr(r) \frac{dr}{r-T^{-1}(s)} \right]$$

Considerando la función de transformación:

$$s = T(r) \int_0^r \text{Pr}(w) dw$$

Los niveles de salida son uniformes y se utiliza para la mejora de la imagen. El resultado de la ecualización del histograma se incrementa en el rango dinámico, lo que dará un mayor contraste. [25]

2.2.6 Reconocimiento de patrones

2.2.6.1 Detención de objetos

Si se nos da una imagen de analizar, como el retrato de grupo en la figura, podríamos tratar de aplicar un algoritmo de reconocimiento de toda la posible subventana en esta imagen.

Tales algoritmos son propensos a ser lentos y propenso a errores. En cambio, es más eficaz para construir detectores de propósito especial, cuyo trabajo consiste en encontrar rápidamente probables regiones donde podrían ocurrir objetos particulares. [26]

2.2.6.2 Características, vectores de características y clasificadores

Los descriptores visuales describen las características de los contenidos dispuestos en imágenes o en vídeos. Describen características elementales tales como la forma, el color, la textura o el movimiento, entre otros. El contenido de una imagen está codificado digitalmente en el valor de cada una de las unidades mínimas de información que la componen llamadas pixels. De esta manera los pixels representan el nexo de unión entre el contenido abstracto de sus valores y las características propias de una imagen que entendemos como relevantes para el humano. Es por ello que cualquier método de gestión de imágenes basado en su contenido deberá guardar algún tipo de relación o actuar sobre el valor de los mismos. Idealmente, un descriptor visual debería poseer las siguientes propiedades: [27]

Tabla 1: Tabla de descriptor

Propiedad	Descripción
Simplicidad	El descriptor debería representar las características extraídas de la imagen de manera clara y sencilla para permitir una fácil interpretación de su contenido.
Repetibilidad	El descriptor generado a partir de una imagen debe ser independiente del momento en el que se genere.
Diferenciabilidad	Dada una imagen, el descriptor generado debe poseer alto grado de discriminación respecto de otras imágenes y al mismo tiempo contener información que permita establecer una relación entre imágenes similares.
Invarianza	Cuando existen deformaciones en la representación de dos imágenes, es deseable que los descriptores que las representan aporten la robustez necesaria para poder relacionarlas aún bajo diferentes transformaciones.
Eficacia	Es deseable que los recursos consumidos para generar el descriptor sean aceptables para poder ser utilizados en aplicaciones con restricciones críticas de espacio y/o tiempo.

Fuente: (Boullosa Gracia 2008)

2.2.7 Redes Neuronales

Las redes neuronales artificiales son un método de resolver problemas, de forma individual o combinadas con otros métodos, para aquellas tareas de clasificación, identificación, diagnóstico, optimización o predicción en las que

el balance datos o conocimiento se inclina hacia los datos y donde, adicionalmente, puede haber la necesidad de aprendizaje en tiempo de ejecución y de cierta tolerancia a fallos. Está inspirado en el comportamiento biológico de las neuronas y en la estructura del cerebro.

Una neurona biológica es una célula especializada en procesar información. Está compuesta por el cuerpo de la célula (soma) y dos tipos de ramificaciones: el axón y las dendritas. La neurona recibe las señales (impulsos) de otras neuronas a través de sus dendritas y transmite señales generadas por el cuerpo de la célula a través del axon. [28]

La arquitectura de una red neuronal está conformada por múltiples unidades procesadoras elementales interconectadas, formando estas un sistema adaptivo que posee un algoritmo para ajustar sus pesos y alcanzar los requerimientos de desempeño del problema basado en muestras representativas.

Existen tres tipos de neuronas artificiales: unidades de entrada, de capa oculta y unidades de salida.

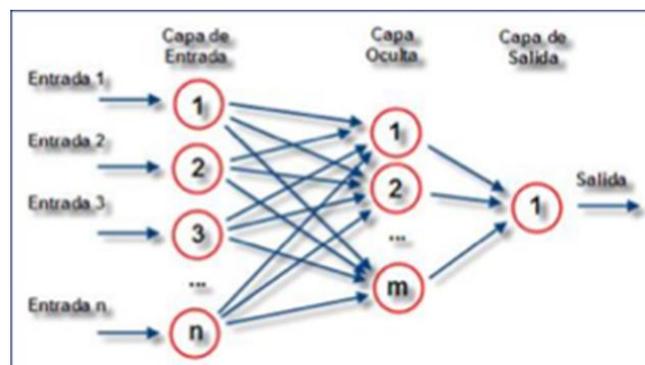


Figura 10: Estructura de una red neuronal multicapa, con una capa de entrada, una capa de oculta y una capa de salida.

Fuente: Florez, R., & Fernandez, J. M.

Las neuronas de la capa de entrada reciben señales del entorno, del exterior, provenientes de sensores u otros medios de información. Las neuronas de la capa oculta reciben las señales recibidas dentro del sistema sin tener contacto con el exterior de la red. Es en esta capa donde ocurre el

procesamiento de la información. Las neuronas envían la información procesada hacia el exterior. [29]

2.2.8 Base de Datos

Los SDBD (sistemas manejadores de base de datos) se han incrementado en los últimos años de forma drástica, pues claro está que cada vez más empresas requieren de software para registrar sus datos.

El objetivo principal del sistema gestor de base de datos es proporcionar una forma de almacenar y recuperar la información de una base de datos de manera que sea tanto práctica como eficiente.

Se compone de un lenguaje de definición de datos (DDL: Data Definition Language), de un lenguaje de manipulación de datos (DML: Data Manipulation Language) y de un lenguaje de consulta (SQL: Structured Query Language) [30]

1.5.1 MYSQL

MySQL es un sistema de gestión de base de datos relacional, multihilo y multiusuario con más de seis millones de instalaciones. MySQL AB —desde enero de 2008 una subsidiaria de Sun Microsystems y ésta a su vez de Oracle Corporation desde abril de 2009 desarrolla MySQL como software libre en un esquema de licenciamiento dual.

Por un lado se ofrece bajo la GNU GPL para cualquier uso compatible con esta licencia, pero para aquellas empresas que quieran incorporarlo en productos privativos deben comprar a la empresa una licencia específica que les permita este uso. Está desarrollado en su mayor parte en ANSI C. [31]

Esto es lo que posibilita el esquema de licenciamiento anteriormente mencionado. Además de la venta de licencias privativas, la compañía ofrece soporte y servicios. Para sus operaciones contratan trabajadores alrededor del mundo que colaboran vía Internet. MySQL AB fue fundado por David Axmark, Allan Larsson y Michael Widenius. [32]

Tabla 2: Características de MySQL

Características	MySQL
Costo	Libre y de pago
Open Source	Si
Plataformas	Linux, Windows y muchas otras
Límite de tamaño de la base de datos	Limitado por el sistema operativo
Compatibilidad ACID	Si
Transacciones	Si
Servicio de reportes	No
Claves Foráneas	Depende del motor
Subconsultas	Si
Replicación	Si
Funciones definidas por el usuario	Si

Fuente: Salamon, Peter (2009)

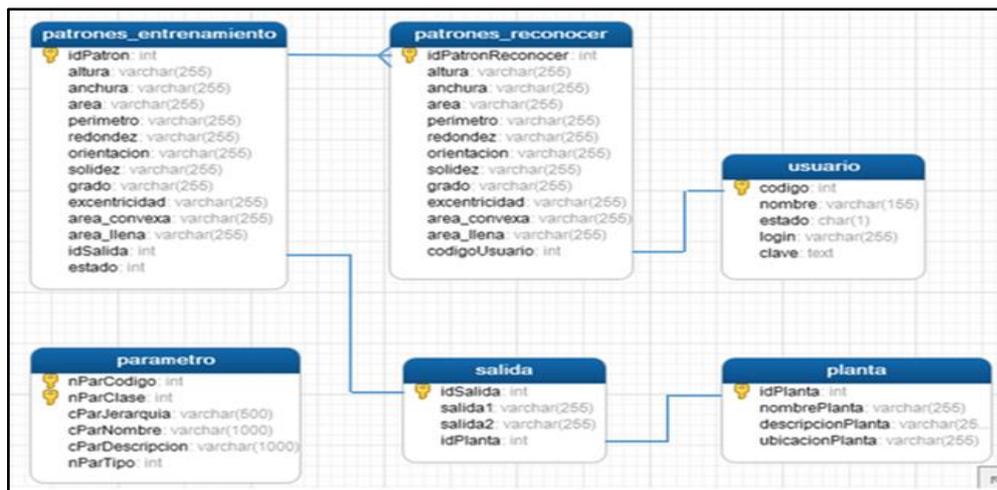


Figura 11: Base de Datos
Fuente: Elaboración Propia

2.2.9 Lenguaje de Programación

Un lenguaje de programación es un idioma artificial diseñado para expresar procesos que pueden ser llevados a cabo por máquinas como las computadoras.

Pueden usarse para crear programas que controlen el comportamiento físico y lógico de una máquina, para expresar algoritmos con precisión, o como modo de comunicación humana.

Está formado por un conjunto de símbolos y reglas sintácticas y semánticas que definen su estructura y el significado de sus elementos y expresiones. Al proceso por el cual se escribe, se prueba, se depura, se compila (de ser necesario) y se mantiene el código fuente de un programa informático se le llama programación. [16]

1.5.2 JAVA

Un lenguaje de programación orientado a objetos desarrollado por Sun Microsystems [33]

- Encapsulación, herencia y polimorfismo
- Fuertemente tipado
- Gestión automática de la memoria (recogida de basura)
- Soporte para concurrencia (multihilo)
- Gestión de excepciones
- Constructores independientes de la arquitectura del procesador

2.2.10 Cultivo de la Quinoa

La quinoa (*Chenopodium quinoa* Willd.), por sus elevadas cualidades nutricionales constituyó a lo largo del tiempo, junto a otros cultivos nativos, uno de los alimentos básicos del hombre andino. Con la llegada hispana a América se introdujo otros cultivos, como los cereales, muchos de los cuales desplazaron a los tradicionales, razón por la cual la quinoa pasó a constituirse en un cultivo marginal practicado por algunas comunidades campesinas de la zona andina.

En la actualidad este cultivo se erige como de gran importancia entre los agricultores del país, como consecuencia de la gran demanda nacional y extranjera que se tiene, ya que este grano nativo destaca por sus bondades nutricionales. Así, la demanda internacional ha experimentado progresos exponenciales lo cual se ve reflejado en la cantidad de toneladas exportadas, que del año 2007 al 2012, pasaron de 1500 a 10400, respectivamente, como señala el Ministerio de Agricultura. [34]

2.2.10.1 Fisiología y fenología del cultivo de la quinua

La quinua es una planta C3 es decir anatómicamente se halla dentro del grupo de plantas que solamente pueden aprovechar el 30% del CO₂ del aire, comparado con plantas C4, que fijan 80-100% del CO₂ del aire y necesitan 2 a 3 veces más agua que las plantas tipo C4 para producir la misma cantidad de masa vegetal. Sin embargo añade que la quinua demuestra gran resistencia a condiciones de sequía extrema gracias a dos comportamientos fisiológicos. El primero se basa en el potencial foliar mínimo (P_{f min}), que viene a ser la fuerza con que la planta extrae el agua del suelo y también la fuerza con que retiene esta agua en sus tejidos. El segundo es la resistencia estomática (RE), que es el mecanismo por el cual las plantas cierran sus estomas en condiciones de sequía, esto, para evitar una transpiración excesiva. [35]

2.2.10.2 Agronomía del cultivo de la quinua

El valor nutritivo de la quinua es excepcional debido a que tiene un excelente balance de carbohidratos, grasas y proteínas para la alimentación humana. La cantidad de proteínas y la calidad son especialmente importantes. El contenido de proteína de los granos de quinua varía entre 14-22%, siendo significativamente mayor que la de cereales. Sin embargo, la ventaja nutricional más importante de la quinua es la composición de aminoácidos esenciales, los cuales están muy cerca de los recomendados por la FAO/OMS/UNU para todos los grupos de edad. [35]

2.2.11 Marco Conceptual

- **GPS:** Sistema americano de navegación y localización mediante satélites.
- **Red Neuronal:** Es un procesador distribuido en paralelo de forma masiva con una propensión natural a almacenar conocimiento experimental y convertirlo en disponible para su uso
- **Inteligencia Artificial:** Es la combinación de algoritmos planteados con el propósito de crear máquinas que presenten las mismas capacidades que el ser humano.

CAPÍTULO II METODOLOGÍA

3.1 Tipo y nivel de investigación

3.1.1 Tipo de investigación

El presente trabajo corresponde a una investigación de tipo experimental debido a que se ha implementado tecnologías aplicadas en un determinado tiempo para determinar sus efectos.

3.1.2 Nivel de investigación

Según el tipo de contrastación de hipótesis la investigación es cuasi-experimental, ya que se observará a los expertos al realizar sus procesos de identificación y localización de manera tradicional y la mejoría que han logrado desde que las iniciaron, del mismo modo se evaluará al realizar sus procesos haciendo el uso del sistema propuesto.

3.2 Diseño de investigación

De acuerdo al tipo de investigación el diseño utilizado es cuasi-experimental utilizando un grupo experimental y un grupo control pretest - posttest cuyo esquema es el siguiente:

$$GE = O_1 X O_2$$

$$GC = O_3 O_4$$

Dónde:

GE: grupo experimental, software propuesto

GC: grupo control, proceso manual

X: estímulo o variable independiente

O1 O2: mediciones aplicadas antes y después en el grupo experimental

Se generará una situación para tratar de explicar cómo afecta y quienes participan en ella en comparación con quienes no lo hacen. En este caso las plantas de quinua serán analizadas en el proceso de identificación y localización.

El diseño de investigación es de dos grupos que son pre y post test, ya que lo que se pretende es conocer la mejoría en la ejecución de los procesos de identificación y localización de las plantas de quinua y con el apoyo tecnológico, que es el sistema que se propone. Es así como se evaluará a dos grupos de plantas de quinua, uno de ellos será el grupo control, por lo que no recibirán el tratamiento experimental, y el otro sí lo hará.

3.3 Población, muestra y muestreo

3.3.1 Población

Nuestra población viene definida en base a 16 especies de plantas de quinua, la cual tiene sub variedades que llegan a un total de 3 mil variedades de plantas de quinua, en la que la empresa se enfoca en 3 variedades de plantas de quinua.

Además, para poder obtener información acerca del proceso de identificación y localización de las diversas variedades de planta de quinua, se consultó al biólogo Carlos Sialer.

Tabla 3: Tipificación de la población

POBLACIÓN	DESCRIPCIÓN	TIPO	CANTIDAD
Plantas de quinua	Las plantas de quinua de interés de la empresa.	Finito	3

Fuente: Elaboración propia

3.3.2 Muestra y Muestreo

El tipo de muestreo aplicado es el no probabilístico. Obtenemos una muestra intencional, basándose en criterios seleccionados según el fin que perseguimos. Emplearemos el juicio de expertos, ya que serán ellos quienes

nos indiquen que plantas de quinua tienen las características necesarias para participar de la muestra.

3.4 Criterios de selección

3.4.1 Criterios de inclusión

- Hojas de la planta sin enfermedades
- Hojas de la planta correctamente fotografiadas

3.4.2 Criterios de exclusión

- Hojas de la planta con enfermedades
- Hojas de la planta mal fotografiadas

3.5 Operacionalización de variables

3.5.1 Variables

3.5.1.1 Variable independiente

-Sistema experto de reconocimientos de patrones en imágenes fotográficas.

3.5.1.2 Variable dependiente

-Detección automática de identificación y localización de las diversas variedades de plantas de quinua

3.5.2 Indicadores

Tabla 4: Indicadores

Objetivo específico	Indicador(es)	Definición conceptual	Unidad de medida	Instrumento	Definición operacional
Mejorar el nivel de certeza de los técnicos de campo para localizar diversas variedades de planta de quinua	N° de plantas localizadas	Información del progreso de los técnicos de campo	Cantidad de plantas localizadas	Registro del sistema de las plantas localizadas	- Σ Plantas reconocidas actualmente - Σ Plantas reconocidas previamente
Incrementar el nivel de eficiencia de los expertos para identificar las diversas variedades de plantas de quinua	N° de plantas identificadas	Información del progreso de los expertos	Cantidad de plantas identificadas	Registro del sistema de las plantas identificadas	- Σ Procesos registrados actualmente - Σ Procesos registradas previamente

Fuente: Elaboración Propia

3.6 Técnicas e instrumentos de recolección de datos

Tabla 5: Métodos y técnicas de recolección de datos.

MÉTODO	INSTRUMENTO	ELEMENTOS DE LA POBLACIÓN
Observación	Ficha de observación (ver anexo N°2, N°3)	Información de la planta de quinua
Entrevista	Guía de Entrevista o Formato de Entrevista (ver anexo N°1)	Técnicos de campo y Expertos

Fuente: Elaboración Propia

3.7 Procedimientos

El desarrollo del modelo de gestión del conocimiento se realizó bajo dos enfoques. La metodología propuesta por John Durkin para el desarrollo de sistemas inteligentes consta de seis fases y la estructura de un sistema de inferencia difuso. La metodología

propuesta por John Durkin para el desarrollo de sistemas inteligentes consta de las siguientes fases [18]

3.8 Plan de procesamiento y análisis de datos

1. Primera Etapa:

En esta etapa se procede a la revisión de la base teórica necesaria para el desarrollo del proyecto de tesis. Sobre todo, con lo referente a términos sobre las redes neurales acerca del reconocimiento de imágenes.

2. Segunda Etapa:

Consiste en el diseño y elección de las herramientas, librerías y de más componentes necesarios para el desarrollo del producto acreditable. Además de la elección de la metodología de desarrollo del software.

3. Tercera Etapa:

Luego de reunirse varias veces con los expertos en este caso un biólogo que brindan el servicio del proceso de una planta, se definirán los procesos del software que se llevara a cabo para la identificación y localización de las plantas de quinua

4. Cuarta Etapa:

Se procederá a la construcción del producto acreditable, en todo momento con la supervisión de los expertos.

5. Quinta Etapa:

Consiste en evaluar la propuesta de solución a través de una muestra piloto en la que un determinado biólogo donde nos dará a saber los pasos y procesos que tiene la planta de quinua. Luego se comparará las plantas reconocidas con otras plantas nuevas. Finalmente se llevará a cabo el informe final de tesis, con las conclusiones, limitaciones encontradas y futuras líneas de investigación.

A continuación se detalla las fases propuestas por John Durkin:

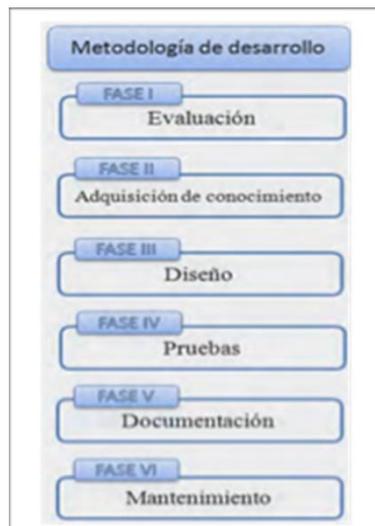


Figura 12: Fases de la metodología de desarrollo de sistemas inteligentes
Fuente: Morales, Mery

- **Fase 1: Evaluación**
 - Motivación para el Esfuerzo.
 - Identificar problemas candidatos.
 - Estudio de viabilidad.
 - Análisis de Costo/Beneficio.
 - Seleccionar el mejor proyecto.
 - Escribir el proyecto propuesto.

- **Fase 2: Adquisición Del Conocimiento**
 - Recolección del conocimiento.
 - Interpretación.
 - Análisis.
 - Diseño de métodos para recolectar conocimiento adicional.

- **Fase 3: Diseño**
 - Seleccionar Técnica de Representación del Conocimiento.
 - Seleccionar Técnica de Control.
 - Seleccionar Software de Desarrollo de Sistema Inteligente.
 - Desarrollo de Prototipo.
 - Desarrollo de Interface.
 - Desarrollo del Producto.

- **Fase 4: Pruebas**
 - Validación del Sistema.
 - Evaluación de la Prueba/Evaluación.

- **Fase 5: Documentación**
 - Relación de temas que deben ser documentados.
 - Organización de la documentación.
 - Documentación Impresa.
 - Documentación en hipertexto.
 - Reporte Final

- **Fase 6: Mantenimiento**
 - Modificaciones probables del sistema.
 - Responsables de mantenimiento.
 - Interfaces de documentación del mantenimiento

3.9 Matriz de consistencia

Tabla 6: Matriz de Consistencia

PROBLEMA	OBJETIVOS	HIPÓTESIS	VARIABLES
<p><u>PROBLEMA GENERAL</u></p> <p>-Proceso de identificación de la planta, ya que se hace manualmente, y eso toma tiempo, por lo que son grande volúmenes de información.</p> <p>-Proceso de localizar de una planta, se hace un registro manual, en la que dificulta volver a encontrar determinada planta</p>	<p><u>OBJETIVO GENERAL</u></p> <p>Implementar un software para identificar y localizar rápidamente las variedades de plantas de quinua</p>	<p><u>HIPÓTESIS GENERAL</u></p> <p>La implementación de un software que permitirá identificar y localizar de manera rápida las variedades de plantas de quinua a través del proceso de reconocimiento de patrones.</p>	<p><u>VARIABLES DE ESTUDIO</u></p> <p>VARIABLE INDEPENDIENTE</p> <p>Sistema experto de reconocimientos de patrones en imágenes fotográficas.</p> <p>VARIABLE DEPENDIENTE</p> <p>Detección automática de identificación y localización de las diversas variedades de plantas de quinua</p>
<p><u>PROBLEMAS ESPECÍFICOS</u></p> <p>Errores en la identificación de las variedades de quinua.</p> <p>Poca certeza en la ubicación de las variedades de las plantas de quinua después de haber sido localizadas</p>	<p><u>OBJETIVOS ESPECÍFICOS</u></p> <p>Mejorar el nivel de certeza de los técnicos de campo para localizar diversas variedades de planta de quinua</p> <p>Incrementar el nivel de eficiencia de los expertos para identificar las diversas variedades de plantas de quinua</p>	<p><u>HIPÓTESIS ESPECÍFICA</u></p> <p>Variación en el proceso de identificación y localización de la planta de quinua después de su proceso</p> <p>Información de variedades de plantas estudiadas por los investigadores no están haciendo aprovechadas de manera optima</p>	<p><u>INDICADORES</u></p> <p>Número de plantas localizadas</p> <p>Resultados por el sistema</p> <p>Evaluaciones de eficiencia a los expertos</p>

3.10 Consideraciones Éticas

Los datos serán utilizados para fines académicos, y la identidad de las personas que proporcionan la información se mantendrá en reserva, salvo los datos autorizados por la empresa.

CAPÍTULO III RESULTADOS

4.1 Selección De Metodología

Para el desarrollo de sistemas inteligentes se ha determinado que la que se usará en este caso será la metodología de Jhon Durkin ya que no solo se caracteriza por la constante interacción con el especialista para poder recolectar, analizar y procesar la información requerida para armar la base de conocimiento, sino que además una de sus fases comprende la documentación, la cual es de suma importancia.

4.2 Fase 1: Evaluación

4.2.1 Motivación para el Esfuerzo.

La identificación de las variedades de plantas de quinua es fundamental para la empresa ya que exporta por diferentes países, además de su valor nutritivo que contiene, es por ello que resulta de gran importancia que esta se lleva a cabo de la manera correcta.

En la actualidad no existe algún software en la que se pueda identificar las diversas variedades de esta planta.

Dada esta situación, se hace necesario implementar un sistema inteligente que elabore el proceso de identificación, según las características que presenta dicha planta de quinua, además que sea capaz de almacenar en una base de datos todas las plantas entrenadas y reconocidas, para una próxima comparación, teniendo también un valor agregado de la localización de la planta, para que la búsqueda sea más rápida.

En conclusión, se determina que la identificación y localización es conducida por el problema ya que se trata de resolver un problema que ya se ha identificado.

4.2.2 Identificar problemas candidatos.

No hay problemas candidatos, porque el problema fue directamente indicado por los interesados.

4.2.3 Estudio de viabilidad.

Tabla 7: Asuntos de Viabilidad del Problema

Puntaje	Peso	Valor Asunto	ASUNTOS DE VIABILIDAD DEL PROBLEMA
72	8	9	Conocimiento experto necesitado
81	9	9	Los pasos de solución de problema son definibles
48	8	6	Conocimiento simbólico usado
100	10	10	El problema es solucionable
72	9	8	Existen sistemas exitosos
56	7	8	El problema es bien enfocado
72	8	9	El problema es razonablemente complejo
72	9	8	El problema es estable
48	8	6	Conocimiento incompleto o incierto utilizado
49	7	7	Solución más de una recomendación
670	83	80	TOTAL

Donde:

$$\text{Puntaje} = \text{Peso} * \text{Valor Asunto}$$

- El peso y valor asunto, se dan con una evaluación del 0 al 10.

Tabla 8: Asuntos de Viabilidad del Experto

Puntaje	Peso	Valor Asunto	ASUNTOS DE VIABILIDAD DE EXPERTO
<i>EXPERTO DE DOMINIO</i>			
72	8	9	El experto puede comunicar el conocimiento
36	6	6	El experto puede dedicar tiempo
49	7	7	El experto es cooperativo
157	21	22	TOTAL
<i>INGENIERO DEL CONOCIMIENTO</i>			
48	8	6	Buenos procesos de identificación y localización
49	7	7	Puede relacionar el problema al software
36	6	6	Tiene destrezas de programación de sistema inteligente
81	9	9	Puede dedicar tiempo
214	30	28	TOTAL
<i>USUARIO FINAL</i>			
42	6	7	El usuario final puede dedicar tiempo
42	7	6	El usuario final es receptivo al cambio
35	5	7	El usuario final es cooperativo
119	18	20	TOTAL
490	69	70	TOTAL FINAL

Donde:

$$Puntaje = Peso * Valor Asunto$$

- El peso y valor asunto, se dan con una evaluación del 0 al 10.

Tabla 9: Asunto de viabilidad del despliegue

Puntaje	Peso	Valor Asunto	ASUNTOS DE VIABILIDAD DEL DESPLIEGUE
64	8	8	El sistema puede ser introducido fácilmente
49	7	7	El sistema puede ser mantenido
56	7	8	El sistema no tiene una ruta crítica
64	8	8	El sistema puede ser integrado con recursos existentes
42	6	7	Entrenamiento disponible
275	36	38	TOTAL

Donde:

$$Puntaje = Peso * Valor Asunto$$

- El peso y valor asunto, se dan con una evaluación del 0 al 10.

Tabla 10: Viabilidad del proyecto

CATEGORIA	PUNTAJE TOTAL	PESO TOTAL
Problema	670	83
Expertos	490	69
Despliegue	275	36
TOTAL	1435	188

Donde:

$$\text{Viabilidad del Proyecto} = 1435/188 = 7.632978723$$

4.2.4 Análisis de Costo/Beneficio

El desarrollo del proyecto que se llevará a cabo, requiere de gastos en materiales y recursos necesarios para la elaboración, ejecución y finalización del proyecto, materiales y recursos como computadora, materiales de oficina, transporte, copias fotostáticas, etc., que a continuación se detalla:

Tabla 11: Tabla de Costo

ÍTEM	DESCRIPCIÓN	UNIDAD	CANTIDAD	COSTO UNITARIO	TOTAL
SERVICIOS	Cámara Fotográfica Digital	UNIDAD	1	450.00	450.00
	Llamadas telefónicas	MINUTOS	100	0.50	50.00
	Luz eléctrica	MES	12	40.00	480.00
	Depreciación Laptop	ANUAL	1	1300.00	1300.00
	Impresora	UNIDAD	1	250.00	250.00
	Internet	MES	12	60.00	720.00
	Movilidad	PASAJES	60	50.00	3000.00
MATERIALES DE ESCRITORIO	Papel bond A4	PAQUETE	1	25.00	25.00
	Lápiz	UNIDAD	3	1.00	3.00
	Lapicero	UNIDAD	3	1.50	4.50
	USB 16GB	UNIDAD	1	45.00	45.00
	Tinta líquida para impresora	UNIDAD	3	10.00	30.00
	Fotocopias de artículos y otros	UNIDAD	150	0.10	15.00
TOTAL					6372.50

Tabla 12: Tabla de Beneficio

Inversión	3000.00
Valor presente de beneficio	15000.00
Valor presente de costo	6372.50
Total porcentaje beneficio	1.6

Donde:

$$\text{Total Porcentaje Beneficio} = \text{beneficio} / (\text{costo} + \text{inversión})$$

4.2.5 Seleccionar el mejor proyecto

Considerando el resultado del análisis de viabilidad del proyecto, el cual ha obtenido un puntaje mayor al 70% obtenidos en la viabilidad del proyecto (Tabla 11), y el de costo/beneficios se determina que el proyecto propuesto es el mejor ya que la inversión es mínima respecto a los beneficios que se le traerán a la organización, la cual mejorará el servicio ofrecido por consiguiente su imagen y prestigio.

4.2.6 Escribir el proyecto propuesto.

PROYECTO PERSONAL	
OBJETIVO	<p>Objetivo General: Implementar un software para identificar y localizar rápidamente las variedades de plantas de quinua</p> <p>Objetivos Específicos:</p> <ul style="list-style-type: none">✓ Mejorar el nivel de certeza de los técnicos de campo para localizar diversas variedades de planta de quinua✓ Incrementar el nivel de eficiencia de los expertos para identificar las diversas variedades de plantas de quinua
PROBLEMA	<ul style="list-style-type: none">▪ Poca certeza en ubicación y localización de las plantas▪ Exceso de tiempo en la identificación de la planta▪ Errores de comparación de las características de una planta al momento de identificar.▪ Pocas variedades de plantas de quinua identificadas por expertos sin experiencia▪ Excesivo volumen de información a procesar para identificar una planta▪ Las tomas fotográficas son deficientes dados a que los expertos tienen falta de conocimientos técnicos sobre fotografía.▪ Escasa capacitaciones para la identificación de planta
SOLUCIÓN	<ul style="list-style-type: none">▪ Disminuir el nivel de certeza de búsqueda para encontrar dicha planta▪ Reducir los errores de comparación para la identificación▪ Reducir el nivel de eficiencia de identificar la planta de Quinua
PLAN	<p>Primera Etapa:</p> <p>En esta etapa se procede a la revisión de la base teórica necesaria para el desarrollo del proyecto de tesis. Sobre todo, con lo referente a términos sobre las redes neurales acerca del reconocimiento de imágenes</p> <p>Segunda Etapa:</p> <p>Consiste en el diseño y elección de las herramientas, librerías y de más componentes necesarios para el desarrollo del producto acreditable. Además de la elección de la metodología de desarrollo del software.</p> <p>Tercera Etapa:</p> <p>Luego de reunirse varias veces con los expertos en este caso un biólogo que brindan el servicio del proceso de una planta, se definirán los procesos del software que se llevara a cabo para la identificación y localización de las plantas de quinua</p>

Cuarta Etapa:

Se procederá a la construcción del producto acreditable, en todo momento con la supervisión de los expertos.

Quinta Etapa:

Consiste en evaluar la propuesta de solución a través de una muestra piloto en la que un determinado biólogo donde nos dará a saber los pasos y procesos que tiene la planta de quinua. Luego se comparará las plantas reconocidas con otras plantas nuevas. Finalmente se llevará a cabo el informe final de tesis, con las conclusiones, limitaciones encontradas y futuras líneas de investigación.

TIEMPOS ESTABLECIDOS**FASE 1: EVALUACIÓN**

- Motivación para el Esfuerzo.
- Identificar problemas candidatos.
- Estudio de viabilidad. 20%
- Análisis de Costo/Beneficio.
- Seleccionar el mejor proyecto.
- Escribir el proyecto propuesto.

FASE 2: ADQUISICIÓN DEL CONOCIMIENTO

- Recolección del conocimiento.
- Interpretación.
- Análisis. 50%
- Diseño de métodos para recolectar conocimiento adicional.
- Historias de usuarios

FASE 3: DISEÑO

- Seleccionar Técnica de Representación del Conocimiento.
- Seleccionar Técnica de Control.
- Seleccionar Software de Desarrollo de Sistema Inteligente.
- Desarrollo de Prototipo.
- Desarrollo de Interface. 80%
- Glosario de términos
- Tarjetas C.R.C
- Desarrollo del Producto.

FASE 4: PRUEBAS

- Validación del Sistema
- Evaluación de la Prueba/Evaluación

FASE 5: DOCUMENTACIÓN

- Relación de temas que deben ser documentados.
- Organización de la documentación.
- Documentación Impresa. 100%
- Documentación en hipertexto.
- Reporte Final

<p>FASE 6: MANTENIMIENTO</p> <ul style="list-style-type: none"> ▪ Modificaciones probables del sistema. ▪ Responsables de mantenimiento. ▪ Interfaces de documentación del mantenimiento
<p>RESULTADO</p> <ul style="list-style-type: none"> -Sistema inteligente que servirá de apoyo para la identificación y localización de las plantas de quinua -Documentación impresa
<p>PERSONAL</p> <ul style="list-style-type: none"> • Alumno: Bach. Cristian Vallejos Fonseca • Asesor: Ing. Luis Augusto Zuñe Bispo

4.3 Fase 2: Adquisición Del Conocimiento

4.3.1 Recolección del conocimiento.

Para cumplir con esta fase se hizo uso de entrevistas a distintas especialistas en el tema. (ANEXO 1)

4.3.2 Interpretación.

Basándonos en las respuestas dadas por las distintas especialistas, se ha logrado interpretar lo siguiente:

- La localización de las plantas se le dificulta al técnico de campo en memorizar, y luego poder ubicar próximamente una determinada planta.
- Dificulta procesar grandes volúmenes de datos para un proceso de identificación, generando así más tiempo.
- Muchas veces los expertos no han podido identificar las diversas plantas que existen en el Perú.
- El experto es capaz de interactuar con dispositivos electrónicos.

4.3.3 Análisis

Tabla 13: Tabla de Indicadores para análisis

Indicador	Conocimiento Experto	Ingeniería de Conocimiento
Identificación de la planta	Se utiliza la comparación de la planta con otros ejemplares que ya estén identificados, consultando un herbario o un jardín entre otros, al utilizar este proceso se requiere de una gran cantidad de tiempo.	La identificación de la planta se puede realizar por dos formas, la primera es por la forma que presenta dicha planta, y lo segundo es por los dientes, ya que son únicos, la cual va a requerir poco tiempo.
Localización de la planta	Para localizar nuevamente una determinada planta, tenemos que memorizar o recordar donde fue ubicada exactamente, esto genera que ya no podamos volver a encontrar dicha planta.	Una vez encontrada la planta, el software almacenara los datos de la ubicación para una próxima búsqueda de una determinada planta.

Fuente: Elaboración Propia

4.3.4 Diseño de métodos para recolectar conocimiento adicional

Tabla 14: Recolectar conocimiento adicional

Conocimiento Adicional	Tamaño de la planta
	El color que presenta
	El color del tallo
	La forma
	La longitud de la planta

Fuente: Elaboración Propia

- Los medios que se usarán para la extracción de conocimiento serán:
 - ✓ Entrevista: Anexo escrita y audio
 - ✓ Observación: Anexo

4.3.5 Historias de usuarios

Tabla 15: Historia Usuario - Abrir imagen

HISTORIA DE USUARIO	
Numero: 1	Usuario: Experto
Nombre de la historia: Abrir imagen	Dependencia para su desarrollo: No aplica
Prioridad en el negocio: Alta	Riesgo en desarrollo: Media
Puntos asignados: 3	Sprint Asignado: 1
Responsable: Vallejos Fonseca Cristian	
Descripción: Como experto quiero obtener la foto de la planta donde se va a proceder una secuencia para la identificación y localización	
Criterio de aceptación: <ul style="list-style-type: none"> ✓ Obtener imagen 	

Tabla 16: Historia Usuario - Filtros

HISTORIA DE USUARIO	
Numero: 2	Usuario: Experto
Nombre de la historia: Filtros	Dependencia para su desarrollo: Abrir imagen
Prioridad en el negocio: Alta	Riesgo en desarrollo: Media
Puntos asignados: 3	Sprint Asignado: 1
Responsable: Vallejos Fonseca Cristian	
Descripción: Filtrar a las plantas de Quinoa consiste en aplicar una transformación para obtener una nueva imagen de forma que ciertas características son acentuadas o disminuidas	
Criterio de aceptación: <ul style="list-style-type: none"> ✓ Filtro ✓ Abrir imagen (puede utilizarse cualquier foto) 	

Tabla 17: Historia Usuario - Histograma de la Imagen

HISTORIA DE USUARIO	
Numero: 3	Usuario: Experto
Nombre de la historia: Histograma de la imagen	Dependencia para su desarrollo: Abrir imagen
Prioridad en el negocio: Alta	Riesgo en desarrollo: Media
Puntos asignados: 3	Sprint Asignado: 1
Responsable: Vallejos Fonseca Cristian	
Descripción: histograma nos permitirá clasificar los píxeles de la imagen en fondo y objeto con total garantía sin más que atender a su nivel de gris	
Criterio de aceptación: ✓ Ecuilizar	

Tabla 18: Historia Usuario – Segmentación

HISTORIA DE USUARIO	
Numero: 4	Usuario: Experto
Nombre de la historia: Segmentación	Dependencia para su desarrollo: Abrir imagen
Prioridad en el negocio: Alta	Riesgo en desarrollo: Media
Puntos asignados: 3	Sprint Asignado: 1
Responsable: Vallejos Fonseca Cristian	
Descripción: Designa una serie de transformaciones que, normalmente a partir de la imagen procesada, la divide en partes, cada una de las cuales muestra una fuerte correlación con algún objeto o región.	
Criterio de aceptación: ✓ Otsu	

Tabla 19: Historia Usuario - Extraer características

HISTORIA DE USUARIO	
Numero: 5	Usuario: Experto
Nombre de la historia: Extraer características	Dependencia para su desarrollo: Abrir imagen
Prioridad en el negocio: Alta	Riesgo en desarrollo: Media
Puntos asignados: 4	Sprint Asignado: 1
Responsable: Vallejos Fonseca Cristian	
Descripción: El experto verificará la extracción de características obtenidas de la imagen para ser reconocida, utilizando Matlab	
Criterio de aceptación: ✓ Obtener características	

Tabla 20: Historia Usuario - Reconocimiento

HISTORIA DE USUARIO	
Numero: 6	Usuario: Experto
Nombre de la historia: Reconocimiento	Dependencia para su desarrollo: Extracción de características
Prioridad en el negocio: Alta	Riesgo en desarrollo: Media
Puntos asignados: 5	Sprint Asignado: 1
Responsable: Vallejos Fonseca Cristian	
Descripción: Si se nos da una imagen de analizar, como el retrato de grupo en la figura, podríamos tratar de aplicar un algoritmo de reconocimiento de toda la posible sub- ventana en esta imagen.	
Criterio de aceptación: ✓ Red neuronal	

Tabla 21: Historia Usuario - Inicio Sesión

HISTORIA DE USUARIO	
Numero: 7	Usuario: Experto
Nombre de la historia: Inicio de Sesión	Dependencia para su desarrollo: Usuarios registrados
Prioridad en el negocio: Alta	Riesgo en desarrollo: Media
Puntos asignados: 3	Sprint Asignado: 1
Responsable: Vallejos Fonseca Cristian	
Descripción: Como experto quiero iniciar sesión al software, ingresando mi correo electrónico y mi contraseña	
Criterio de aceptación: ✓ Iniciar sesión	

Tabla 22: Historia Usuario - Cerrar Sesión

HISTORIA DE USUARIO	
Numero: 8	Usuario: Experto
Nombre de la historia: Cerrar sesión	Dependencia para su desarrollo: Usuarios registrados
Prioridad en el negocio: Media	Riesgo en desarrollo: Baja
Puntos asignados: 3	Sprint Asignado: 3
Responsable: Vallejos Fonseca Cristian	
Descripción: Como experto requiero cerrar sesión para que otras personas no realicen modificaciones a los datos	
Criterio de aceptación: ✓ Cerrar sesión	

4.4 Fase 3: Diseño

4.4.1 Seleccionar técnica de representación de conocimiento

Se hará uso del método basado en reglas para poder definir los problemas y que así el sistema logre dar respuestas, en este caso a través de entrenamiento de imágenes.

4.4.2 Seleccionar Técnica de control

Se hará uso de indicadores (hipótesis) que el experto debe y se comprueba la veracidad de esta haciendo uso de las distintas herramientas y sus resultados.

4.4.3 Seleccionar software para el desarrollo del sistema inteligente

Se hará uso de Matlab para la extracción de características de las imágenes de las plantas y el gestor de base de datos MySQL lo cual es de suma importancia ya que la interfaz gráfica del sistema inteligente será escritorio y será desarrollada con el lenguaje de JAVA.

4.4.4 Prototipo de interfaces

Diseño de Entrada Al Software

Pantalla de cómo se verá al iniciar el software



*Figura 13: Diseño de entrada del software
Fuente: Elaboración Propia*

Inicio de Sesión para los Usuarios

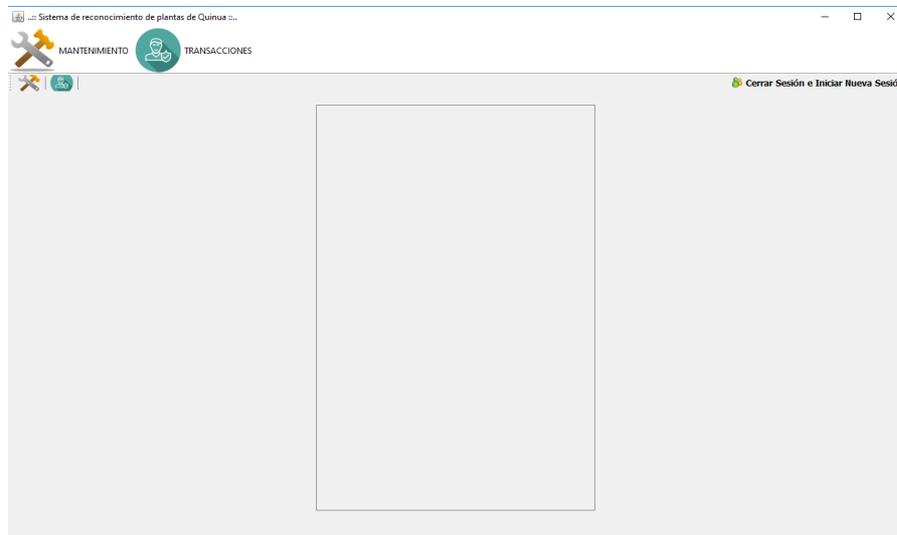
Ingresar datos específicos correctos, para el ingreso



*Figura 14: Inicio de sesión para los usuarios
Fuente: Elaboración Propia*

Menú Principal

De esta manera se verá el software al iniciar correctamente



*Figura 15: Menú Principal
Fuente: Elaboración Propia*

Mantenimiento

Acá podemos observar el mantenimiento que se realizará para poder reconocer una determinada imagen

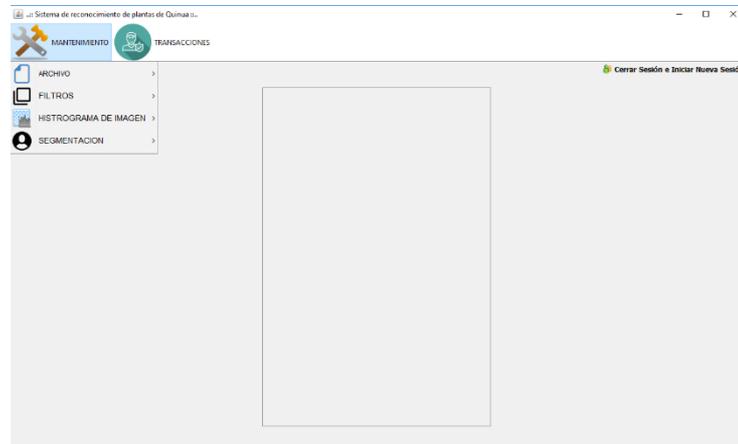


Figura 16: Mantenimiento
Fuente: Elaboración Propia

Mantenimiento – Abrir Imagen

Primero se debe seleccionar la plata que queremos identificar

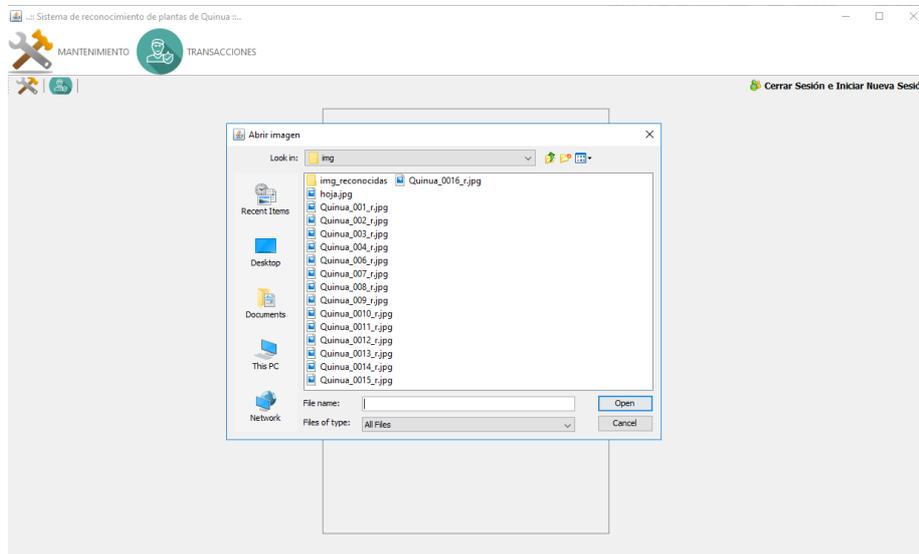


Figura 17: Mantenimiento - Abrir imagen
Fuente: Elaboración Propia

Mantenimiento – Filtros

Acá vamos a mostrar los diferentes filtros que se van a realizar, la cual se ha realizado de una forma automática que será escogido por el sistema.

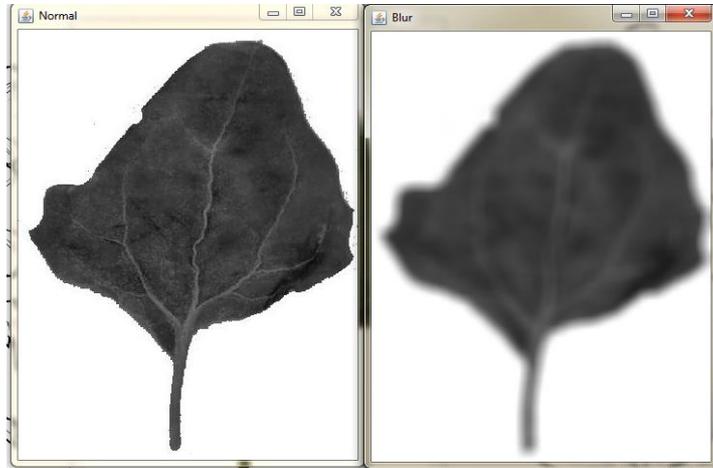


Figura 18: Mantenimiento - Filtros
Fuente: Elaboración Propia

Mantenimiento – Histograma

Se realizara la ayuda de un histograma para suavizar la imagen

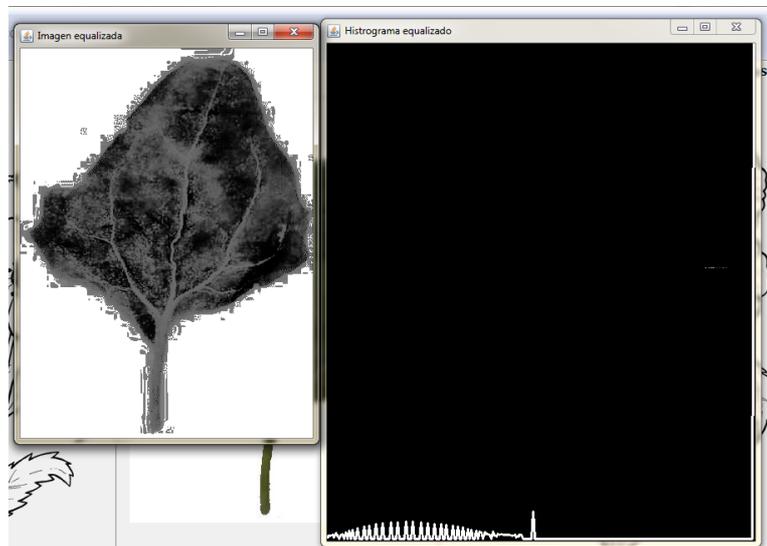


Figura 19: Mantenimiento - Histograma
Fuente: Elaboración Propia

Mantenimiento – Segmentación

Con la ayuda de la segmentación podemos convertir la imagen a blanco y negro

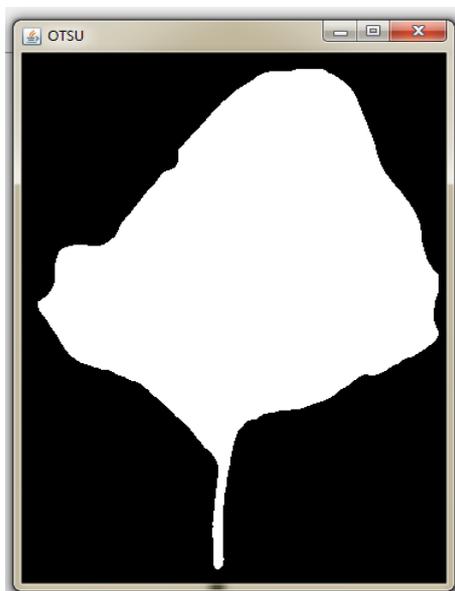


Figura 20: Mantenimiento - Segmentación
Fuente: Elaboración Propia

Transacciones

Las transacciones consiste en dos partes, la cual es la extracción de las características de la imagen, y ya el proceso de reconocimiento

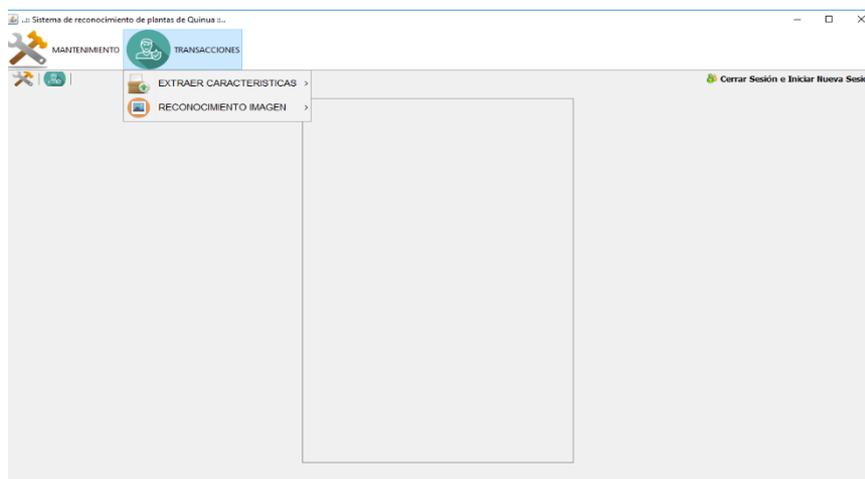


Figura 21: Transacciones
Fuente: Elaboración Propia

Transacciones – Extracción De Características

Tenemos que seleccionar la plata, luego ya seleccionada vamos a tener que realizar la obtención de características

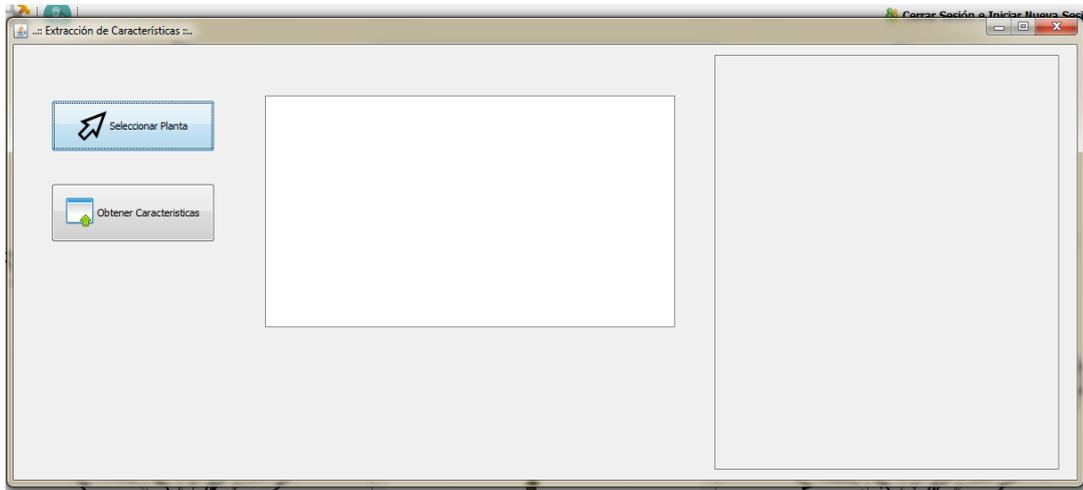


Figura 22: Transacciones - Extracción de características
Fuente: Elaboración Propia

Transacciones – Reconocimiento De Imagen

Luego de haber obtenido las características, pasamos al proceso de identificación, la cual vamos a tener que entrenar dicha planta, y luego reconocerla



Figura 23: Transacciones - Reconocimiento de imagen
Fuente: Elaboración Propia

4.4.5 Glosario de términos:

4.5.1 Sistema inteligente:

Programa de inteligencia artificial diseñado especialmente para resolver problemas complejos y tomar decisiones.

4.5.2 Ingeniero de conocimiento:

Es aquel individuo que ha construido el sistema en su totalidad, por lo tanto, conoce a profundidad las funcionalidades del software.

4.4.6 Tarjeta CRC

LOGUEO

Responsabilidad	Colaboración
Verifica	Experto

EXPERTO

Responsabilidad	Colaboración
Entrenamiento	Experto
Reconocimiento	Experto

4.4.7 Producto

4.4.7.1 Demostración del Software

4.4.7.1.1 Iniciar Sesión

El inicio de sesión permitirá la entrada de usuarios de los que están registrados en la base de datos, con su correo electrónico y contraseña



The image shows a software window titled "INICIO DE SESION". Inside the window, there is a logo on the left consisting of a stack of books with a green plant sprout growing from the top. To the right of the logo, the text "INICIAR SESIÓN" is displayed in a large, bold, black font. Below this, there is a section titled "DATOS PERSONALES" containing two input fields: "USUARIO" with the text "cristian@hotmail.com" and "CONTRASEÑA" with a masked password of seven dots. At the bottom of the form, there are two buttons: "ACEPTAR" with a green plus icon and "SALIR" with a red minus icon.

Figura 24: Demostración - Inicio sesión
Fuente: Elaboración Propia

4.4.7.1.2 Adquisición de Imágenes

La adquisición de las imágenes es uno de los pasos más importantes en un sistema de visión artificial, ya que la calidad de las imágenes a analizar debe tener una calidad óptima, que no afecte al rendimiento del sistema.

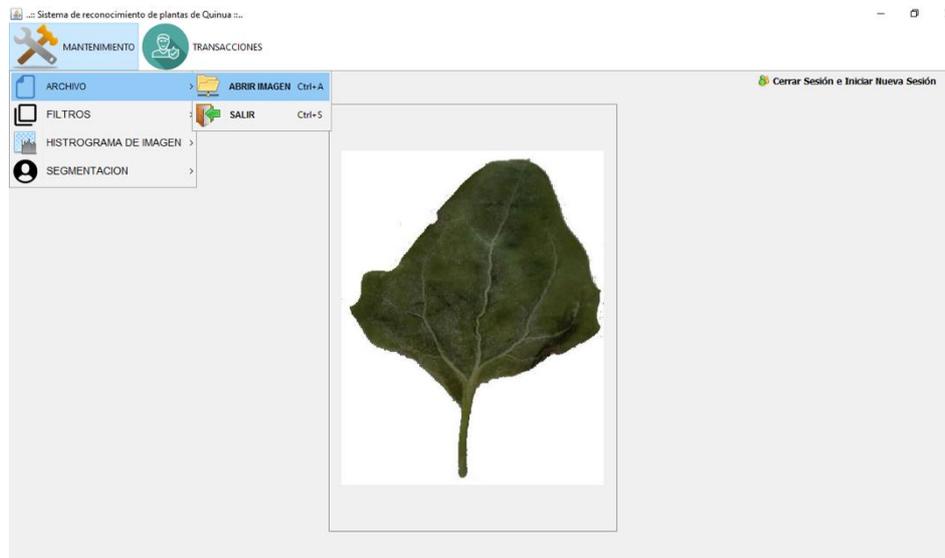


Figura 25: Adquisición de imágenes
Fuente: Elaboración Propia

4.4.7.2 Pre-procesamiento

En la etapa de pre procesamiento se pretende corregir defectos que puedan producirse durante la captura y digitalización de las imágenes las placas radiográficas. Si bien se ha tomado las precauciones para obtener una muestra lo más limpia posible.

4.4.7.2.1 Filtrado de imagen

Es necesario filtrar la imagen para quitar ruido, para realizar esto es necesario reducir la amplitud de las variaciones de la imagen, una forma simple de hacer esto, es reemplazar cada píxel por la media del valor de los píxeles de alrededor, de esta forma las variaciones rápidas de intensidad pueden ser suavizadas y reemplazadas por una transición más gradual.

4.4.7.2.2 Blur

La fórmula matemática en la que se basa el filtro blur es la siguiente:

$$g(i, j) = \sum_{k, l} f(i + k, j + l)h(k, l)$$

En la cual el valor de un pixel de salida (es decir, $g(i, j)$) se determina como una suma ponderada de los valores de los pixeles de entrada (es decir $f(i + k, j + l)$). La implementación del filtro blur en open CV está determinado de la siguiente manera:

```
Mat src = Funciones.cargar(ruta, false, false, 0);
Imgproc.cvtColor(src, src, Imgproc.COLOR_BGR2GRAY);
Funciones.mostrar(src, 700, 300, "Normal");
Mat blur = new Mat();
Imgproc.blur(src, blur, new Size(15, 15));
Funciones.mostrar(blur, 700, 300, "Blur");
```

Si aplicamos la función Blur de OpenCV sobre la imagen original obtenemos el siguiente resultado:



Figura 26: Imagen filtro Blur
Fuente: Elaboración Propia

4.4.7.2.3 Gaussian Blur

La fórmula matemática en la que se basa el gaussian blur es la siguiente:

$$f(x) = ae^{-(x-b)^2/c^2}$$

En la cual la distancia entre el pixel que se modificara (es decir, $(x - b)$) y el pixel del que se está tomando el valor (es decir a y c) valen 1 lo cual resulta un número real entre 0 y 1 que representara el porcentaje de cambio.

La implementación del filtro blur en open CV está determinado de la siguiente manera:

```

Mat src = Funciones.cargar(ruta, false, false, 0);
Imgproc.cvtColor(src, src, Imgproc.COLOR_BGR2GRAY);
Funciones.mostrar(src, 700, 300, "Normal");
Mat GaussianBlur = new Mat();
//GaussianBlur (origen, destino, tamaño kernel gaussiano, desviacion en X)
Imgproc.GaussianBlur(src, GaussianBlur, new Size(11, 11), 0);
Funciones.mostrar(GaussianBlur, 700, 300, "GaussianBlur");

```

Si aplicamos la función Gaussian Blur de OpenCV sobre la imagen original obtenemos el siguiente resultado:

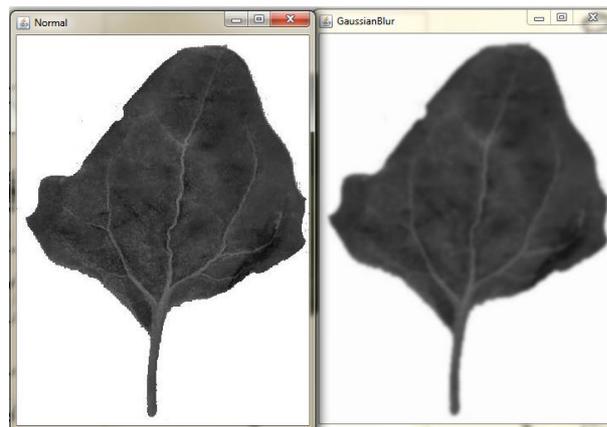


Figura 27: Imagen filtro Gaussian Blur

Fuente: Elaboración Propia

4.4.7.2.4 Filtro Bilateral

La fórmula matemática en la que se basa el filtro bilateral es la siguiente:

$$W_p = \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|I_p - I_q|).$$

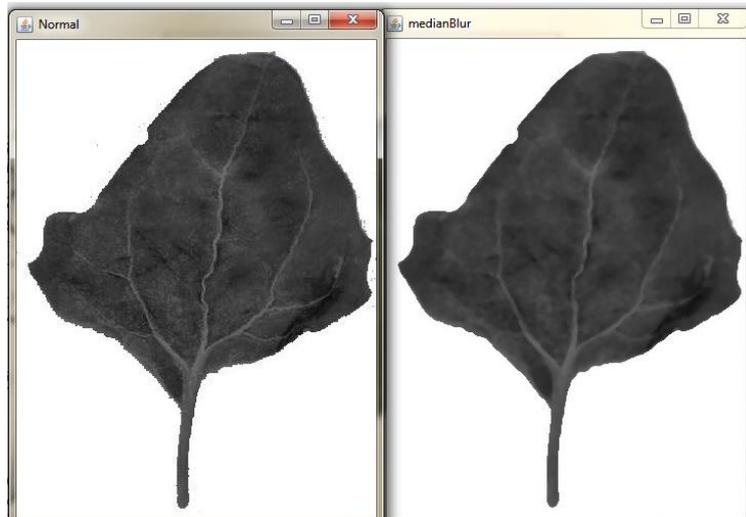
La implementación del filtro blur en open CV está determinado de la siguiente manera:

```

Mat src = Funciones.cargar(ruta, false, false, 0);
Imgproc.cvtColor(src, src, Imgproc.COLOR_BGR2GRAY);
Funciones.mostrar(src, 700, 300, "Normal");
Mat bilateralFilter = new Mat();
//bilateralFilter (origen, destino, tamaño, sigma color, sigma espacio)
Imgproc.bilateralFilter(src, bilateralFilter, 9, 75, 75);
Funciones.mostrar(bilateralFilter, 700, 300, "bilateralFilter");

```

Si aplicamos la función Filtro Bilateral de OpenCV sobre la imagen original obtenemos el siguiente resultado:



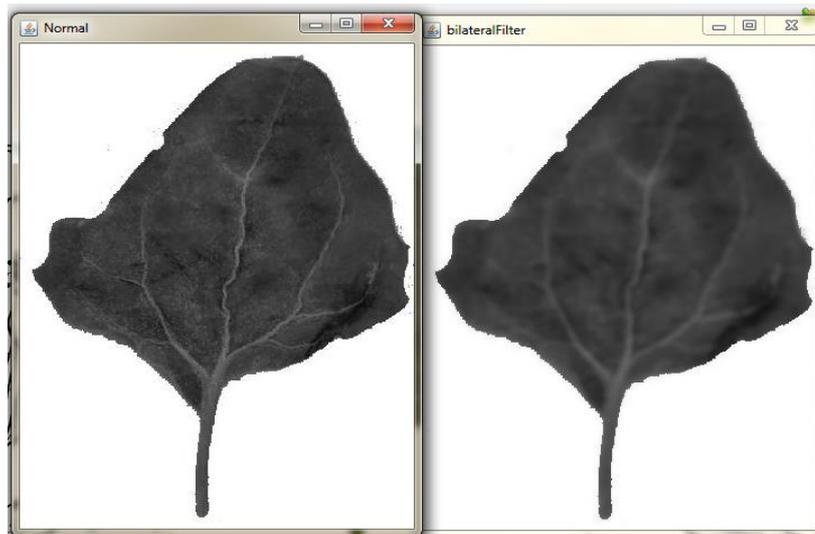
*Figura 28: Imagen Filtro Bilateral
Fuente: Elaboración Propia*

4.4.7.2.5 Median Blur

La implementación del filtro blur en OpenCV está determinado de la siguiente manera:

```
Mat src = Funciones.cargar(ruta, false, false, 0);  
Imgproc.cvtColor(src, src, Imgproc.COLOR_BGR2GRAY);  
Funciones.mostrar(src, 700, 300, "Normal");  
Mat medianBlur = new Mat();  
//medianBlur (origen, destino, tamaño)  
Imgproc.medianBlur(src, medianBlur, 5);  
Funciones.mostrar(medianBlur, 700, 300, "medianBlur");
```

Si aplicamos la función Median Blur de OpenCV sobre la imagen original obtenemos el siguiente resultado:

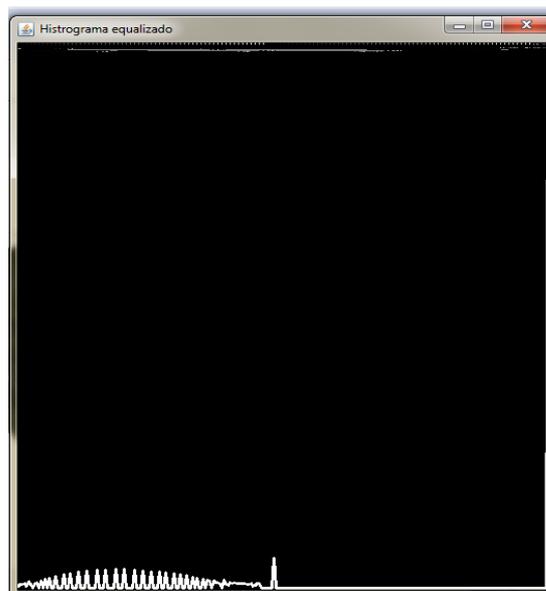


*Figura 29: Imagen Median Blur
Fuente: Elaboración Propia*

4.4.7.2.6 Histograma

Ecualización del Histograma

El histograma después de ecualización se muestra en la siguiente figura:



*Figura 30: Imagen Histograma
Fuente: Elaboración Propia*

Esto nos puede dar perfecta distribución de los valores de los píxeles para su posterior procesamiento.

4.4.7.2.7 Segmentación

A continuación, están los resultados después de la binarización de la imagen:

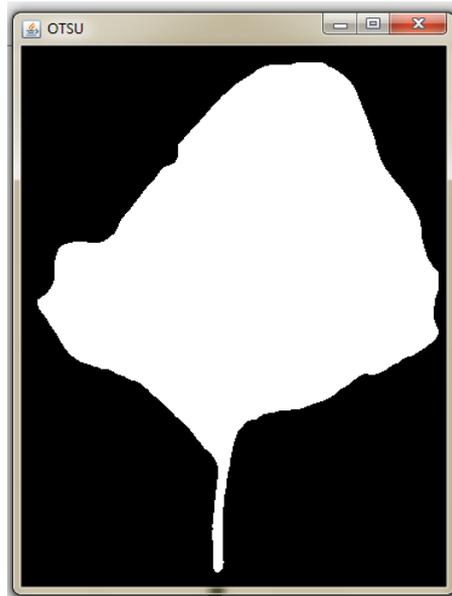


Figura 31: Imagen Segmentación
Fuente: Elaboración Propia

```

Imgproc.threshold(src, hsvImg, 60, 179.0, Imgproc.THRESH_BINARY_INV);
Imgproc.threshold(src, hsvImg2, 60, 179.0, Imgproc.THRESH_BINARY);

Imgproc.blur(hsvImg, hsvImg, new Size(5, 5));
Imgproc.blur(hsvImg2, hsvImg2, new Size(5, 5));

Imgproc.dilate(hsvImg, hsvImg, new Mat(), new Point(-1, 1), 6);
Imgproc.erode(hsvImg, hsvImg, new Mat(), new Point(-1, 1), 6);

Imgproc.dilate(hsvImg2, hsvImg, new Mat(), new Point(-1, 1), 6);
Imgproc.erode(hsvImg2, hsvImg, new Mat(), new Point(-1, 1), 6);

Imgproc.threshold(hsvImg, hsvImg, 60, 179.0, Imgproc.THRESH_BINARY_INV);
Imgproc.threshold(hsvImg2, hsvImg2, 60, 179.0, Imgproc.THRESH_BINARY);

Mat foreground = new Mat(src.size(), CvType.CV_8UC3, new Scalar(255, 255, 255));
Mat foreground2 = new Mat(src.size(), CvType.CV_8UC3, new Scalar(255, 255, 255));

```

4.4.7.2.8 Método Otsu

A continuación, están los resultados después de la segmentación con el método OTSU.

```

Mat src = Funciones.cargar(ruta, false, false, 0);
Imgproc.cvtColor(src, src, Imgproc.COLOR_BGR2GRAY);

Imgproc.blur(src, src, new Size(5, 5));

Imgproc.equalizeHist(src, src);

Imgproc.threshold(src, src, 60, 255, Imgproc.THRESH_OTSU);;

```

4.4.7.2.9 Extracción de Características

Para la extracción de características se procedió a utilizar la herramienta de visión artificial y reconocimiento de patrones Balu Toolbox, que será implementada en Matlab y permitirá extraer las características de las imágenes.

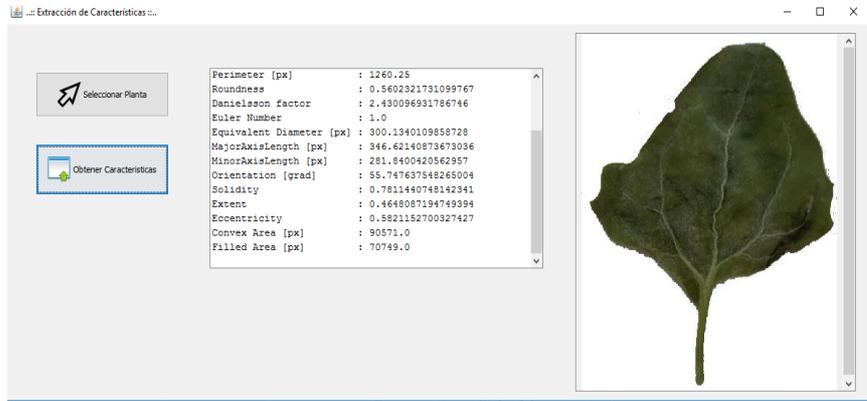


Figura 32: Imagen de extracción de características
Fuente: Elaboración Propia

Extracción de características geométricas básicas

Característica	Valor
Center of Grav i [px]	114,7829
Center of Grav j [px]	188,8274
Height [px]	220,0000
Width [px]	348,0000
Area [px]	23729,0000
Perimeter [px]	1957,0000
Roundness	0,0779
Danielsson Factor	10,1742
Euler Number	1,0000
Equivalent Diameter [px]	173,5945
MajorAxisLength [px]	380,2082
MinorAxisLength [px]	182,8533
Orientation [grad]	-4,3068
Solidity	0,3541
Extent	0,3091
Eccentricity	0,8768
Convex Area [px]	66838,0000
FilledArea [px]	23668,0000

Figura 33: Extracción de vector de características geométricas básicas

La conexión mediante netbeans se da de la siguiente manera:

```
MatlabProxyFactoryOptions options = new MatlabProxyFactoryOptions.Builder()
    .setUsePreviouslyControlledSession(true)
    .setHidden(true)
    .setMatlabLocation(null).build();
```

```
MatlabProxyFactory factory = new MatlabProxyFactory(options);
MatlabProxy proxy = factory.getProxy();
```

Luego se procede a enviar la imagen y recibir sus características

```
proxy.eval("I = imread('" + ruta + "');\n"
    + "R = Bim_segbalu(I);\n"
    + "[X,Xn] = Bfx_basicgeo(R);\n"
    + "Bio_printfeatures(X,Xn)");
```

```
double[] result = ((double[]) proxy.getVariable("X"));
String[] result2 = ((String[]) proxy.getVariable("Xn"));
```

```
for (int i = 0; i < result.length; i++) {
    System.out.println(result2[i] + " : " + result[i]);
    JTextArea2.append(result2[i] + " : " + result[i] + "\n");
}
```

```
MatlabProxyFactoryOptions options = new MatlabProxyFactoryOptions.Builder()
    .setUsePreviouslyControlledSession(true)
    .setHidden(true)
    .setMatlabLocation(null).build();
```

```
MatlabProxyFactory factory = new MatlabProxyFactory(options);
MatlabProxy proxy = factory.getProxy();
```

```
//proxy.eval("disp('hello world')");
//Set a variable, add to it, retrieve it, and print the result
proxy.eval("I = imread('" + ruta + "');\n"
    + "R = Bim_segbalu(I);\n"
    + "[X,Xn] = Bfx_basicgeo(R);\n"
    + "Bio_printfeatures(X,Xn)");
```

```
double[] result = ((double[]) proxy.getVariable("X"));
String[] result2 = ((String[]) proxy.getVariable("Xn"));
```

```
for (int i = 0; i < result.length; i++) {
    System.out.println(result2[i] + " : " + result[i]);
    JTextArea2.append(result2[i] + " : " + result[i] + "\n");
}
```

```
//Disconnect the proxy from MATLAB
proxy.disconnect();
```

Acá observamos los resultados dados por el software, verificando la disminución de errores:

Tabla 23: Disminuir los errores de comparación

Planta	Cantidad Errores (Antes)	Mejoras de Errores (Después)
1	1 mes	10 minutos
2	1 mes	10 minutos
3	1 mes	10 minutos
4	1 mes	10 minutos
5	1 mes	10 minutos
6	1 mes	10 minutos
7	1 mes	10 minutos
8	1 mes	10 minutos
9	1 mes	10 minutos
10	1 mes	10 minutos
11	1 mes	10 minutos
12	1 mes	10 minutos
13	1 mes	10 minutos
14	1 mes	10 minutos
15	1 mes	10 minutos
16	1 mes	10 minutos

- Los 10 minutos es la demora de procesar la disminución de errores de identificación de la planta de quinua.

Clasificación:

Para la clasificación de las muestras obtenidas en base a las características extraídas por los descriptores se ha escogido dos clasificadores: Redes Neuronales

4.4.7.3 Red Neuronal Back Propagation

Para determinar la estructura de la red neuronal se procedió a hacer un análisis de sensibilidad, lo cual es necesario para obtener un error global más cerca de 0.

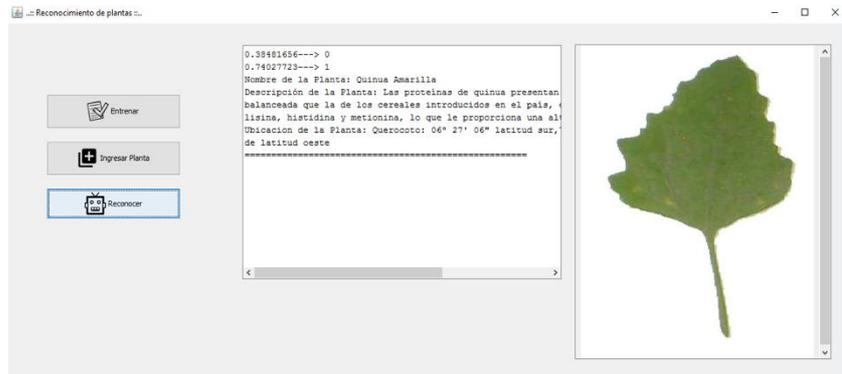


Figura 34: Red Neuronal Back Propagation
Fuente: Elaboración Propia

En esta tabla verificamos el tiempo de demora que da por resultado el software, generando así una mayor factibilidad en reducción de tiempo en el proceso de identificación.

Tabla 24: Reducir el tiempo de identificación

Planta	Tiempo
1	74 miliseg
2	21 miliseg
3	22 miliseg
4	20 miliseg
5	21 miliseg
6	23 miliseg
7	21 miliseg
8	21 miliseg
9	19 miliseg
10	23 miliseg
11	20 miliseg
12	22 miliseg
13	23 miliseg
14	21 miliseg
15	23 miliseg
16	23 miliseg

- *Los milisegundos dependerá a la demora que tome el proceso de los filtros para la identificación de la planta de quinua.*

4.4.7.4 Base de Datos

En la base de datos se van almacenar todas las plantas que se van a identificar, para una próxima comparación, para ellos se creó un módulo llamado “patrones_reconocer”.

patrones_reconocer	
IdPatronReconocer	int
altura	varchar(255)
anchura	varchar(255)
area	varchar(255)
perimetro	varchar(255)
redondez	varchar(255)
orientacion	varchar(255)
solidez	varchar(255)
grado	varchar(255)
excentricidad	varchar(255)
area_convexa	varchar(255)
area_llena	varchar(255)
codigoUsuario	int

Figura 35: Tabla Patrones_Reconocer

La consulta que se utilizó para visualizar las plantas identificadas es la siguiente:

```
select * from patrones_reconocer
```

Gracias a esta consulta se pudo visualizar con mayor exactitud las plantas identificadas durante el proceso de prueba del sistema, el acceso a identificar la planta de Quinua fue más factible

Tabla 25: Plantas identificadas

Plantas Identificadas (Antes)	Plantas Identificadas (Después)
3	16

- *Cabe resaltar que esta identificación se dio durante el proceso de mantenimiento y enseñanza a los expertos de la empresa, actualmente se han realizado más entrenamientos por parte de la empresa.*

- ❖ Aquí se muestra todas las plantas identificadas, almacenadas en una base de datos

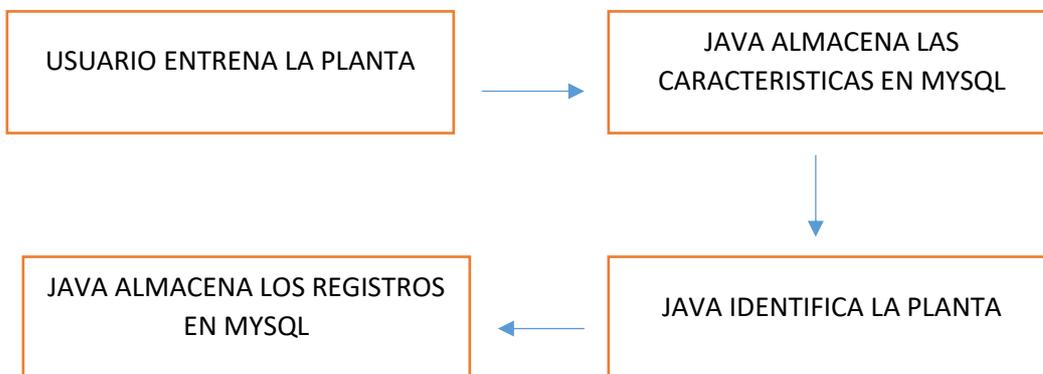
Figura 36: Tabla de Patrones Reconocer

idPatronReconocer	altura	anchura	area	perimetro	redondez	orientacion	solidez	grado	excentricidad
1	449.0	339.0	70806.25	1260.25	0.5602321731099767	55.747637548265004	0.7811440748142341	0.4648087194749394	0.58211527
2	421.0	280.0	48233.875	1214.0	0.4112677395085029	74.6213857025511	0.6936962874458376	0.4087970817780794	0.51762336
3	449.0	339.0	70806.25	1260.25	0.5602321731099767	55.747637548265004	0.7811440748142341	0.4648087194749394	0.58211527
4	421.0	280.0	48233.875	1214.0	0.4112677395085029	74.6213857025511	0.6936962874458376	0.4087970817780794	0.51762336
5	421.0	280.0	48233.875	1214.0	0.4112677395085029	74.6213857025511	0.6936962874458376	0.4087970817780794	0.51762336
6	449.0	339.0	70806.25	1260.25	0.5602321731099767	55.747637548265004	0.7811440748142341	0.4648087194749394	0.58211527
7	449.0	339.0	70806.25	1260.25	0.5602321731099767	55.747637548265004	0.7811440748142341	0.4648087194749394	0.58211527
8	421.0	280.0	48233.875	1214.0	0.4112677395085029	74.6213857025511	0.6936962874458376	0.4087970817780794	0.51762336
9	425.0	306.0	61303.625	1156.0	0.5764748324924976	66.28201962941127	0.7924726035372812	0.47098039215686277	0.57945699
10	444.0	289.0	54544.25	1275.75	0.42114115557420967	-69.8870038961927	0.7007678555351194	0.4246079900246265	0.57558895
11	438.0	290.0	53547.625	1244.0	0.4348198046886952	-60.75825278881294	0.709659874774488	0.4211698945048024	0.69028501
12	427.0	303.0	52813.875	1238.5	0.43267905788103267	-89.28857067596053	0.693166592226735	0.4078496824658505	0.44070128
13	435.0	318.0	63438.375	1185.5	0.567229404360668	60.23645433650204	0.7774061082148596	0.4582014024434324	0.55322798
14	449.0	339.0	70806.25	1260.25	0.5602321731099767	55.747637548265004	0.7811440748142341	0.4648087194749394	0.58211527
15	444.0	289.0	54544.25	1275.75	0.42114115557420967	-69.8870038961927	0.7007678555351194	0.4246079900246265	0.57558895
16	449.0	339.0	70806.25	1260.25	0.5602321731099767	55.747637548265004	0.7811440748142341	0.4648087194749394	0.58211527

Fuente: Elaboración Propia

4.5 Fase 5: Documentación

4.5.1 Funcionamiento del Sistema Inteligente



4.6 Resultados en Base a los objetivos

A continuación se detalla los resultados en base a los objetivos establecidos en el proyecto de investigación:

4.6.1 Composición de la Muestra

Para el análisis de los objetivos específicos, se realizó el modelo de Pre Test y Post Test, en donde se contrasta la situación actual del proceso con

el sistema tradicional y la situación posterior con la aplicación del sistema propuesto, por ser una muestra menor a 30 elementos, se aplicó la prueba t-student:

Tabla 26: Composición de la muestra de las plantas de quinua del distrito de Querocoto - Cajamarca

GRUPO	PLANTAS
EXPERIMENTAL	8 plantas de quinua
CONTROL	8 plantas de quinua

Fuente: Elaboración propia

4.6.1.1 Análisis estadísticos del objetivo: Mejorar el nivel de certeza de los técnicos de campo para localizar las diversas variedades de la planta de Quinua.

a) Formulación de la Hipótesis

Ho: $\mu_c = \mu_e$: El rendimiento de resolución de problemas de combinación es igual en el grupo control y en el grupo experimental.

H1: $\mu_c < \mu_e$: El rendimiento de resolución de problemas de combinación es mayor en el grupo experimental.

b) Nivel de significancia: $\alpha = 0.05$

c) Estadístico de prueba

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_c^2}{n_1} + \frac{S_e^2}{n_2}}}$$

$$S_c^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

d) Región Crítica

$$t_{(1-\alpha, n_1 + n_2 - 2)} = t_{(0.05, 14)} = 1.7613$$



e) Valor Estadístico

Tabla 27: Prueba t para dos muestras suponiendo varianzas iguales.
Localización de las diversas variedades de plantas de quinua

	Control	Experimental
Media	0.5	0.875
Varianza	0.28571429	0.125
Observaciones	8	8
Varianza agrupada	0.20535714	
Diferencia hipotética de las medias	0	
Grados de libertad	14	
Estadístico t	-1.65503185	
P(T<=t) una cola	0.06007692	
Valor crítico de t (una cola)	1.76131014	
P(T<=t) dos colas	0.12015383	
Valor crítico de t (dos colas)	2.14478669	

Fuente: Elaboración propia

f) Decisión

$t_c = -1.65 \in$ a la Región Crítica

✓ Se rechaza H_0

g) Conclusión

A un 95% de confianza se estima la certeza de la localización de las plantas de quinua es mayor en el grupo experimental

• **Prueba de hipótesis de muestras pareadas (pre – post test en el nivel de certeza de los técnicos de campo para localizar las diversas variedades plantas de Quinua) –GRUPO EXPERIMENTAL**

a) Formulación de la Hipótesis

$H_0: \mu_{pre} = \mu_{post}$: El sistema para las plantas de quinua no tiene efecto significativo en la localización de plantas de quinua basado en la resolución de problemas de combinación.

$H_1: \mu_c < \mu_e$: El sistema para las plantas de quinua tiene efecto significativo en la localización de plantas de quinua basado en la resolución de problemas de combinación.

b) Nivel de significancia: $\alpha = 0.05$

c) Estadístico de prueba

$$T = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}} \quad \text{una t con n-1 g.l.}$$

d) Región Crítica

$$t_{(1-\alpha, n_1 + n_2 - 2)} = t_{(0.05, 7)} = 1.89$$



e) Evaluación Crítica

Tabla 28: Prueba t para dos muestras suponiendo varianza relacionadas.
Localización de las diversas variedades de plantas de quinua.

	Control	Experimental
Media	0.5	0.875
Varianza	0.28571429	0.125
Observaciones	8	8
Coefficiente de correlación de Pearson	0.37796447	
Diferencia hipotética de las medias	0	
Grados de libertad	7	
Estadístico t	-2.04939015	
P(T<=t) una cola	0.03980101	
Valor crítico de t (una cola)	1.89457861	
P(T<=t) dos colas	0.07960201	
Valor crítico de t (dos colas)	2.36462425	

Fuente: Elaboración propia

f) Decisión

$t_c = -2.04 \in$ a la Región Crítica

✓ Se rechaza H_0

g) Conclusión

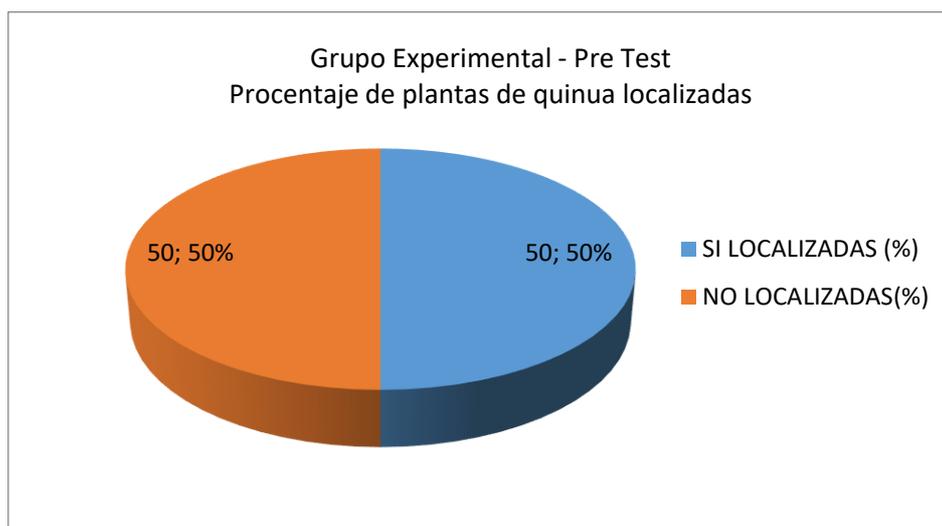
Con un 95% de confianza se estima que el software tiene efecto significativo en la localización de las plantas de quinua basado en la resolución de problemas de combinación.

- **Porcentaje de la localización de las plantas de quinua, pre test-post test - GRUPO EXPERIMENTAL.**

*Tabla 29: Porcentaje de la localización de las plantas de quinua.
Grupo Experimental*

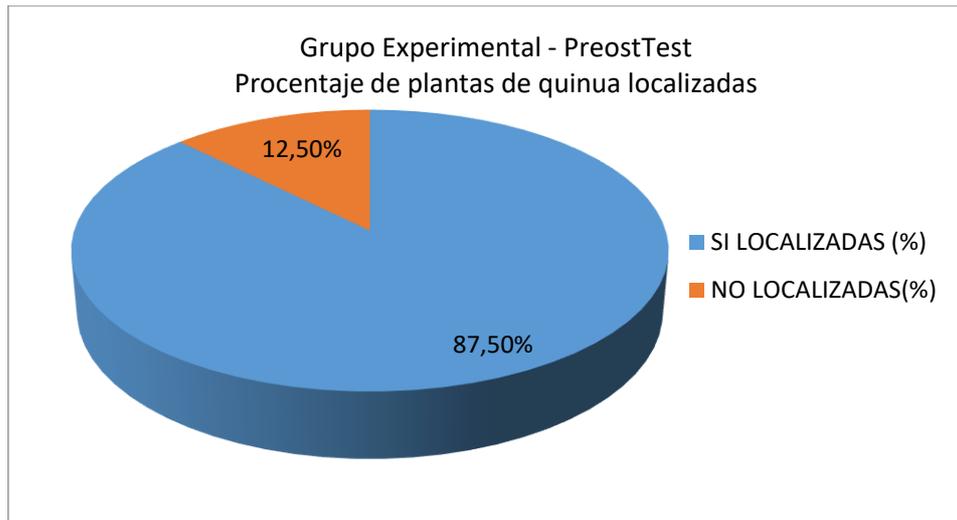
GRUPO EXPERIMENTAL		
Plantas de quinua	PRE TEST	POST TEST
	Localización	Localización
1	SI	SI
2	SI	SI
3	NO	SI
4	NO	SI
5	SI	SI
6	NO	SI
7	SI	SI
8	NO	NO
SI (%)	50%	87.5%
NO (%)	50%	12.5%

Fuente: Elaboración propia



*Figura 37: Porcentaje de plantas de quinua localizadas.
Grupo Experimental – Pre Test
Fuente: Elaboración propia*

Como se muestra en la figura, se hizo una evaluación pre-test al grupo experimental y obtuvimos los siguientes resultados, plantas localizadas 50% y un 50% no lograron localizar.



*Figura 38: Porcentaje de plantas de quinua localizadas.
Grupo Experimental – Post Test*
Fuente: Elaboración propia

Una vez aplicado el software de localización de las diversas variedades plantas de quinua, el grupo experimental en el post test, los resultados fueron favorables obteniendo el 87.5% de plantas de quinua localizadas y un 12.5 % de plantas de quinua no localizadas.

4.6.1.2 Análisis estadístico del objetivo: Incrementar el nivel de eficiencia de los expertos para identificar las diversas variedades de la planta de Quinua.

a) Formulación de la Hipótesis

Ho: $\mu_c = \mu_e$: El rendimiento de resolución de problemas de combinación es igual en el grupo control y en el grupo experimental.

H1: $\mu_c < \mu_e$: El rendimiento de resolución de problemas de combinación es mayor en el grupo experimental.

b) Nivel de significancia: $\alpha = 0.05$

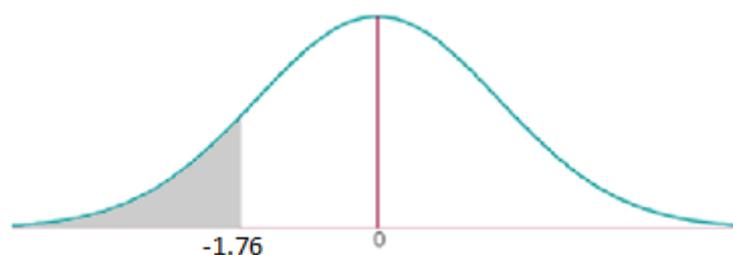
c) Estadístico de prueba

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_c^2}{n_1} + \frac{S_e^2}{n_2}}}$$

$$S_c^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

d) Región Crítica

$$t_{(1-\alpha, n_1 + n_2 - 2)} = t_{(0.05, 14)} = 1.76$$



e) Valor Estadístico

Tabla 30: Prueba t para dos muestras suponiendo varianzas iguales. Identificación de las diversas variedades de plantas de quinua.

	Control	Experimental
Media	0.625	1
Varianza	0.267857143	0
Observaciones	8	8
Varianza agrupada	0.133928571	
Diferencia hipotética de las medias	0	
Grados de libertad	14	
Estadístico t	-2.049390153	
P(T<=t) una cola	0.029823083	
Valor crítico de t (una cola)	1.761310136	
P(T<=t) dos colas	0.059646166	
Valor crítico de t (dos colas)	2.144786688	

Fuente: Elaboración propia

f) Decisión

$t_c = -2.04 \in$ a la Región Crítica

✓ Se rechaza H_0

g) Conclusión

A un 95% de confianza se estima la eficiencia de la identificación de las plantas de quinua es mayor en el grupo experimental

- **Prueba de hipótesis de muestras pareadas (pre – post test en el nivel de eficiencia de los expertos para la identificación de las diversas variedades plantas de Quinua) –GRUPO EXPERIMENTAL**

a) Formulación de la Hipótesis

Ho: $\mu_{pre} = \mu_{post}$: El sistema para las plantas de quinua no tiene efecto significativo en la identificación de plantas de quinua basado en la resolución de problemas de combinación.

H1: $\mu_c < \mu_e$: El sistema para las plantas de quinua tiene efecto significativo en la identificación de plantas de quinua basado en la resolución de problemas de combinación.

b) Nivel de significancia: $\alpha = 0.05$

c) Estadístico de prueba

$$T = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}} \quad \text{una } t \text{ con } n-1 \text{ g.l.}$$

d) Región Crítica

$$t_{(1-\alpha, n_1 + n_2 - 2)} = t_{(0.05, 7)} = 1.89$$



e) Evaluación Crítica

Tabla 31: Prueba t para dos muestras suponiendo varianzas relacionadas. Identificación de las diversas variedades de plantas de quinua.

	Control	Experimental
Media	0.625	1
Varianza	0.267857143	0
Observaciones	8	8
Coefficiente de correlación de Pearson	0,0	
Diferencia hipotética de las medias	0	
Grados de libertad	7	
Estadístico t	-2.049390153	
P(T<=t) una cola	0.039801006	
Valor crítico de t (una cola)	1.894578605	
P(T<=t) dos colas	0.079602012	
Valor crítico de t (dos colas)	2.364624252	

Fuente: Elaboración propia

- f) Decisión
 $t_c = - 2.04 \epsilon$ a la Región Crítica
 ✓ Se rechaza H_0

- g) Conclusión
 Con un 95% de confianza se estima que el software tiene efecto significativo en la identificación de las plantas de quinua basado en la resolución de problemas de combinación.

• **Porcentaje de la identificación de las plantas de quinua, pre test-post test - GRUPO EXPERIMENTAL.**

Tabla 32: Porcentaje de la identificación de las plantas de quinua. Grupo Experimental

GRUPO EXPERIMENTAL		
Plantas de quinua	PRE TEST	POST TEST
	Identificación	Identificación
1	SI	SI
2	NO	SI
3	SI	SI
4	SI	SI
5	SI	SI
6	NO	SI
7	SI	SI
8	NO	SI
SI (%)	62.5%	100%
NO (%)	37.5%	0%

Fuente: Elaboración propia

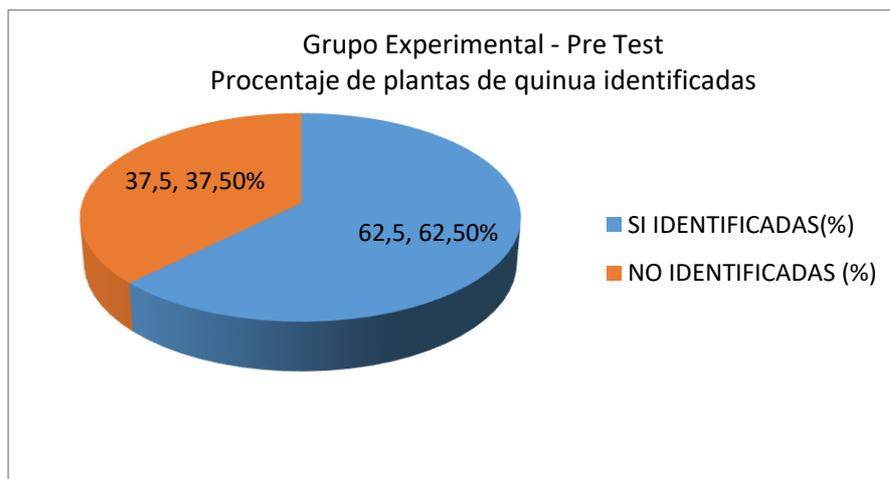
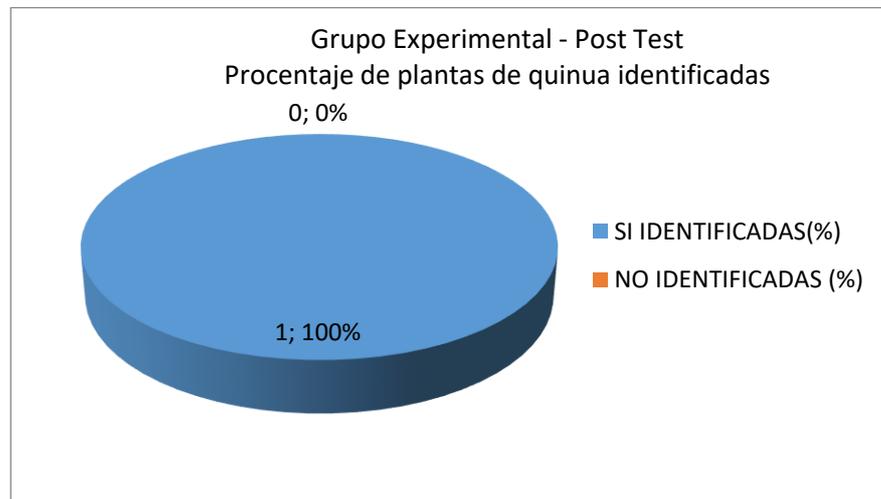


Figura 39: Porcentaje de plantas de quinua identificadas. Grupo Experimental – Pre Test
 Fuente: Elaboración propia

Como se muestra en la figura, se hizo una evaluación pre-test al grupo experimental y obtuvimos los siguientes resultados, plantas identificadas 62.5% y un 37.5% no lograron identificar.



*Figura 40: Porcentaje de plantas de quinua identificadas.
Grupo Experimental - Post Test
Fuente: Elaboración propia*

Una vez aplicado el software de identificación de las diversas variedades plantas de quinua, el grupo experimental en el post test, los resultados fueron favorables obteniendo el 100% de plantas de quinua identificadas y un 0 % de plantas de quinua no identificadas.

CAPÍTULO IV DISCUSIÓN

En este capítulo se hará en análisis de resultados de software del sistema inteligente desarrollado. La evaluación realizada se enfoca en los objetivos definidos en el informe:

Tabla 33: Objetivos

OBJETIVOS
Mejorar el nivel de certeza de los técnicos de campo para localizar las diversas variedades de la planta de quinua.
Incrementar el nivel de eficiencia de los expertos para identificar las diversas variedades de la planta de quinua.

Para cumplir el primer objetivo que es mejorar el nivel de certeza de los técnicos de campo para localizar las diversas variedades de la planta de quinua, se utilizó la teoría de Luis Lanuza Sandoval [37], nos dice que el GPS nos permite en determinar todo el mundo la posición de un objeto, una persona, un carro o un avión, con una precisión hasta de centímetros usando un GPS diferencial, aunque habitual son unos pocos metros. Este sistema fue desarrollado e instalado y actualmente operado por el Departamento de Defensa de los Estados Unidos. En el proyecto de investigación de Luis Lanuza Sandoval propone una solución para el problema de la localización de carros robados basado en un sistema con en el uso de las tecnologías GPS, GPRS Y GSM, gracias a los avances de la tecnología es posible usar estas tres tecnologías en conjunto para poder determinar la ubicación de un carro, generando una forma sencilla y económica para determinar dicha ubicación y velocidad en cualquier punto en la tierra. Es por ello que para esta investigación se utilizó la posibilidad de usar las cámaras fotográficas de los teléfonos inteligentes y uso de los GPS integrados en estos teléfonos, ya que en los metadatos quedan registradas las coordenadas geográficas del lugar donde se tomó la foto. (Figura 20)

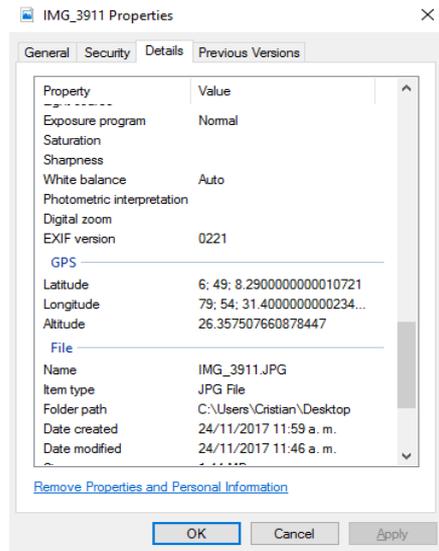


Figura 41: Demostración de GPS de una imagen

Fuente: Elaboración propia

La consulta que se utilizó para verificar la ubicación de la planta con más exactitud es la siguiente:

```
select * from planta
```

idPlanta	nombrePlanta	descripcionPlanta	ubicacionPlanta
1	Quinua Blanca	Es un cereal andino, de excepcional valor nutritivo rico	Querocoto: 06° 27' 06" latitud sur, 78° 57' 41" de latitud norte de latitud oeste
2	Quinua Morada	En comparación con la blanca, la quinua morada prop	Querocoto: 06° 27' 06" latitud sur, 78° 57' 41" de latitud norte de latitud oeste
3	Quinua Amarilla	Las proteínas de quinua presentan una proporción de	Querocoto: 06° 27' 06" latitud sur, 78° 57' 41" de latitud norte de latitud oeste

Figura 42: Tabla planta

Fuente: Elaboración propia

Por otro lado según Pérez, Botella y Muñoz [38] la geolocalización hace referencia al posicionamiento de un punto espacial en un sistema de coordenadas y datos, estos sistemas se llaman sistemas de información geográfica y se definen como una integración de hardware, software y datos geográficos con el fin de capturar, almacenar, manipular y analizar la información geográfica referenciada de ellos; un sistema de información geográfica es un modelo de parte de la realidad referido a unas coordenadas geográficas que tienen el objetivo de satisfacer necesidades concretas de información, en estos sistemas podemos incluir puntos de interés que son puntos espaciales con información geográfica como: longitud, latitud o la altura; la ventaja de la geolocalización se pueden observar en el que mediante el etiquetado de esos puntos, el usuario puede encontrar una amplia variedad de información sobre un lugar específico, es por ello que

esta teoría se eligió para la presente investigación, ya que nos ayuda a tener un registro más ordenado y eficiente de las diversas localizaciones que presentan las variedades de plantas de quinua, y así ser utilizadas para una próxima búsqueda con mayor rapidez, disminuyendo el tiempo de su búsqueda.

Para cumplir el segundo objetivo que es incrementar el nivel de eficiencia de los expertos para identificar las diversas variedades de la planta de quinua, se utilizó la teoría de Hugo Vega Huerta [39], nos dice que el uso computacional de las redes neuronales artificiales, viene creciendo exponencialmente, ya que mediante su empleo se encuentra cada vez más soluciones de problemas complejos en las diferentes áreas de conocimiento humano, no como secuencia de pasos, sino como la evaluación de sistemas computacionales inspirados en el cerebro humano y que por tanto dotados de cierta inteligencia operan de forma paralela. El proyecto de investigación de Hugo Vega es un sistema experto que permite reconocer los mangos cuyas formas cumplen con los estándares de calidad exigidos por los clientes de Estados Unidos y Europa, es por ello que el software está basado en el proceso de imágenes digitales de mangos mediante redes neuronales artificiales, dando como conclusión que el diseño de una red neuronal artificial ha sido capaz de aprender a reconocer a los mangos de la muestra, generando mejoras para la empresa. La teoría propuesta por Hugo Huerta ayudó en la implementación de la clasificación e identificación de las diversas variedades de plantas de quinua, luego con la utilización del software, se toma una imagen digital de cada una de ellas, con dichas imágenes entrenamos una red neuronal para que reconozca acertadamente todos los elementos de la muestra.

Así mismo, Ponce Cruz [28] que nos dice que las redes neuronales artificiales son un método de resolver problemas, de forma individual o combinadas con otros métodos, para aquellas tareas de clasificación, identificación, diagnóstico, optimización o predicción en las que el balance de conocimiento se inclina hacia los datos y donde, adicionalmente, puede haber la necesidad de aprendizaje en tiempo de ejecución y de cierta tolerancia a fallos. Las redes neuronales comúnmente se pueden configurar para representar funciones o reconocer patrones. La ventaja de la configuración es que el tiempo de entrenamiento se reduce considerablemente en relación con otros algoritmos y además se garantiza una buena velocidad de convergencia. La teoría propuesta por Ponce Cruz se eligió para este trabajo, diseñándose un módulo de registro a través del cual los expertos podrían visualizar las plantas ya reconocidas de una manera más rápida, sin

necesidad de estar buscando en un herbario u otro proceso de identificación, dicha información se almacenó en la tabla “patrones_reconocer”. (Figura 15).

Finalmente, cabe resaltar que, según los datos obtenidos, se considera verdadera la hipótesis propuesta al inicio del desarrollo de la tesis: “La implementación de un software que permita identificar y localizar de manera rápida las variedades de quinua a través del proceso de reconocimiento de imágenes”. Dado que se ha logrado alcanzar notablemente los objetivos propuestos.

CONCLUSIONES

El desarrollo de este trabajo de investigación de tesis ha permitido contribuir correctamente a la identificación y localización de las diversas variedades de plantas que se presentan en nuestro país, permitiendo así que los técnicos de campo y a los expertos puedan acceder con facilidad a sus procesos, se concluye que:

1. Se implementó una función del sistema que mejoró la certeza en la localización de una determinada muestra de planta de quinua; en el pre test se obtuvo un porcentaje del 50% en plantas de quinua localizadas y un 50 % de plantas de quinua no localizadas, y en el pos test con un 87.5% de plantas de quinua localizadas y un 12.5% de plantas de quinua no localizadas. Así mismo, el tiempo de búsqueda manual era de 1:30 a 3:00 horas, y ahora con el software el tiempo de búsqueda disminuye entre 30 minutos a 1 hora.
2. La aplicación de las tecnologías de reconocimiento de imágenes, incrementó la eficiencia de los expertos en la identificación de las plantas de quinua, disminuyendo los errores en el proceso de identificación, en el pre test se obtuvo un porcentaje del 62.5% de plantas de quinua identificadas y un 37.5% de plantas de quinua no identificadas, en el post test con un 100% de plantas de quinua identificadas y un 0% de plantas de quinua no identificadas, esto se debe a que el software propuesto siempre dará un resultado, ya sea correcto o similar a una planta de quinua. Así mismo, el tiempo de identificación de una hoja en forma manual tomo aproximadamente 30 días, y con el software toma aproximadamente 10 minutos.

Finalmente, se debe tener en cuenta que a pesar de los tiempos de pruebas de resultados ha disminuido considerablemente gracias a la automatización de los procesos, la cual ha permitido disminuir los tiempos de identificación y localización, aunque aún es necesario que se actualicen cada vez estos procesos ya mencionados.

RECOMENDACIONES

1. Esta investigación logro identificar las plantas de quinua con imágenes en formato jpeg, pero se recomienda para futuras investigaciones realizar la detección en imágenes con formato dcm.
2. Esta investigación reconoce las plantas de quinua basada a un aprendizaje con redes neuronales y redes bayesianas, dejando la posibilidad para futuras investigaciones puedan realizar el reconocimiento con SVM o Lógica difusa aplicada a las imágenes.
3. Se sugiere proponer más capacitaciones o talleres sobre reconocimiento, o identificación de plantas a los expertos.
4. Se sugiere brindar una orientación fotográfica a los expertos, para facilitar el proceso de identificación.
5. Esta investigación logro identificar las plantas de quinua utilizando el lenguaje de programación Java, pero se recomienda implementar los algoritmos de pre procesamiento y de procesamiento en otro lenguaje de programación.
6. Se recomienda implementar a futuras un sistema de geolocalización, para la localización de las plantas de quinua.

REFERENCIAS BIBLIOGRÁFICAS

- [1] P. C. Araujo Farro, «Repositorio de Producción Científica e Intelectual de la UNICAMP,» 29 marzo 2008. [En línea]. Available: <http://repositorio.unicamp.br/handle/REPOSIP/256477>.
- [2] A. C. Candela, «"SISTEMA DE IDENTIFICACIÓN, EXTRACCIÓN Y RECUPERACIÓN DE,» España, 2011.
- [3] S. V. Pérez, «MÓVILES, IDENTIFICACIÓN DE ESPECIES VEGETALES UTILIZANDO DISPOSITIVOS,» Chile, 2013.
- [4] M. d. C. A. Vélez, «"INTEGRACIÓN DEL DISEÑO CENTRADO EN USUARIO CON,» Perú, 2015.
- [5] P. Garcia, «Reconocimiento de imagenes utilizando redes neuronales artificiales,» España, 2013.
- [6] G. Ivan, «Visión artificial y procesamiento digital de imágenes usando Matlab,» ISBN, Ecuador, 2008.
- [7] A. Marín, «Desarrollo de prototipo de aplicacion (app), para dispositivos móviles basados en el sistema operativo ios, para el reconocimiento de objetos "hojas" en imagenes,» 2014.
- [8] J. S. H. P. Javier Redolfi, «Identificación de hojas de plantas usando vectores de Fisher,» Argentina, 2015.
- [9] D. Betancourt, «Sistema de visión por computadora para detectar hierba no deseada en prototipo de cultivo de frijol usando ambiente controlado,» Bogotá, 2014.
- [10] E. Garcia, «Desarrollo de un sistema de vision artificial por capturas y en tiempo real, como primera de un macroproyecto para la inspeccion visual automatica de cultivos de frijol,» Mexico, 2013.
- [11] N. Vignale, «Identificacion Micrografica de las hojas de Smallanthus sonchifolius,» Buenos Aires, 2005.
- [12] A. H. R. B. B. B. Stefany Gonzales, «Identificacion de especies meloidogyne asociadas a plantas ornamentales de altura en costa rica,» Costa Rica, 2015.
- [13] J. Gutierrez, «Desarrollo de un sistema inteligente usando redes hopfield para el reconocimiento de rostros,» Tujillo, 2013.
- [14] B. B. A. T. Amparo Betanzos, Ingeniería del conocimiento, Madrid: Farescp, S.A, 2004.
- [15] «EcuRed,» 14 Diciembre 2010. [En línea]. Available: http://www.ecured.cu/Lenguaje_de_Programaci%C3%B3n.
- [16] A. Gutierrez, «Ingeniería del conocimiento segun durkin,» 2008.

- [17] A. G. y. N. d. A. Raul Pino, «Sistemas expertos, redes neuronales artificiales y computación evolutiva,» Servicios de Publicaciones Universidad de Oviedo, Oviedo, 2001.
- [18] R. C. & W. R. E. Gonzales, Procesamiento de Imágenes Digitales, 2008.
- [19] E. G. Santillán, «Detección y clasificación de objetos dentro,» Mexico, 2006.
- [20] E. Sucar, «Visión computacional,» Mexico, 2008.
- [21] J. Z. H. & Z. J. Wu, Digital Image Processing: Part I., Madrid, 2010.
- [22] P. P. G. García, «Reconocimiento de imágenes utilizando redes neuronales artificiales,» Madrid, 2013.
- [23] C. A. S. GUTIÉRREZ, «SISTEMA DE LOCALIZACIÓN EN ESPACIOS INTERIORES PARA UN ROBOT MÓVIL UTILIZANDO NUBES DE PUNTOS Y MODELOS OCULTOS DE MARKOV,» Mexico, 2017.
- [24] E. D. L. F. F.M.TRESPADERNE, «Imágenes binarias».
- [25] R. W. P. Gonzalez, «Procesamiento digital de imágenes,» 1996.
- [26] R. Szeliski, «Computer Vision: Algorithms and Applications,» 2010.
- [27] O. Boullosa Gracia, Estudio comparativo de descriptores visuales para la detección de escenas cuasi-duplicadas, 2008.
- [28] P. Ponce Cruz, Inteligencia Artificial con Aplicaciones a la Ingeniería. Alfaomega, 2010.
- [29] R. & F. J. M. Florez, Las Redes Neuronales Artificiales, Fundamentos teóricos y aplicaciones prácticas, 2008.
- [30] P. G. A. A. Diaz Marcos, «Sistemas Manejadores de Base de Datos,» Mexico, 2010.
- [31] «Rozvo Ware Solutions,» [En línea]. Available: <http://www.rozvo.com/rz/index.php/blog/blog-3/29-cum-sociis-natoque-8>.
- [32] P. Salamon, «Grupo Publimail,» 2009. [En línea]. Available: <http://www.publimail.pe/index.php/soporte/preguntas-frecuentes/119-que-es-mysql>.
- [33] J. Pavón M., «Programación Orientada a Objetos Facultad de Informática,» Madrid, 2008.
- [34] L. Gomez, «Estado del arte de la quinua en el mundo,» Lima, 2014.
- [35] Y. B. Diaz, «Rendimiento de cuatro variedades de quinua,» 2016.
- [36] M. Morales, «Modelo de gestión del conocimiento basado en las difusas para evaluar el desempeño laboral,» 2010.
- [37] L. D. L. Sandoval, «Sistema de localización por medio de SMS o GPRS,» Guatemala, 2008.

- [38] B. y. M. Pérez, «Introducción a los sistemas de información geográfica y geotelemática,» UOC, España, 2011.
- [39] H. F. V. Hierta, «Redes neuronales para el reconocimiento de la calidad morfológica de mangos exportables para la empresa Biofruit del Perú S.A.C,» Lima, 2011.
- [40] «A&J&G MICH MEXICO,» [En línea]. Available:
<http://www.tiendasvirtualesycomercioweb.com/iqes-base-de-datos-mysql>.
- [41] M. Azucena, «Desarrollo de prototipo de aplicación, para dispositivos móviles basados en el sistema operativo ios, para el reconocimiento de objetos "hojas" en imágenes,» España, 2014.
- [42] J. Durkin, «Expert System: Design and Development,» Macmillan, New York, 1994.

ANEXOS
ANEXO N° 1



Facultad de Ingeniería
Escuela de Ingeniería de Sistemas y
Computación

PARTICIPANTES: Experto en el proceso de plantas

INSTRUCCIONES: Se deben de responder todas las preguntas de forma correcta, de acuerdo a sus conocimientos.

1. **¿Cuánto tiempo tarda un exportador en identificar una variedad de plantas de quinua?**
2. **¿Cuáles son las variedades de quinua?**
3. **¿Qué dificultad que tiene un exportador para poder localizar las variedades de plantas de quinua?**
4. **¿Cuánto puede afectar las enfermedades que presenta la quinua en la aplicación móvil?**
5. **¿Cómo es el proceso de identificación de la planta de quinua?**
6. **¿Cómo clasifican las plantas de quinua?**
7. **¿Qué enfermedad presenta la quinua?**
8. **¿Cuál es la evaluación de quinua?**
9. **¿Cuáles son las fases que presenta la quinua?**
10. **¿Qué tipo de quinua ha tenido más éxito en el mercado mundial?**
11. **¿Cuál es el país que más quinua exporta?**
12. **¿Cómo le gustaría que le ayude mi aplicación?**

ANEXO N° 2



Facultad de Ingeniería
Escuela de Ingeniería de Sistemas y
Computación

ENTREVISTA AL BIÓLOGO

Nombre: Carlos Sialer

Cargo: Biólogo

Fecha: 27/06/2016

1. ¿Cuánto tiempo tarda un exportador en identificar una variedad de plantas de quinua?

Aproximadamente un mes

2. ¿Cuáles son las variedades de quinua?

En el Perú existen alrededor de 18 variedades. Entre ellos se eligió dos variedades por motivo que estas son sembradas en lugar donde se desarrollara el dicho proyecto.

- Chenopodium
- Pasankalla

3. ¿Qué dificultad que tiene un exportador para poder localizar las variedades de plantas de quinua?

Las limitaciones tecnológicas y las dificultades que experimentan los productores de menores recursos.

4. ¿Cuánto puede afectar las enfermedades que presenta la quinua en el software?

El deterioro puede comprometer la hoja y no se lograría la identificación de la imagen de la planta

5. ¿Cómo es el proceso de identificación de la planta de quinua?

Para la identificación de la planta son los siguientes pasos:

- ✓ Tamaño y modo de vida (hierba, arbusto, árbol)
- ✓ Color de las flores y frutos
- ✓ Localidad donde obtuvo la planta (ser lo más específico posible)
- ✓ La fecha en que usted colectó el material de planta
- ✓ Especificar si la muestra es cultivada o silvestre
- ✓ Enumere cada muestra

6. ¿Cómo clasifican las plantas de quinua?

La podemos clasificar de acuerdo a los dientes que presenta cada planta y por su forma que muestra

7. ¿Qué enfermedad presenta la quinua?

-Mildiu: enfermedad que presenta la quinua, en la que pone amarilla a la hoja

8. ¿Cuál es la evaluación de quinua?

-Evaluación de la actividad antagonista en plántulas de quinua: a 25 días de germinación.

- a) Plántulas de quinua infectadas sin la aplicación de compuestos bioactivos (control positivo)
- b) Plántulas de quinua infectadas con fusarium
- c) Plántulas de quinua infectadas con rhizoctonia sp. Tratadas con el compuesto bioactivo de la cepa

-Evaluación de plagas y enfermedades en el cultivo de quinua

9. ¿Cuáles son las fases que presenta la quinua?

- **Fases Fenológicas de la Quinua**
 - Germinación
 - Emergencia
 - Dos hojas verdaderas
 - Cuatro hojas verdaderas
 - Seis hojas verdaderas
 - Rominificación
 - Inicio de panojamiento
 - Panojamiento
 - Inicio de floración

10. ¿Qué tipo de quinua ha tenido más éxito en el mercado mundial?

El tipo de quinua que más se exporta es el grupo de la llamada **quinua real**, que se cultiva en las tierras altas del sur de Bolivia, se caracteriza por un grano de gran tamaño, cuyo diámetro supera los 2,2 mm. La quinua orgánica, independientemente de su variedad, es objeto de una gran demanda en mercados diferenciados.

11. ¿Cuál es el país que más quinua exporta?

Bolivia es el principal exportador de quinua en el mundo, seguido del Perú y el Ecuador.

Los principales importadores de quinua boliviana son actualmente los Estados Unidos, Francia, los Países Bajos, Alemania, el Canadá, Israel, el Brasil y el Reino Unido.

12. ¿Cómo le gustaría que le ayude mi aplicación?

A poder determinar cualquier modificación que pueda existir a nivel morfológico en las variedades de quinua al ser sembradas y producidas bajo diferentes condiciones edafoclimáticas y como estas pueden afectar su productividad.

ANEXO N° 3



Facultad de Ingeniería
Escuela de Ingeniería de Sistemas y
Computación

ENTREVISTA

PARTICIPANTES: Experto en el proceso de plantas

INSTRUCCIONES: La información proporcionada será anónima. Se agradece que responda a las siguientes preguntas.

1. ¿Cuánto tiempo tarda un exportador en identificar una variedad de plantas de quinua?

Aproximadamente tres meses

2. ¿Cuáles son las variedades de quinua?

Hay 16 variedades de plantas de quinua en el Perú.

3. ¿Qué dificultad que tiene un exportador para poder localizar las variedades de plantas de quinua?

Las dificultades de volver a encontrar dicha planta, ya que no se utiliza un registro.

4. ¿Cuánto puede afectar las enfermedades que presenta la quinua en el software?

Afectaría en que la planta al presentar esos rasgos, no se podría identificar correctamente.

5. ¿Cómo es el proceso de identificación de la planta de quinua?

- El método más rápido y fiable es recurrir a un experto, pero no siempre lo tenemos a disposición.

- Un segundo método consiste en la comparación de la planta con otros ejemplares que ya estén identificados, por ejemplo, consultando un herbario, un jardín botánico, un libro con descripciones completas e ilustraciones, etc., pero este método necesita una gran cantidad de tiempo y el contar con colecciones completas y fiables.

- El tercer método es utilizando claves de identificación.

6. ¿Cómo clasifican las plantas de quinua?

La calificación se da por medio de la forma que muestra cada variedad de plantas de quinua ya que son únicas

7. ¿Qué enfermedad presenta la quinua?

Las enfermedades se clasifican en:

- enfermedades del follaje
- enfermedades del tallo
- enfermedades de la raíz

8. ¿Cuál es la evaluación de quinua?

La evaluación de líneas mejoradas, es un proceso previo a la liberación de nuevas variedades, los caracteres agronómicos son frecuentemente expuestos al sol sobre lonas por separado, posteriormente se tomaron las muestras para evaluar las variables cuantitativas, peso y altura de planta, diámetro y longitud de panoja.

9. ¿Cuáles son las fases que presenta la quinua?

- **Fases Fenológicas de la Quinua**

Germinación

Emergencia

Dos hojas verdaderas

Cuatro hojas verdaderas

Seis hojas verdaderas

10. ¿Qué tipo de quinua ha tenido más éxito en el mercado mundial?

En el mercado se requiere la quinua de color, especialmente para su uso en la industria alimentaria o agrícola. Sin embargo, debido al desarrollo de su potencial gastronómico la demanda de quinua de color también se está incrementando.

11. ¿Cuál es el país que más quinua exporta?

Bolivia es el principal exportador de quinua en el mundo

12. ¿Cómo le gustaría que le ayude mi aplicación?

A poder identificar y localizar las diversas variedades de plantas que se presentan en nuestro país con mayor facilidad, evitando realizar un proceso que toma un determinado tiempo, y poder acceder muy fácilmente a ellas.