Detección y localización de malezas en cultivos de arroz basados en el análisis de imágenes de reflectancia y RGB tomadas desde drones.

Alejandro Carvajal Montealegre

Facultad de Ingeniería
Programa de Ingeniería Electrónica
Ibagué, 2020

Detección y localización de malezas en cultivos de arroz basados en el análisis de imágenes de reflectancia y RGB tomadas desde drones.

Alejandro Carvajal Montealegre

Trabajo de grado que se presenta como requisito parcial para optar al título de:
Ingeniero Electrónico

Director (a):
PhD. Ing. Oscar Barrero Mendoza.
Profesor Universidad de Ibagué

Facultad de Ingeniería
Programa de Ingeniería Electrónica
Ibagué, 2020

Este trabajo se lo dedico a mis padres, que gracias a su apoyo incondicional, esfuerzo y sacrificio, he podido llegar hasta donde me encuentro el día de hoy. Sin ellos nada de esto hubiera sido posible.

Agradecimientos

Mis más sinceros agradecimientos al tutor de esta tesis, el PhD. Ing. Oscar Barrero Mendoza. Sus conocimientos y apoyo brindados durante el desarrollo de este trabajo permitieron lograr los resultados obtenidos.

Agradezco a la Gobernación del Tolima y su financiación al proyecto “Sistema de alertas tempranas en cultivos de arroz, basado en imágenes de reflectancia, térmicas y RGB de alta resolución tomadas desde drones, con el fin de mejorar la productividad de los cultivos de arroz del departamento del Tolima”. Gracias a los recursos destinados se pudo desarrollar el proyecto, y por ende esta tesis que surge de allí.

Agradezco también a los distritos de riego Usocoello, Ususaldaña y Asorrecio por haber brindado los lotes donde se llevó a cabo el estudio de malezas. De igual modo agradezco a la Federación Nacional de Arroceros (Fedearroz), que en conjunto con los distritos de riego, brindaron personal y conocimiento necesario para el desarrollo del proyecto.

Por ultimo quiero agradecer a todas las personas que contribuyeron de alguna manera al desarrollo de esta tesis.

Resumen

Las malezas son plantas que crecen de forma silvestre en zonas cultivadas y que compiten por recursos tales como agua, luz y nutrientes con las plantas del cultivo. En la actualidad, la mayoría de agricultores hacen un control homogéneo de malezas con herbicidas. Esto es poco eficiente, ya que la distribución de malezas no es uniforme en los lotes, además trae consecuencias negativas como un gran impacto ambiental y un alto costo en los insumos para realizar dicha tarea. La agricultura de precisión se ha venido desarrollando últimamente, y consiste en la aplicación de tecnología para la observación y monitoreo de cultivos, con el fin de actuar solo en los sitios que presenten problemas, esto trae como beneficios un menor consumo de insumos y aumento de la producción, generando así un mayor beneficio a los agricultores y a la región. En este trabajo, se describe un método para detectar zonas con presencia de malezas en cultivos de arroz y todo su proceso, desde la toma de información usando vehículos aéreos no tripulados, hasta la generación de ortomosaicos, procesamiento de las imágenes, extracción de características y desarrollo de sistemas de clasificación supervisados. Finalmente se obtuvo como resultado tres sistemas de clasificación desarrollados bajo la modalidad de Random Forest, Maquina de Soporte Vectorial y K Vecinos más Cercanos. Los tres diferentes sistemas demostraron tener resultados muy similares, con desempeños cercanos al 90%.

Palabras clave: malezas, agricultura de precisión, vehículos aéreos no tripulados, procesamiento de imágenes, técnicas de clasificación supervisadas.

Abstract

Weeds are plants that grow in a wild way in cultivated areas, and compete for resources such as water, light and nutrients with crop plants. Nowadays, most farmers apply a homogeneous weed control with herbicides over the crop land. This is less productive due to the non- uniform weed distribution, besides of having negative consequences like great environmental impact and a high supplies cost to perform such task. Precision agriculture has been developing lately and consists of technology application to crop observation and monitoring, with the purpose of acting only in areas that present any problems, this technique bring great benefits such as: yield increase and less supplies consumption, generating in that way, a benefit to the farmer and the region. In this work, it is described a weed detection method in rice crops and all its process, beginning from capturing information using unmanned aerial vehicles, until orthomosaic generation, image processing, feature extraction and development of classification systems. Finally as a result, three classification systems under the modality of Random Forest, Support Vector Machine and K Nearest Neighbors were obtained. All three classification systems proved to have similar results, with performances close to 90%

Keywords: weeds, precision agriculture, unmanned aerial vehicles, image processing, supervised classification techniques.

Tabla de contenido

Resumen ..... **VIII**

Lista de figuras **1**

Lista de tablas **3**

1. Generalidades **4**

1.1. Introducción 4
1.2. Estado del arte 6
1.3. Objetivo general 9
1.3.1. Objetivos específicos 9

2. Marco teórico **10**

2.1. Arroz 10
2.2. Malezas 10
2.3. Teledetección 11
2.4. Índices vegetativos 12
2.4.1. Índice de Vegetación de Diferencia Normalizada (NDVI) 12
2.5. Segmentación 13
2.5.1. Binarización 13
2.6. Análisis de imágenes basado en objetos (OBIA) 15
2.6.1. Etiquetado de componentes conectados 15
2.7. Análisis de objetos basado en características de forma 16
2.7.1. Momentos invariantes de Hu 17
2.8. Análisis de objetos basados en características de textura 18
2.8.1. Matriz de co-ocurrencia 19
2.8.2. Descriptores de Haralick 21
2.9. Sistema de clasificación 23
2.9.1. Random Forest 25
2.9.2. Máquina de Soporte Vectorial 26
2.9.3. K-Nearest Neighbors (Vecinos más cercanos) 27

3. Materiales y métodos **28**

3.1. Materiales 28
3.1.1. Dron 28
3.1.2. Cámara RGB 29
3.1.3. Cámara multiespectral 29
3.2. Sitio de estudio 30
3.3. Localizacion puntos de control terrestre sobre los lotes 31
3.4. Adquisición de las imágenes 33

4. Desarrollo **35**

4.1. Generación Ortomosaicos 35
4.2. Procesamiento de las imágenes 38
4.3. Definición descriptores propuestos 41
4.4. Elección descriptores más importantes 43
4.4.1. Distribución normal de los datos de cada clase 43
4.4.2. Análisis de Componentes Principales (PCA) 47
4.4.3. Selección Secuencial de Características 50
4.5. Desarrollo sistemas de clasificación 51

4.5.1. Random Forest 52
4.5.2. Máquina de Soporte Vectorial 52
4.5.3. K Vecinos más Cercanos 52
4.6. Generación mapas con zonas de malezas 52

5. Resultados 54

5.1. Random Forest 55
5.2. Máquina de Soporte Vectorial 58
5.3. K Vecinos más Cercanos 60
5.4. Comparación entre los tres métodos de clasificación. 63

6. Conclusiones y recomendaciones 68

6.1. Conclusiones 68
6.2. Recomendaciones 69

Lista de figuras

Fig. 1.1 Publicaciones relacionados con las palabra "precision agriculture" a lo largo del tiempo vs número de publicaciones en la plataforma Scopus..... 7
Fig. 1.2 Publicaciones relacionadas con las palabras "weed detection" a lo largo del tiempo vs número de publicaciones en la plataforma Scopus..... 7
Fig. 2.1 Ortomosaico NDVI de un cultivo 12
Fig. 2.2 a) Imagen en escala de grises, b) binarización mediante el método Otsu de la imagen en escala de grises. 14
Fig. 2.3 Conectividad de cuatro y ocho pixeles..... 16
Fig. 2.4 a) Imagen binarizada, b) etiquetado de la imagen binarizada 16
Fig. 2.5 a) Ventana de 5x5 pixeles y en el centro el pixel que recibe el valor del cálculo, b) posición nueva de la ventana, c) pixeles de la imagen que reciben el valor del cálculo.. 19
Fig. 2.6 Los 8 vecinos del pixel de referencia X, de acuerdo al ángulo usado para el cálculo de la matriz de co-ocurrencia, para una distancia de 1 20
Fig. 2.7 a) Imagen en escala de grises, y b) su matriz de co-ocurrencia 20
Fig. 2.8 Clustering de datos de acuerdo a su ubicación espacial..... 24
Fig. 2.9 Representación gráfica de un sistema de clasificación supervisado 25
Fig. 2.10 Metodología de clasificación de un sistema Random Forest..... 26
Fig. 2.11 Selección del mejor hiperplano en una máquina de soporte vectorial lineal..... 26
Fig. 2.12 Dato a etiquetar y sus vecinos más cercanos 27
Fig. 3.1 Dron DJI Matrice 200..... 28
Fig. 3.2 Cámara Zenmuse X4S 29
Fig. 3.3 Cámara multiespectral + DLS 30
Fig. 3.4 a) Mapa de Colombia y el departamento de Tolima en azul, b) Mapa del Tolima y los municipios de Lérida, Espinal y Saldaña en amarillo..... 31
Fig. 3.5 Tablero de punto de control 31
Fig. 3.6 Ubicación de los marcadores en un polígono rectangular..... 32
Fig. 3.7 Ubicación de los marcadores en un polígono irregular 32
Fig. 3.8 Lote con: a) plantas muy pequeñas para hacer detección de malezas, b) lote con plantas muy grandes para hacer detección de malezas 33
Fig. 3.9 Ejemplo de la planeación de una ruta de vuelo..... 34
Fig. 3.10 Panel de calibración utilizado para la corrección radiométrica de las imágenes capturadas con la cámara multiespectral..... 34
Fig. 4.1 Visualización de una ruta de vuelo y los puntos en los que fueron adquiridos las imágenes..... 35
Fig. 4.2 Proyección nube de puntos de un lote a partir de las imágenes capturadas..... 36
Fig. 4.3 Ubicación de los puntos de control terrestres en la nube de puntos 36
Fig. 4.4 Ortomosaico generado a partir de las nubes de puntos a) RGB, b) NDVI en escala de grises 37
Fig. 4.5 a) Ortomosaico RGB, b) Ortomosaico NDVI, c) Ortomosaico NDVI con diferencia de colores más marcada 38
Fig. 4.6 a) Orientación real del ortomosaico NDVI, b) binarización del ortomosaico NDVI 40
Fig. 4.7 Etiquetado de los objetos presentes en la imagen binarizada..... 41
Fig. 4.8 Ortomosaico RGB + etiquetado de los objetos 41
Fig. 4.9 Función de densidad de probabilidad para cada una de las clases de los siguientes descriptores a) Contraste, b) Correlación, c) Energía, d) Homogeneidad, e) Longitud del eje mayor, f) Longitud del eje menor, g) Área, h) Diámetro equivalente, i)

Excentricidad, j) Solidez, k) Extensión, l) Momento 1, m) Momento 2, n) Momento 3, o) Momento 4, p) Momento 5, q) Momento 6, r) Momento 7, s) NDVI 44
Fig. 4.10 Descriptores resultantes del algoritmo que realiza el proceso de selección secuencial de características 51
Fig. 4.11 a) Ortomosaico RGB, b) Ortomosaico RGB + etiquetado de todos los objetos, c) Ortomosaico RGB + objetos etiquetados como maleza utilizando Random Forest y descriptores resultantes del PCA 53
Fig. 5.1 Matriz de confusión para los algoritmos de clasificación del tipo Random Forest utilizando los descriptores resultantes de: a) distribución normal, b) análisis de componentes principales (PCA), c) selección secuencial de características (SFS). 55
Fig. 5.2 a) Zona de un lote con presencia de malezas. Clasificación del tipo Random Forest utilizando los algoritmos: b) distribución normal, c) PCA, d) SFS. 57
Fig. 5.3 Matriz de confusión para los algoritmos de clasificación del tipo Maquina de Soporte Vectorial utilizando los descriptores resultantes de: a) distribución normal, b) PCA, c) selección secuencial de características (SFS). 58
Fig. 5.4 a) Zona de un lote con presencia de malezas. Clasificación del tipo Maquina de Soporte Vectorial utilizando los algoritmos: b) distribución normal, c) PCA, d) SFS. 59
Fig. 5.5 Matriz de confusión para los algoritmos de clasificación del tipo K Vecinos más Cercanos utilizando los descriptores resultantes de: a) distribución normal, b) PCA, c) selección secuencial de características (SFS). 60
Fig. 5.6 a) Zona de un lote con presencia de malezas. Clasificación del tipo K Vecinos más Cercanos utilizando los algoritmos: b) distribución normal, c) PCA, d) SFS. 61
Fig. 5.7 Matriz de confusión sobre el 15% del total de los datos con los sistemas de clasificación: a) Random Forest, b) Maquina de Soporte Vectorial, c) K Vecinos más Cercanos. 63
Fig. 5.8 a) Imagen RGB de una zona de un lote, b) etiquetado de todos los objetos. Clasificación: c) Random Forest, d) Maquina de Soporte Vectorial, e) K Vecinos más Cercanos 64

Lista de tablas

Tabla 4.1 Descriptores propuestos a utilizar en el clasificador 42
Tabla 4.2 Definición descriptores a utilizar 42
Tabla 4.3 Descriptores resultantes del análisis de la distribución normal de los datos de cada clase 47
Tabla 4.4 Matriz con los coeficientes del PCA de la base de datos normalizada 48
Tabla 4.5 Abreviatura de cada descriptor en la matriz de coeficientes 48
Tabla 4.6 Varianzas de cada componente principal 49
Tabla 4.7 Suma ponderada de los descriptores a lo largo de las primeras cinco componentes 50
Tabla 5.1 Estructura de la matriz de confusión 54
Tabla 5.2 Desempeño de los algoritmos de clasificación del tipo Random Forest. 56
Tabla 5.3 Numero de objetos clasificados como maleza por los algoritmos de clasificación Random Forest 57
Tabla 5.4 Desempeño de los algoritmos de clasificación del tipo Maquina de Soporte Vectorial 59
Tabla 5.5 Numero de objetos clasificados como maleza por los algoritmos de clasificación Maquina de Soporte Vectorial 60
Tabla 5.6 Desempeño de los algoritmos de clasificación del tipo K Vecinos más Cercanos 61
Tabla 5.7 Numero de objetos clasificados como maleza por los algoritmos de clasificación K Vecinos más Cercanos 62
Tabla 5.8 Desempeño de los algoritmos resultantes de cada sistema de clasificación ... 64
Tabla 5.9 Numero de objetos clasificados como maleza en el ortomosaico 67

1. Generalidades

1.1. Introducción

Colombia es un país que tiene como una de las principales bases de la economía la agricultura, siendo uno de los productos más cosechables el arroz. De acuerdo al boletín técnico de la encuesta nacional de arroz mecanizado, durante el primer semestre del 2019, se sembraron 352.850 hectáreas de arroz, produciendo 1.008.081 toneladas de producto en el país [1]. Además es considerado como un insumo básico por la mayoría de familias Colombianas, y su consumo por hogar se estima en 5,2 libras a la semana.

Las malezas son plantas que crecen de forma silvestre en zonas cultivadas e interfieren con el desarrollo y crecimiento normal de los cultivos, ya que compiten por recursos tales como agua, luz solar, nutriente y fertilizante con las plantas sembradas, además de atraer plagas. Esto hace que el rendimiento disminuya y aumenten los costos de producción, de allí surge la necesidad de hacer un control de malezas en diferentes etapas del cultivo. Existen diferentes métodos para hacer control de malezas, pero el más usado en la actualidad es el control químico, que consiste en rociar herbicidas de manera uniforme sobre el lote cultivado. Debido a su facilidad de aplicación y mayor eficiencia con respecto a otros métodos, es el preferido a utilizar por los agricultores. Este método ha demostrado tener grandes inconvenientes debido al tratamiento homogéneo que se aplica sobre el lote, ya que se ignora los sitios donde no es necesario aplicar herbicida debido a la ausencia de malezas, como también los sitios donde la densidad de estas plantas es muy alta, ocasionando consecuencias negativas como un mayor impacto ambiental usando herbicidas en zonas donde no es necesario, o disminuyendo el rendimiento del cultivo al no aplicar suficiente en zonas donde la densidad de malezas es muy alta.

La agricultura de precisión es un término que surgió hace poco, e involucra la aplicación de tecnología para monitorear y determinar la variabilidad dentro de un cultivo, para posteriormente tomar decisiones de acuerdo a las necesidades de determinadas zonas. Los principales elementos utilizados en la agricultura de precisión son: Sistemas de Posicionamiento Global (GPS), sensores, cámaras, vehículos aéreos no tripulados, imágenes satelitales, entre otros. El manejo de malezas específicas del sitio consiste en la aplicación de la agricultura de precisión para determinar zonas con malezas dentro de los cultivos, y aplicar solo en ellas la cantidad la de herbicida necesario. Esto resulta muy beneficioso para los agricultores, debido a que un menor consumo de herbicidas, y mayor rendimiento del cultivo, implican un mayor ingreso económico.

El departamento del Tolima es el segundo mayor productor de arroz en Colombia [1, p. 3], por lo tanto dicho producto es una de las principales fuente de ingresos económicos de la región. Sin embargo los altos costos de producción, los tratados de libre comercio y el gran número de exportaciones que llegan al país, hacen que el arroz local sea poco competitivo. Es por ello que la Gobernación del Tolima, en conjunto con la Federación Nacional de

Arroceros Fedearroz, los distritos de riego[1] Usocoello, Ususaldaña, Asorrecio y la Universidad de Ibagué invierten recursos para desarrollar herramientas tecnológicas de alerta temprana que detecten problemas agronómicos a tiempo, de modo que se puedan tomar las medidas correctivas en el momento oportuno, con el fin de mejorar la productividad y obtener un mayor beneficio económico para la región. Más específicamente se utilizarán drones, cámaras RGB, térmica y multiespectral para monitorear lotes de cultivos de arroz destinados a objeto de estudio, y analizar condiciones de irrigación y planimetría en la etapa de pre-siembra, detectar zonas con presencia de malezas a inicios de la etapa vegetativa, problemas de estrés hídrico, nutricional y fitosanitario durante el final de la etapa vegetativa y comienzos de la reproductiva. De uno de estos problemas a detectar, es de donde surge este trabajo.

En este trabajo se propone un método para la detección de zonas con presencia de malezas en etapas tempranas en cultivos de arroz, comprendidas desde 15 días antes de la siembra, hasta 30 días de emergencia de la planta. Como objeto de estudio, se realizaron dos ciclos de siembra en distintas épocas del año, en cada ciclo se realizó la siembra en tres lotes distintos pertenecientes a los distritos de riego de Ususaldaña, Usocoello y Asorrecio, ubicados en los municipios de Saldaña, Espinal y Lérida respectivamente, en el departamento del Tolima. Para monitorear los cultivos se utilizan vehículos aéreos no tripulados y dos tipos de cámaras: RGB y multiespectral. Una vez capturada la información, se generan mapas de reflectancia y RGB, a los cuales se les hace un procesamiento de imágenes, posteriormente se definen características o descriptores que permitan discriminar las malezas de las plantas de arroz, para luego desarrollar un sistema de clasificación y finalmente generar un mapa de colorimetría que indique los lugares con presencia de maleza dentro del lote.

[1] Los distritos de riego son instituciones que brindan servicios a los agricultores y distribuyen agua a los cultivos de la región.

1.2. Estado del arte

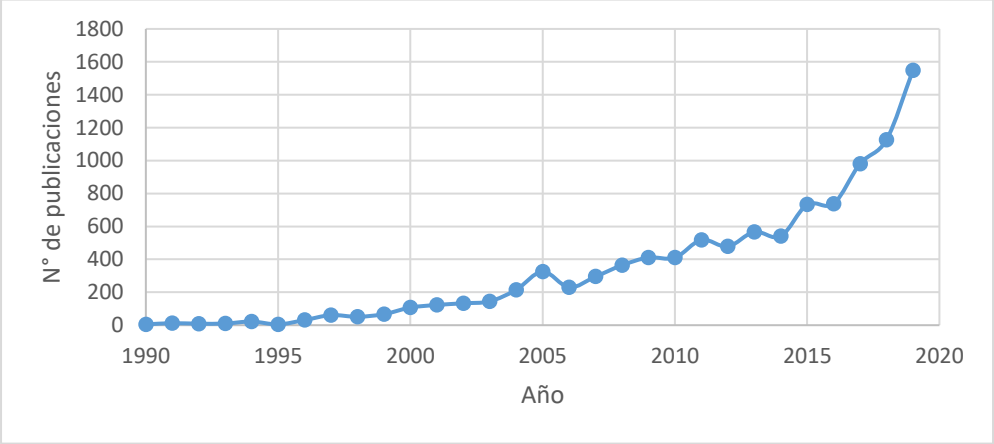
La agricultura es el cultivo de la tierra para la producción de alimentos. Es una actividad practicada por los seres humanos desde hace miles de años, inicialmente era desarrollada por personas o comunidades para producir sus propios alimentos, y con el paso del tiempo se fue volviendo en una actividad comercial. A partir de la edad media se empezó a aplicar innovaciones tecnológicas a la agricultura, como el uso de caballos para arar la tierra, y la implementación de sistemas de riego, esto con el fin de optimizar el rendimiento del cultivo, y disminuir el trabajo humano. En tiempos modernos, se empezó a utilizar fertilizantes, herbicidas y pesticidas, para estimular el crecimiento de la planta, y controlar factores externos que puedan afectar la calidad y el rendimiento del cultivo.

A inicios del siglo XX y mediante estudios de diversas universidades de Estados Unidos se determinó que el rendimiento del cultivo variaba dentro de un mismo lote cultivado, y eso se debe a que el suelo presenta propiedades heterogéneas [2]. A mediados de la década de 1980's, como objeto de estudio se comenzó a emplear el uso de sensores para sensar la materia orgánica del suelo en diferentes puntos dentro de un mismo cultivo [3], y unos diez años después, se incrementó el uso de la tecnología para el monitoreo de cultivos y la aplicación de tratamientos y correctivos en el momento y lugar adecuado. Esto se conoce como agricultura de precisión, y hoy en día involucra diversas tecnologías como: uso de imágenes satelitales, aéreas, sensores espectroscópicos, térmicos, y algoritmos computacionales para detectar la radiación electromagnética reflejada por las plantas y el suelo. Estas herramientas permiten capturar y procesar información de vital importancia para la toma de decisiones sobre el cultivo, lo cual se verá reflejado en un mayor rendimiento y mayor ingreso para el agricultor.

Desde finales del siglo XX hasta hoy en día, ha venido aumentando el número de investigaciones que se realizan en el campo de la agricultura de precisión. Al consultar las palabras "precision agriculture" en la plataforma Scopus[2], se puede observar como en el año 1991 o antes hay menos de diez publicaciones por año sobre este tema. Para el año 2000 por primera vez hay más de cien publicaciones que relacionan estos términos, en el año 2011 sobrepasa las 500 publicaciones por año, y el año 2018 sobrepasa por primera vez más de mil publicaciones en el año (ver figura 1.1). Esto demuestra el creciente número de investigaciones realizadas en el área.

[2] Scopus es una de las plataformas de bases de datos más grandes del mundo. Cubre más de 36.000 títulos, de más de 11.000 editores.

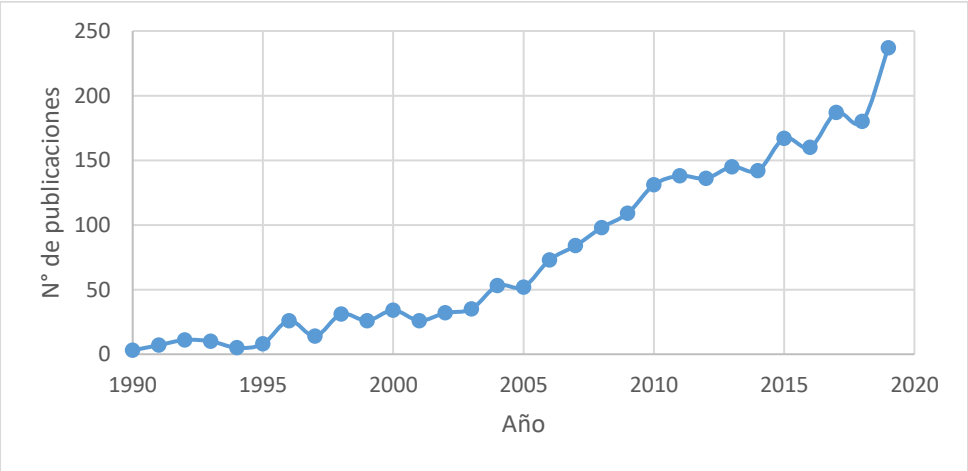
Fig. 1.1 Publicaciones relacionados con las palabra "precision agriculture" a lo largo del tiempo vs número de publicaciones en la plataforma Scopus



Fuente: Autor

Las malezas son plantas que afectan gravemente los cultivos, y son uno de los más grandes problemas que deben de afrontar los agricultores. Es por ello que dentro del campo de la agricultura de precisión, la detección de malezas es un área ampliamente estudiada, y se han publicado diversos artículos a lo largo del tiempo, donde el número de publicaciones ha aumentado notablemente año a año (ver figura 1.2)

Fig. 1.2 Publicaciones relacionadas con las palabras "weed detection" a lo largo del tiempo vs número de publicaciones en la plataforma Scopus



Fuente: Autor

Uno de los artículos más citados en el ámbito de detección de malezas es *"Autonomous robotic weed control systems: A review"* [4]. Allí proponen el diseño de un robot terrestre

no tripulado que utiliza visión por computador para detectar malezas. Para ello utilizan diferentes patrones de reconocimiento como: las características geométricas de las hojas y el tallo, la reflectancia de las plantas, donde principalmente se emplea el uso de índices vegetativos. Adicionalmente se utilizaron características de textura como la matriz de co ocurrencia para detectar dichas plantas. El vehículo cuenta con un sistema de posicionamiento global (GPS) de navegación cinética satelital en tiempo real (RTK), por lo tanto se realiza un mapeo del cultivo, donde cada sitio donde haya presencia de maleza esta georreferenciado correctamente.

Otro tipo de vehículos usados para detectar malezas son los vehículos aéreos, los cuales poseen grandes ventajas y también desventajas a la hora de realizar la detección de malezas en cultivos. López-Granados en su artículo *"Weed detection for site-specific weed management: mapping and real-time approaches"* describe los diferentes métodos existentes de teledetección, como las cámaras usadas para ello. Según el autor, los vehículos aéreos no tripulados (UAV) poseen la ventaja de abarcar áreas más grandes que los vehículos terrestres, pero la resolución espacial de las imágenes es menor, por lo que realizar detección de malezas en etapas tempranas es un poco tedioso. Aunque en general el método más viable para hacer la detección, es el aéreo. Las técnicas de procesamiento de la información pueden ser las mismas para los diferentes tipos de vehículos que puedan utilizarse.

A pesar de haber múltiples investigaciones sobre detección de malezas en cultivos, para el caso de arroz son muy pocos los estudios para cultivos de este tipo específico de planta. Uno de los artículos más recientes sobre detección de malezas en cultivos de arroz es *"Deep localization model for intra-row crop detection in paddy field"* [5]. Allí los investigadores tienen dos lotes de estudio, de una extensión de 2 hectareas en total. Adquieren las imágenes desde un vehículo terrestre, arman una base de datos de más de 10.000 ejemplares y utilizan redes neuronales convolucionales para detectar la maleza presente entre los surcos del arroz.

Al buscar las palabras "detección de malezas" en algunas plataformas académicas son muy pocos los resultados obtenidos, ya que en español no se encuentra mucha información sobre el tema. Uno de los trabajos más notables es un artículo de conferencia llamado "Detección en tiempo real de malezas a través de técnicas de visión artificial" [6], donde detallan las técnicas de análisis para la detección de malezas. Allí destacan las técnicas de análisis de color, análisis de forma, más específicamente de las hojas, y análisis de textura. El artículo plantea la utilización de una cámara digital para la captura de la información, y su posterior procesamiento y análisis para detectar malezas. Al buscar las palabras "detección de malezas en cultivos de arroz" en plataformas académicas en español, los resultados obtenidos fueron nulos, o poco relacionados con el procedimiento de detección considerado como estándar en el campo de investigación, el cual consiste en la captura de las imágenes, procesamiento de la información, implementación de un sistema de clasificación y resultados.

1.3. Objetivo general

Detectar y localizar malezas en cultivos de arroz basados en el análisis de imágenes RGB y de reflectancia tomadas desde drones.

1.3.1. Objetivos específicos

- Generar mapas de reflectancia multiespectral y RGB del lote experimental del cultivo de arroz desde 15 días antes de la siembra hasta 1 mes después de emergencia de las plantas.

- Definir los descriptores a usar en el sistema de clasificación con base en la información de los mapas de reflectancia y RGB.

- Desarrollar un sistema de clasificación que permita discriminar las plantas de arroz de la maleza en etapa temprana del cultivo.

- Desarrollar mapas de color donde se muestre la localización georreferenciada de la maleza dentro del cultivo de arroz.

2. Marco teórico

2.1. Arroz

El arroz es un cereal perteneciente a la familia de las gramíneas y es considerado por muchas regiones y culturas como un alimento básico. Es uno de los cereales más producidos del mundo y que más variedades presentan, habiendo cerca de 10 mil tipos de arroz. Su gran consumo se debe a muchos factores, siendo unos de los principales su gran contenido calórico y de vitaminas. Debido a que no tiene consecuencias digestivas, la mayoría de la población lo puede consumir, su durabilidad permite un almacenamiento prolongado, inclusive por meses, sin que se dañe, además de ser un cereal cultivado por milenios, lo cual ha hecho que se haya vuelto parte de muchas culturas del mundo su consumo.

Colombia es uno de los mayores productores de arroz en Latinoamérica, produciendo más de 2 millones de toneladas de producto en más de 400.000 hectáreas. Existen diferentes métodos de siembra en el país, pero en general se dividen en dos grandes grupos: sistema de arroz irrigado y secano. El arroz irrigado representa al 54% de las hectáreas de arroz en Latinoamérica, y el 81% de la producción de Colombia [7]. Entre los principales métodos de siembra dentro de este sistema, se encuentran: Arroz trasplantado en suelo fangueado o batido, siembra directa con plantas trasplantadas, siembra en suelo inundado y siembra directa sobre suelo seco. El sistema de arroz secano es poco usado en la región, debido a que este método de siembra adquiere el agua para el cultivo enteramente de la lluvia, conllevando a grandes inconvenientes como falta de agua en temporadas de sequía y una mayor cantidad de malezas. Entre todos los métodos de siembra de arroz, se debe tener en cuenta cual es el método más eficiente, para prevenir problemas que puedan afectar el rendimiento del cultivo, de acuerdo a la disponibilidad de recursos del sitio de siembra como agua y características del suelo.

2.2. Malezas

Las malezas son plantas que crecen de forma silvestre en ambientes no deseados. En el ámbito de la agricultura, las malezas disminuyen el rendimiento de los cultivos, ya que compiten por recursos tales como agua, luz, nutrientes con las plantas cultivadas, además de atraer plagas. Esto ocasiona que la cantidad y calidad del cultivo cosechado disminuya drásticamente. Se estima que las malezas ocasionan una pérdida de alrededor 10% de la cosecha mundial de arroz, inclusive habiendo realizado un control de malezas, y a un cultivo al cual no se le hace ningún tipo de control, puede tener pérdidas entre el 44% y 96% de la producción [8]. Las familias de malezas más comunes en cultivos de arroz son: gramíneas, ciperáceas, hoja ancha y el arroz rojo.

Además del uso de herbicidas, existen múltiples métodos para aminorar el crecimiento de malezas durante toda la etapa de cultivo, dependiendo del terreno y los recursos

disponibles, pero en general los métodos más usados son: Introducir en el lote una lámina de agua de al menos 5cm de profundidad durante la mayor parte del ciclo de crecimiento de las plantas, sembrar el arroz mediante trasplante, es decir sembrar las semillas en una cama de siembra, y trasplantarlas ya germinadas al lote, para evitar el crecimiento de semillas correspondientes a malezas [7]. Otras prácticas usadas son: fertilizar las plantas después de haber realizado el control de malezas.

2.3. Teledetección

La teledetección es la adquisición de información de un objeto sin entrar en contacto con este, utilizando elementos de escaneo como sensores y cámaras, los cuales van montados sobre plataformas como aviones, satélites, vehículos aéreos no tripulados. La teledetección es ampliamente usada para obtener información de grandes áreas, un ejemplo de ello es la adquisición de datos meteorológicos de determinadas zonas, evaluación de cultivos, observación de bosques y cuencas hidrográficas, entre otros. Existen diferentes métodos de adquirir información de grandes áreas por medio de teledetección, siendo los más utilizados el radar, los altímetros, el LIDAR, la fotografía estereoscópica, y las plataformas multiespectrales. Estas capturan información en diferentes longitudes de onda del espectro electromagnético, normalmente entre 3 y 15 bandas, permitiendo así capturar información valiosa que se encuentra fuera del espectro visible. Los dispositivos más comunes para capturar imágenes multiespectrales son los satélites, las cámaras multiespectrales e hiperespectrales.

En el ámbito de la agricultura de precisión, la captura de información mediante imágenes multiespectrales es ampliamente usada para obtener características morfológicas, fisiológicas, estructurales y fenológicas de las plantas [9], lo cual es útil para evaluar y monitorear el crecimiento y desarrollo de un cultivo. Las plantas reciben radiación proveniente del sol en diferentes longitudes de onda, tanto en el espectro visible como fuera de este. Una parte de esta energía es absorbida por las plantas en diferentes longitudes de onda y el resto es reflejada, a este proceso se le conoce como reflectancia. Estas longitudes de onda reflejadas pueden ser captadas por un sensor óptico, y puede variar de acuerdo a las características de las plantas. Las bandas que brindan mayor información sobre los cultivos son las bandas infrarrojo cercano y borde rojo. El infrarrojo cercano se sitúa entre las longitudes de onda de 780nm y 2500nm y sirve para detectar la calidad de las hojas y enfermedades en las plantas. El borde rojo tiene longitudes de onda entre 700nm y 800nm y es de gran utilidad, ya que puede detectar enfermedades en edades más tempranas de la planta que las detectadas por la banda infrarrojo cercano, también brinda información más precisa en etapas tardías del cultivo, cuando la densidad de este es mayor, debido a que toma información de toda la planta y no solo de la parte superior o las hojas.

2.4. Índices vegetativos

Un índice vegetativo según M. Gilbert *et al* [10] “es un parámetro calculado a partir de los valores de reflectancia a diferentes longitudes de onda y que pretende extraer de los mismos la información relacionada con la vegetación, minimizando la influencia de perturbaciones como las debidas al suelo y a las condiciones atmosféricas”.

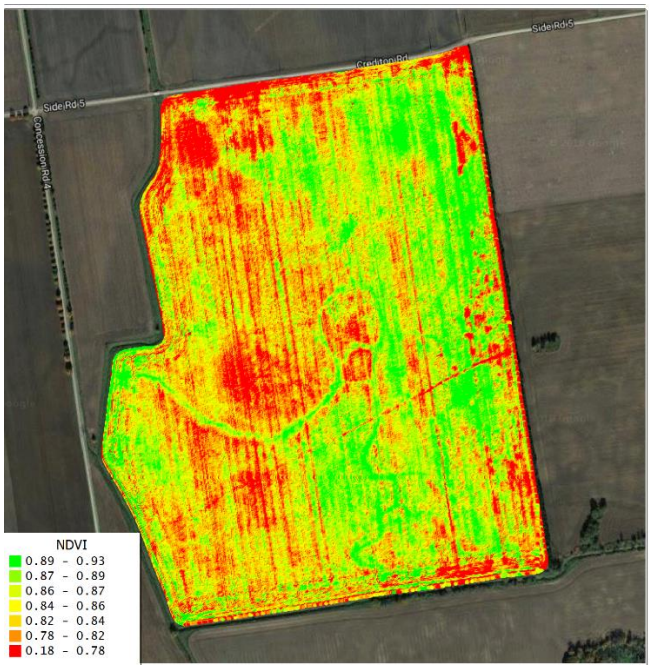
2.4.1. Índice de Vegetación de Diferencia Normalizada (NDVI)

El índice de vegetación de diferencia normalizada (NDVI) utiliza las bandas rojo e infrarrojo cercano para estimar la calidad, cantidad y desarrollo de la vegetación. También es bastante útil en el procesamiento de imágenes pertenecientes a cultivos, ya que realza los pixeles pertenecientes a plantas y atenúa el resto, lo cual permite hacer una segmentación (ver sección 2.5) más acertada de la imagen. El NDVI se calcula con la siguiente ecuación

$$NDVI = \frac{NIR - R}{NIR + R} \tag{2.1}$$

Donde NIR es la banda del infrarrojo cercano y R es la banda del rojo.

Fig. 2.1 Ortomosaico NDVI de un cultivo



Fuente: Pix4D, “Optimizing the ROI of fungicides with NDVI imagery”. Disponible [online] https://www.pix4d.com/blog/pix4dmapper-optimizing-the-ROI-of-fungicides-with-NDVI

2.5. Segmentación

La segmentación de imágenes es un método ampliamente usado para la simplificación del análisis de las imágenes y tiene como función dividir una imagen en varias partes u objetos. Se utiliza principalmente para localizar y encontrar los bordes de un objeto dentro de una imagen. El objetivo de la segmentación es encontrar características visuales similares entre los pixeles, como color, intensidad o textura para hacer una agrupación, o regiones donde la diferencia de las características mencionadas anteriormente sea grande, para hacer una detección de bordes. Usualmente se trabaja la segmentación en imágenes que se encuentran en escala de grises.

Los métodos más ampliamente usados para realizar una segmentación son: agrupamiento, métodos basados en el histograma, detección de bordes, y umbralización. Para el método de agrupamiento o clustering, el algoritmo más usado es K-means, el cual consiste en posicionar sobre la imagen puntos aleatorios que serán el centro de agrupamiento y asignar a cada pixel un valor de cluster, luego se realizan varias iteraciones donde los puntos del centro de los cluster van cambiando, hasta que el valor de la distancia euclidiana entre el centro y el valor del pixel sea el menor posible [11], es decir hasta que este valor llegue a su mínimo y no varíe a través de las iteraciones. En los métodos basados en el histograma, se calcula un histograma de todos los pixeles de una imagen, y de acuerdo a los valores máximos y mínimos de la gráfica, se realiza el agrupamiento de pixeles. La detección de bordes detecta puntos donde las características de un pixel con respecto a otro cambian drásticamente, y traza una línea para dividir la región. La umbralización consiste en fijar unos umbrales (usualmente el valor del pixel), y dependiendo entre los umbrales que se encuentren los pixeles, así mismo se agrupan. Entre los algoritmos más usados en la umbralización, se encuentra la binarización.

2.5.1. Binarización

La binarización es el método más simple de umbralización y consiste en convertir una imagen en escala de grises a una donde solo tenga dos valores de pixel posibles: blanco y negro. Para determinar a qué grupo (blanco o negro) pertenece el pixel, se compara el valor de dicho pixel en escala de grises con un valor umbral, y de acuerdo de si es mayor o menor se clasifica. La binarización se expresa matemáticamente como:

$$g(x,y) = \begin{cases} 0, si\ f(x,y) < T \\ 1, si\ f(x,y) \geq T \end{cases} \tag{2.2}$$

Donde $g(x,y)$ es el valor que tendrá el pixel después de la umbralización, $f(x,y)$ es el valor del pixel en escala de grises, y T es el valor umbral.

El valor umbral puede ser definido de manera manual, pero normalmente se calcula de manera automática. El método más usado para determinar el umbral de manera

automática es el basado en el histograma. Se ubican los máximos locales del histograma, que usualmente es bimodal, y se calcula su valor promedio para determinar su valor umbral. Otra variante un poco más efectiva de este método es establecer como umbral el valor más bajo entre los picos. Estos métodos son propensos a errores, y por ende poco usados. El método más ampliamente usado, y considerado como estándar es el método Otsu. Este método se basa en el análisis del histograma para determinar su valor umbral, basándose en el valor mínimo de varianza dentro de cada objeto o pixeles con valores similares, y valor máximo para objetos de diferentes clases, con el fin de tener una diferencia más marcada y realizar una binarización más precisa. Las varianzas se expresan matemáticamente como:

$$\sigma_w{}^2 = W_b\sigma_b{}^2 + W_f\sigma_f{}^2 \tag{2.3}$$

$$\sigma_B{}^2 = W_b W_f(\mu_b - \mu_f)^2 \tag{2.4}$$

Donde $\sigma_w{}^2$ es la varianza dentro de los mismos objetos, W_b la cantidad en porcentaje de los pixeles correspondientes al fondo, W_f cantidad de pixeles correspondientes al frente, $\sigma_b{}^2$ y $\sigma_f{}^2$ son las varianzas de los pixeles correspondientes al fondo y al frente de la imagen respectivamente. $\sigma_B{}^2$ es la varianza entre las dos diferentes clases, μ_b y μ_f son los promedios ponderados de pixeles correspondientes al fondo y al frente de la imagen respectivamente.

El algoritmo inicialmente toma varios valores umbrales comprendidos entre 0 - 255, y dependiendo de si el pixel tiene un valor menor o mayor a ese umbral lo clasifica como fondo o frente, posteriormente calcula ambas varianzas utilizando la ecuación (2.3) y (2.4). Finalmente escoge como valor umbral definitivo donde el valor de $\sigma_w{}^2$ sea mínimo y $\sigma_B{}^2$ sea máximo.

Fig. 2.2 a) Imagen en escala de grises, b) binarización mediante el método Otsu de la imagen en escala de grises.

a) b)

Fuente: The Lab Book Pages, "Otsu Thresholding". Disponible [online] http://www.labbookpages.co.uk/software/imgProc/otsuThreshold.html

2.6. Análisis de imágenes basado en objetos (OBIA)

El análisis de imágenes basado en objetos consiste en agrupar pixeles que tengan características en común tales como: reflectancia, color, forma, relación con los pixeles adyacentes y asignarles una misma etiqueta a este conjunto. Los dos pasos principales del OBIA es realizar una segmentación y posteriormente una clasificación. Existen diversos métodos de hacer segmentación (ver sección 2.5) y clasificación. Entre los métodos más usados de clasificación se encuentran detección de bordes y etiquetado de componentes conectados. La ventaja de realizar un análisis de imágenes basado en objetos es hacer un procesamiento de imágenes más fácil y rápido, debido a que se analiza un grupo de pixeles como un solo objeto.

2.6.1. Etiquetado de componentes conectados

En el etiquetado de componentes conectados, se transforma na imagen binaria a una donde todos los pixeles correspondientes a objetos que están conectados entre sí, se les asigna una única etiqueta [12]. El etiquetado es necesario en aplicaciones que requieran un reconocimiento de objetos. Existen diversos algoritmos que realizan esta tarea, pero principalmente se utilizan 4 clases: Algoritmos de múltiple escaneo que hacen un barrido de arriba hacia abajo y de izquierda a derecha, en cada barrido asigna etiquetas provisionales a los objetos, y se realizan N barridos hasta que las etiquetas de los objetos no cambien. Los algoritmos de dos escaneos realizan un primer barrido de la imagen asignando etiquetas provisionales a los objetos y guardándolas en un vector de una dimensión, luego un segundo barrido sobre dicho vector reemplaza las etiquetas provisionales por el equivalente más pequeño. Existen también métodos híbridos, que son una combinación de los dos métodos mencionados anteriormente y consiste en realizar primero un escaneo de arriba hacia abajo y de izquierda a derecha, y almacenar las etiquetas en un vector de una dimensión, para en posteriores barridos resolver los problemas de equivalencia en las etiquetas. Resultados experimentales han demostrado que el número máximo de barridos en la imagen es cuatro para que este método sea eficiente.

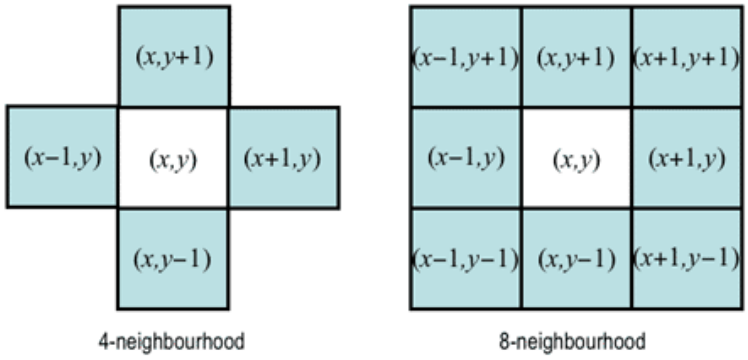
El método más ampliamente usado y considerado casi estándar, es el algoritmo que realiza dos barridos sobre la imagen. Esto se debe a su alta precisión y menor tiempo de ejecución con respecto a los otros métodos que realizan esta tarea. En el algoritmo de dos barridos, se debe establecer si se va a manejar una conectividad de cuatro o de ocho puntos (ver figura 2.2) y se realiza el siguiente procedimiento para hacer el etiquetado:

1. Se realiza un barrido de arriba hacia abajo o de izquierda a derecha sobre la imagen hasta encontrar el primer pixel correspondiente al frente.
2. Se revisan los cuatro u ocho pixeles adyacentes (dependiendo de la conectividad).
3. Si alguno de los pixeles vecinos tienen alguna etiqueta, se le asigna la etiqueta de menor valor al pixel. En caso contrario se asigna una nueva etiqueta.

4. Se guarda el valor de etiqueta del objeto.
5. Se repiten los pasos anteriores hasta realizar un barrido sobre toda la imagen.

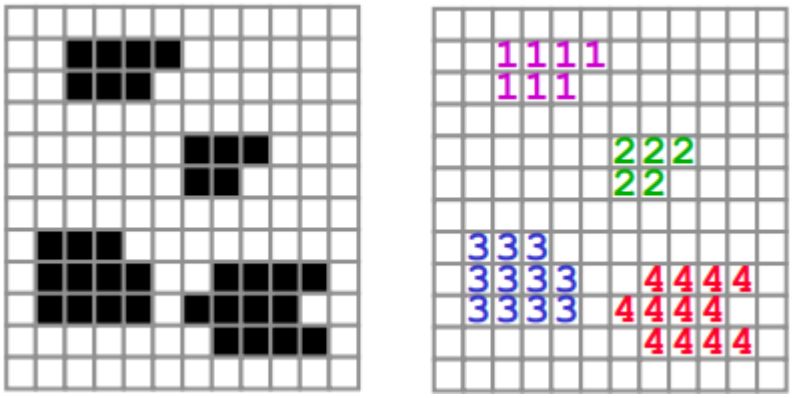
En el segundo barrido se resigna a los pixeles correspondientes al frente la menor etiqueta posible, de acuerdo a los pixeles que se encuentren conectados entre sí. El objetivo de este nuevo barrido, es corregir posibles etiquetas mal asignadas durante el primero.

Fig. 2.3 Conectividad de cuatro y ocho pixeles



Fuente: ResearchGate, "Shape based optimal edge detection paper review". Disponible [online] https://www.researchgate.net/publication/278000510_Shape_based_optimal_edge_detection_paper_review

Fig. 2.4 a) Imagen binarizada, b) etiquetado de la imagen binarizada



a) *b)*

Fuente: Universidad de Zaragoza, "Lección 2: Conectividad". Disponible [online] http://webdiis.unizar.es/~neira/12082/12082conectividad.pdf

2.7. Análisis de objetos basado en características de forma

Las características de forma son primordiales a la hora de realizar una clasificación de objetos después de haber realizado una segmentación y etiquetado, ya que permiten discriminar objetos entre sí en base a sus características visuales como forma, área, etc. Existen dos grandes métodos para extraer características por forma y son el análisis

basado en la forma completa del objeto, y los basados exclusivamente en el borde. Para realizar una correcta discriminación entre objetos basados en sus características visuales o forma, se recomienda cumplir con ciertos criterios como que las características a calcular sean invariantes a la rotación, traslación o escalado de las imágenes. Esto quiere decir que si la imagen cambia de resolución, orientación o posición, las características del objeto no deben cambiar. También se debe de procurar que las características sean inmunes a la distorsión y posible ruido presente en la imagen [13].

2.7.1. Momentos invariantes de Hu

Los momentos invariantes en el reconocimiento de objetos y patrones fueron introducidos por primera vez por Hu en 1962 [14], y describen una serie de características que son invariantes a la posición, traslación, rotación y tamaño del objeto. Son utilizados ampliamente en el reconocimiento de patrones, registro y reconstrucciones de imágenes. También es de gran utilidad a la hora de extraer características de objetos, para hacer una clasificación. Una de las pequeñas desventajas que presentan estos momentos, es que pueden variar de acuerdo a la presencia de ruido proveniente de la captura de las imágenes, o su procesamiento.
Según Hu [14], para una función continua de una imagen de dos dimensiones $f(x,y)$, el momento de orden $p+q$ esta dado por:

$$M_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)\, dxdy \tag{2.5}$$

Donde M_{pq} es el momento, x, y las posiciones, p el momento en el eje x, q el momento en el eje y, $f(x,y)$ la función de la imagen. p y q toman valores discretos (0, 1, 2, 3, …). Estos momentos pueden ser susceptibles a traslaciones, rotaciones y escalado [15], por lo tanto las invariantes pueden obtener utilizando los momentos centrales para evitar el problema de traslación:

$$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\bar{x})^p (y-\bar{y})^q f(x,y)\, dxdy \tag{2.6}$$

Donde $\bar{x}, \bar{y}$ son los centroides de la imagen $f(x,y)$, y se calculan a partir de las siguientes dos ecuaciones:

$$\bar{x} = \frac{M_{10}}{M_{00}} \tag{2.7}$$

$$\bar{y} = \frac{M_{01}}{M_{00}} \tag{2.8}$$

Se puede obtener una invariancia en el escalado de la imagen, al dividir los momentos centrales (ecuación 2.6), en un momento central de orden cero escalado correctamente:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\left(1+\frac{p+q}{2}\right)}} \tag{2.9}$$

Donde η_{pq} son los momentos centrales normalizados invariantes al escalado, μ_{pq} son los momentos centrales, μ_{00} es el momento central de orden cero, p, q son el orden de los momentos en los ejes x, y respectivamente.

A partir de la invariancia al escalado y la traslación, Hu [14] introdujo siete momentos invariantes con momentos inferiores o iguales a orden tres, que también son invariantes a la rotación:

$$\phi_1 = \eta_{20} + \eta_{02} \tag{2.10}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \tag{2.11}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \tag{2.12}$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \tag{2.13}$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \tag{2.14}$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \tag{2.15}$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \tag{2.16}$$

2.8. Análisis de objetos basados en características de textura

La textura es la variación de los niveles de intensidades de un pixel, en una pequeña porción de la imagen, o vecindario. Hay varios métodos para extraer características de textura de una imagen, de acuerdo a [16] se pueden distinguir cuatro categorías: métodos estadísticos, métodos estructurales, métodos basados en modelos, y métodos basados en transformadas.

El análisis estadístico de textura describe la textura de regiones en una imagen a través de momentos de orden mayor de sus histogramas[3] en escala de grises. El método más usado dentro de esta categoría, se basa en la extracción de características de textura de una matriz de co-ocurrencia (ver sección 2.8.1). El análisis estructural de textura describe una textura como la composición de elementos de textura bien definidos. Varios métodos

[3] Histograma es la representación gráfica de distribuciones de frecuencias de datos en una función.

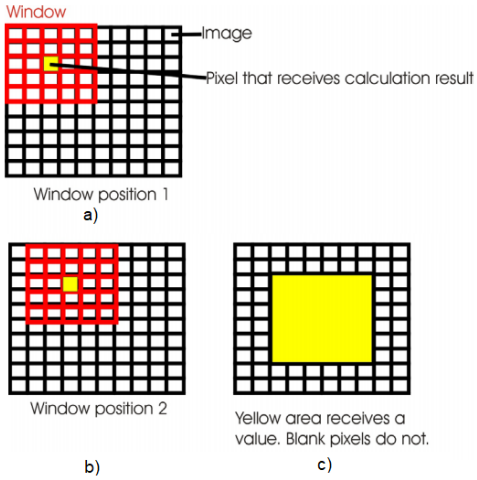
de análisis estructural han sido propuestos, que van desde usar diferentes formas de elementos estructurantes, hasta concebir estructuras reales como versiones distorsionadas de texturas ideales. Este método es poco usado en general, debido a que solo puede describir texturas regulares. El análisis de textura basado en modelos genera un modelo de cada pixel en la imagen, basado en el promedio de las intensidades de los pixeles vecinos. Por último, el análisis de textura basado en transformadas convierten la imagen en una nueva figura usando las propiedades espaciales de frecuencia de las variaciones de intensidad de los pixeles. Los tipos de transformadas más utilizadas son: la Transformada Rápida de Fourier (FFT), Análisis de Componentes Principales (PCA), la Transformada de Wavelet y el Filtro de Gabor.

2.8.1. Matriz de co-ocurrencia

La matriz de co-ocurrencia en escala de grises o GLCM (del inglés Gray Level Co-ocurrence Matrix) es un método estadístico de segundo orden para calcular la textura de una imagen. Esta matriz describe la frecuencia de un nivel de gris que aparece en una relación espacial específica con otro valor de gris dentro del área de una ventana determinada a una distancia δ y un angulo θ el uno del otro [16].

El tamaño de la ventana debe de ser cuadrada y con un número impar de pixeles [17]. El resultado del cálculo es un número que representa la ventana completa, y es situado en el pixel central de la ventana, luego la ventana se mueve un pixel, y se realiza nuevamente el cálculo, y así sucesivamente hasta que se realiza un barrido sobre toda la imagen

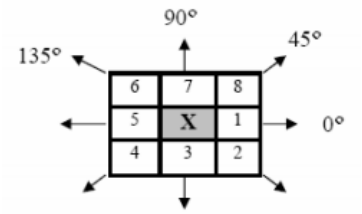
Fig. 2.5 a) Ventana de 5x5 pixeles y en el centro el pixel que recibe el valor del cálculo, b) posición nueva de la ventana, c) pixeles de la imagen que reciben el valor del cálculo



Fuente: Calgary University, "GLCM Texture: A Tutorial" Disponible [online] https://prism.ucalgary.ca/bitstream/handle/1880/51900/texture%20tutorial%20v%203_0%20180206.pdf?sequence=11&isAllowed=y

La relación espacial entre el pixel de referencia y su vecino puede ser en cualquiera de las 8 direcciones de los pixeles que lo rodean, aunque normalmente se suelen tomar solo 4 direcciones ya que las otras cuatro son direcciones opuestas, y se pueden determinar más fácilmente mediante una matriz simétrica. Se puede obtener una invariancia en la relación de los pixeles, cuando se calcula la matriz de co-ocurrencia en las direcciones $\theta = 0°,\ 45°,\ 90°\ y\ 135°$ y se promedian. Las direcciones suelen expresarse también en forma de coordenadas, siendo para el caso de $\theta = 0°$ (1, 0). Esto quiere decir que se toma un pixel en la dirección del eje X positivo, y cero pixeles en la dirección del eje Y, siendo el pixel de referencia el origen del plano cartesiano. Para el caso de $\theta = 45°$, la dirección seria (1, 1), para $\theta = 90°$ seria (0, -1), y para $\theta = 135°$ seria (1, -1).

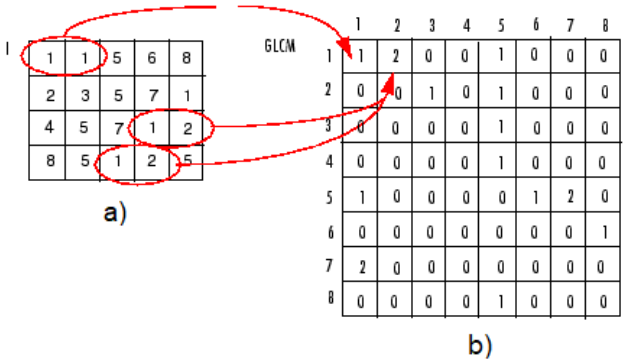
Fig. 2.6 Los 8 vecinos del pixel de referencia X, de acuerdo al ángulo usado para el cálculo de la matriz de co-ocurrencia, para una distancia de 1



Fuente: Presutti, M. "La matriz de co-ocurrencia en la clasificación multiespectral". 4ª Jornada de Educação em Sensoriamento Remoto no Âmbito do Mercosul Disponible [online] http://www3.inpe.br/unidades/cep/atividadescep/jornada/programa/t-9_trab_27.pdf

En la figura 2.7 a se tiene la representación de una imagen en una ventana de 4x4 pixeles, con 8 diferentes niveles de grises. Allí se calcula la matriz de co ocurrencia en la dirección (1, 0), a una distancia $\delta = 1\ pixel$. Se puede observar cómo está presente en la imagen la combinación de valores de grises 1 y 2 dos veces, por lo tanto en la posición (2, 1) en la matriz de co-ocurrencia (figura 2.7 b) se sitúa el número 2, y así se realiza el proceso con todos los pixeles presentes en la imagen, a excepción de la fila situada al extremo derecho, ya que estos pixeles no tienen vecinos en la posición (1, 0).

Fig. 2.7 a) Imagen en escala de grises, y b) su matriz de co-ocurrencia



Fuente: MathWorks, "Graycomatrix". Disponible [online] https://la.mathworks.com/help/images/ref/graycomatrix.html

La matriz de co-ocurrencia es cuadrada, y tiene un número de filas y columnas igual a la cantidad de niveles de grises en la imagen. Normalmente las imágenes manejan datos de 8 bits, por lo tanto la matriz de co-ocurrencia en ese caso sería de 256x256 (2^8 = 256). En el caso de la figura 2.7 a, los datos del pixel son de 3 bits, por lo tanto la matriz de co-ocurrencia es de 8x8 (2^3 = 8).

Normalmente para hacer análisis de textura mediante la matriz de co-ocurrencia, es necesario que esta sea simétrica, es decir que la cantidad de veces que aparece la combinación (1,2) sea igual al número de veces que aparece la combinación (2,1). El método más sencillo para lograr la simetría es sumándole la matriz traspuesta a la matriz original. Adicionalmente se requiere normalizar la matriz de co ocurrencia, por lo tanto se divide el número de observaciones de cada clase, en el número de observaciones total. Así se obtiene el porcentaje de ocurrencia de cada una de las clases.

2.8.2. Descriptores de Haralick

Haralick et al [18] propusieron un análisis cuantitativo de una matriz de co-ocurrencia (GLCM) normalizada a través de 14 descriptores de textura a través de la probabilidad $p(i,j)$ de los valores de intensidad de dos pixeles (i,j):

1. Segundo momento angular:

$$f_1 = \sum_i \sum_j p(i,j)^2 \tag{2.17}$$

2. Contraste:

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\}, |i-j| = n \tag{2.18}$$

Donde N_g es el nivel de gris a evaluar.

3. Correlación:

$$f_3 = \frac{\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{2.19}$$

Donde μ_x, μ_y es la media, y σ_x, σ_y la desviación estándar de p_x y p_y.

4. Suma de cuadrados: Varianza

$$f_4 = \sum_i \sum_j (i-\mu)^2 p(i,j) \tag{2.20}$$

5. Momento de diferencia inversa: Homogeneidad

$$f_5 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j) \tag{2.21}$$

6. Suma promedio:

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i) \tag{2.22}$$

Donde x, y son las coordenadas de una entrada en la matriz de co-ocurrencia, y $p_{x+y}(i)$ es la probabilidad de la matriz de co-ocurrencia en $x+y$.

7. Suma de varianzas:

$$f_7 = \sum_{i=2}^{2N_g} \left(i-\mu_{x+y}\right)^2 p_{x+y}(i) \tag{2.23}$$

Donde μ_{x+y} es la suma promedio

8. Suma de entropías:

$$f_8 = -\sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\} \tag{2.24}$$

9. Entropía:

$$f_9 = -\sum_i \sum_j p(i,j) \log\big(p(i,j)\big) \tag{2.25}$$

10. Diferencia de varianzas:

$$f_{10} = \sum_{i=0}^{N_g-1} \left(i-\mu_{x-y}\right)^2 p_{x-y}(i) \tag{2.26}$$

11. Diferencia de entropías:

$$f_{11} = -\sum_{i=0}^{N_g-1} p_{x-y}(i)\log\{p_{x-y}(i)\} \tag{2.27}$$

12. Información de medición de correlación 1:

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}} \tag{2.28}$$

Donde:

$$HXY = -\sum_i \sum_j p(i,j)\log\big(p(i,j)\big) \tag{2.29}$$

$$HXY1 = -\sum_i \sum_j p(i,j)\log\{p_x(i)p_y(j)\} \tag{2.30}$$

13. Información de medición de correlación 2:

$$f_{13} = \sqrt{1 - e^{-2(HXY2-HXY)}} \tag{2.31}$$

Donde:

$$HXY2 = -\sum_i \sum_j p_x(i)p_y(j)\log\{p_x(i)p_y(j)\} \tag{2.32}$$

14. Máximo coeficiente de correlación:

$$f_{14} = \sqrt{segundo\ eigenvalue^{4}\ mas\ largo\ de\ Q} \tag{2.33}$$

Donde:

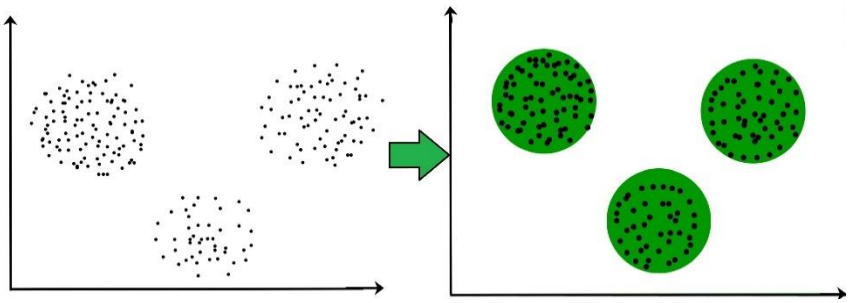
$$Q(i,j) = \sum_k \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)} \tag{2.34}$$

2.9. Sistema de clasificación

[4] Eigenvalue o valor propio es el factor por el cual un vector característico o eigenvector es escalado al aplicar una transformación lineal

En aprendizaje automático, un sistema de clasificación se encarga de agrupar datos de entrada en diferentes categorías. Existen dos tipos de sistemas de clasificación: los supervisados y no supervisados. Los sistemas de clasificación no supervisados se caracterizan por no tener información preexistente y simplemente clasifican la información de acuerdo a características en común que presentan los datos. Este tipo de aprendizaje se utiliza principalmente para encontrar diferentes grupos dentro de los elementos de los datos, esto suele conocerse como "clustering" (ver figura 2.8). Existen diferentes algoritmos de agrupamiento, siendo las más comunes los métodos de: K-means, modelamiento de agrupamiento gaussiano, entre otros.

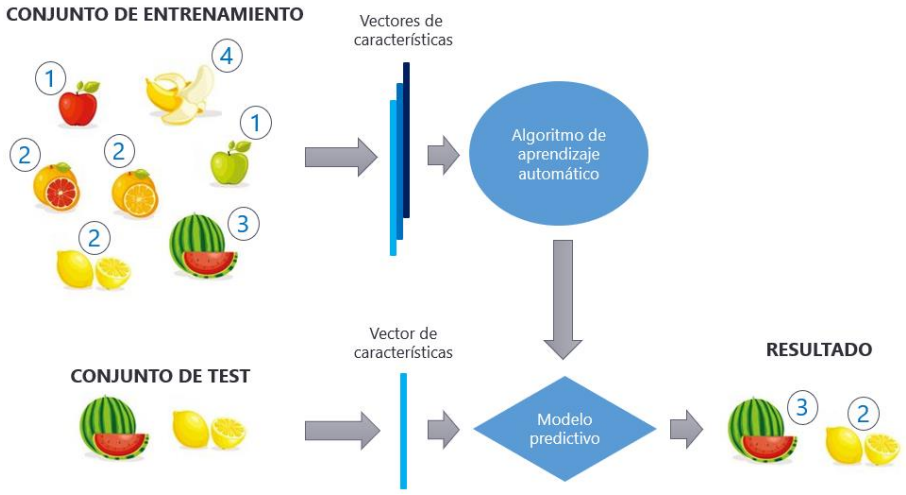
Fig. 2.8 Clustering de datos de acuerdo a su ubicación espacial



Fuente: GeekforGeeks, "Clustering in Machine Learning". Disponible [online] https://www.geeksforgeeks.org/clustering-in-machine-learning/

Los sistemas de clasificación supervisados realizan deducciones de funciones, basados en el comportamiento de datos existentes o datos de entrenamiento, y que han sido etiquetados y caracterizados previamente [19]. Una etiqueta es la salida que han mostrado los datos de entrenamientos. Los datos a clasificar se le calculan dichas características (también conocidas como descriptores) y se le asignara una etiqueta en base a la similitud de estas características con la de los objetos de entrenamiento. La predicción obtenida se representa por medio de una función donde las entradas representan las características analizadas, y la salida la variable que se quiere predecir. En sistemas de clasificación esta salida es de tipo categórico.

Fig. 2.9 Representación gráfica de un sistema de clasificación supervisado



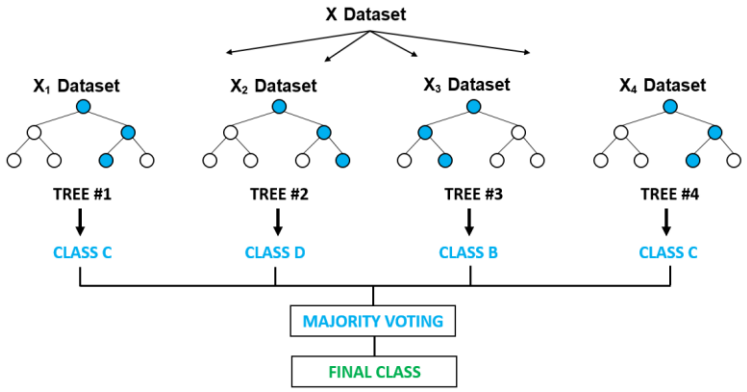
Fuente: Diego Calvo, "Aprendizaje supervisado". Disponible [online] http://www.diegocalvo.es/aprendizaje-supervisado/

En el ámbito de procesamiento de imágenes, las características más utilizadas para discriminar un objeto presente en una imagen con respecto a otros, son los descriptores basados en forma y en textura. Los sistemas de clasificación supervisada más ampliamente usados son: redes neuronales artificiales (ANN), máquina de soporte vectorial (SVM), Bosques Aleatorios o Random Forest, Clasificador Bayesiano, AdaBoost, K Vecinos más Cercanos (KNN), entre otros.

2.9.1. Random Forest

Random Forest o bosques aleatorios es un método de regresión y clasificación supervisada, que consiste en la agrupación de una gran cantidad de árboles de decisión. Los arboles de decisión son un modelo de predicción basado en diagramas de árbol, donde en cada nodo es el punto donde se ha de tomar una decisión de entre varias posibles, y los últimos nodos donde se etiqueta la información [20]. Este modelo ha mostrado ser muy sensible a la varianza de los datos, es decir la diferencia entre muestras de una misma clase en la etapa de entrenamiento. Random Forest divide el número de datos de entrenamiento de una manera aleatoria, para realizar varias predicciones con diferentes arboles de decisión, y finalmente etiquetar el dato a la clase con mayor número de predicciones por los arboles de decisión [21] (ver figura 2.10), esto también se conoce como *Bagging*. Este tipo de clasificación es bastante útil debido a su alta eficiencia, a pesar de que los datos de entrada presenten una alta varianza entre ellos, adicionalmente reduce el *overfitting* o el sobreajuste, ya que cada uno de los árboles de decisión posee solo una porción de los datos de entrenamiento.

Fig. 2.10 Metodología de clasificación de un sistema Random Forest.

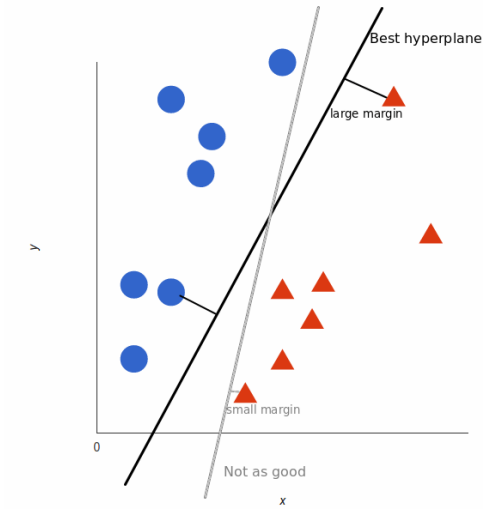


Fuente: Modificado de Global Software Support, "Random Forest Classifier – Machine Learning". Disponible [online] https://www.globalsoftwaresupport.com/random-forest-classifier/

2.9.2. Máquina de Soporte Vectorial

Las máquinas de soporte vectorial es un algoritmo de clasificación binaria supervisada y regresión. Este algoritmo representa los datos de entrenamiento en el espacio, y construye un hiperplano definido como el vector entre los dos puntos más cercanos entre una clase y la otra, conocido normalmente como vector de soporte. Los nuevos datos a clasificar se sitúan en el espacio, y de acuerdo a su ubicación, se clasifican como de una clase u otra.

Fig. 2.11 Selección del mejor hiperplano en una máquina de soporte vectorial lineal



Fuente: MonkeyLearn, "An introduction to Support Vector Machines". Disponible [online] https ://monkeylearn.com/blog/introduction-to-support-vector-machines-svm/

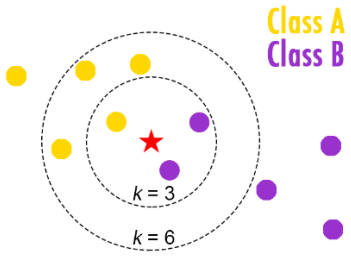
Cuando los datos de entrenamiento no son lineales, se debe de añadir una dimensión adicional, donde se calcula el hiperplano, para posteriormente ser representado en el espacio con las dimensiones originales. Lo anteriormente descrito puede resultar computacionalmente costoso al envolver un gran número de dimensiones, muchas de estas con cálculos bastante complejos. Para resolver este inconveniente se utilizan kernels, que es el producto escalar[5] entre los vectores que representan los hiperplanos *[22]*. Esto es de gran beneficio ya que los cálculos se pueden realizar sin necesidad de conocer las nuevas dimensiones, lo cual reduce el costo computacional significativamente.

2.9.3. K-Nearest Neighbors (Vecinos más cercanos)

K vecinos más cercanos o KNN (Nearest Neighbors) es uno de los algoritmos de clasificación supervisada y regresión más sencillos. Se basa en la estimación de la probabilidad de que un elemento pertenezca a una clase u otra de acuerdo a la proximidad espacial con los datos de entrenamiento [23]. Para ello se calcula la distancia entre el dato a clasificar y los datos de entrenamiento, se establece el valor de K, que es la cantidad de vecinos cercanos a evaluar con respecto al objeto a etiquetar. Se etiqueta el dato nuevo a la clase con mayor frecuencia entre los vecinos cercanos K.

La elección del valor correcto de K es fundamental, y es el éxito en este método de clasificación. En la figura 2.12 se puede observar como al variar el valor de K, puede variar la etiqueta asignada al dato. En caso de K=3, el dato seria etiquetado a la Clase B, en cambio sí se escoge un valor de K=6, el dato seria asignado a la Clase A. La mejor elección de K depende enteramente de los datos de entrenamiento. Valores pequeños de K hacen al clasificador poco estable y susceptible a errores. En cambio valores altos de K disminuyen el ruido en la clasificación, pero valores demasiados altos pueden generar errores de clasificación, y hacer el costo computacional bastante elevado. Para escoger el valor correcto de K, se aconseja variar este parámetro, hasta que el número de errores se reduzca, y la exactitud del clasificador sea lo suficientemente buena.

Fig. 2.12 Dato a etiquetar y sus vecinos más cercanos



Fuente: Modificado de Data Science and Analytics, "Classification Series 5 – K-Nearest Neighbors (knn)". Disponible [online] https://dslytics.wordpress.com/2017/11/16/classification-series-5-k-nearest-neighbors-knn/

[5] Producto escalar es la proyección de un vector con respecto a otro y viceversa.

3. Materiales y métodos

3.1. Materiales

A continuación se describen los materiales utilizados para la captura de las imágenes.

3.1.1. Dron

Un dron o vehículo aéreo no tripulado es una aeronave manejada a distancia, capaz de mantener vuelo y que utiliza motores para propulsarse. En el ámbito de la agricultura de precisión, se suelen utilizar los drones para capturar información de los cultivos, es por ello que llevan consigo cámaras RGB o multiespectrales. Los dos tipos de drones para realizar estas tareas son los drones multirotor y de ala fija. Los drones multirotor constan de cuatro, seis u ocho motores para controlar el vuelo, despego y aterrizaje de la aeronave. Son los más usados en el mercado debido a sus grandes ventajas como: gran maniobrabilidad, facilidad de uso, mayor capacidad de carga y tamaño compacto. Los drones de ala fija se caracterizan por tener la forma de un avión, y son ampliamente usados específicamente en el ámbito de la agricultura de precisión, debido a que son capaces de cubrir una mayor área que los drones multirotor, además de tener mayor estabilidad debido a su diseño.

Para realizar este trabajo, se utilizó un dron multirotor DJI Matrice 200, ya que es un proyecto de investigación que se está llevando en lotes de corta extensión. Este dron cuenta con cuatros motores, y tiene unas dimensiones de 89x88x38cm una vez armado. Posee un rango de operación de 8 kilómetros, una autonomía de vuelo de 38 minutos, protección IP43 contra el polvo y agua. Además cuenta con el sistema DJI Airsense, el cual detecta aeronaves cercanas en tiempo real, y envía una alerta al piloto, además de evitar obstáculos automáticamente [24]. Posee una precisión de GPS de 0.5 metros en el eje vertical, y 1.5 metros en los ejes horizontales. El dron también cuenta con la opción de realizar vuelos automáticamente, esto se puede lograr programando una ruta de vuelo desde cualquier aplicación de vuelo de drones.

Fig. 3.1 Dron DJI Matrice 200

Fuente: DJI, "Matrice 200 Series". Disponible [online] https://www.dji.com/matrice-200-series

3.1.2. Cámara RGB

Una cámara RGB es cualquier cámara digital que utiliza un sensor de imagen que convierte la luz en señales eléctricas, que posteriormente es digitalizada y almacenada en una memoria. La resolución de la imagen viene determinada por el sensor, el cual se compone de millones de pixeles que se cargan en respuesta a determinadas longitudes de onda, más específicamente rojo, verde y azul (red, green, blue en inglés, de allí las siglas RGB). Mediante la aplicación de algoritmos internos en el dispositivo de la cámara, se combinan para representar colores completos. En el ámbito de la agricultura de precisión, las cámaras RGB son utilizadas para detectar diferentes características del lote y cultivo

La cámara utilizada para la captura de imágenes RGB en este proyecto es la Zenmuse X4S. Esta cámara de la empresa DJI está diseñada para tomar imágenes y videos de alta resolución desde drones. Tiene un lente 8.8mm/F2.8-11 de baja distorsión y un sensor de CMOS de 1 pulgada que captura imágenes a una resolución de 20 mega pixeles, con un ISO máximo de 12.800. Posee una apertura focal de 84°, lo cual permite abarcar una gran cantidad de área en una sola imagen, además integra tecnología para el procesamiento de imágenes propia de DJI, que permite grabar videos a 4K/60 H.264 and 4K/30 H.265 con una tasa de bits de 100 megabits por segundo. En modo ráfaga puede capturar 14 imágenes por segundo en una resolución de 20 mega pixeles [25].

Fig. 3.2 Cámara Zenmuse X4S

Fuente: DJI, "Zenmuse X4S". Disponible [online] https://www.dji.com/zenmuse-x4s

3.1.3. Cámara multiespectral

Una cámara multiespectral es capaz de captar diferentes rangos de longitudes de onda, normalmente entre 5 y 15, incluyendo algunas que se sitúan fuera del espectro visible. Las bandas más comunes presentes en tipo de cámaras son las situadas en el infrarrojo como por ejemplo: el infrarrojo cercano, borde rojo y el infrarrojo térmico. Las imágenes multiespectrales permiten obtener información que con el ojo humano no es posible de detectar. En el ámbito de la agricultura de precisión, este tipo de cámaras son ampliamente usadas, debido a que las plantas y el suelo reflejan la energía en las longitudes de onda fuera del espectro visible de manera diferente cuando el estado de los cuerpos cambia,

por lo tanto es posible determinar enfermedades en las plantas, problemas de estrés hídrico y nutricional, fertilidad, problemas de irrigación del suelo, entre otros.

En este proyecto se utilizó la cámara Altum, de la marca Micasense para la fotografía multiespectral. Esta cámara captura información en seis bandas espectrales diferentes: azul, verde, rojo, infrarrojo cercano, borde rojo e infrarrojo térmico. Tiene una resolución de sensor de 3,2 mega pixeles, para todas las bandas a excepción de la térmica. Posee una velocidad de obturación de 1 segundo, una longitud focal de 8mm para todas las bandas multiespectrales, y 1.7mm para la banda térmica [26]. Además integra un sensor de luz descendente (DLS) de segunda generación y un panel de calibración que permite realizar una muy buena corrección radiométrica cuando no se tienen las condiciones de luz adecuadas.

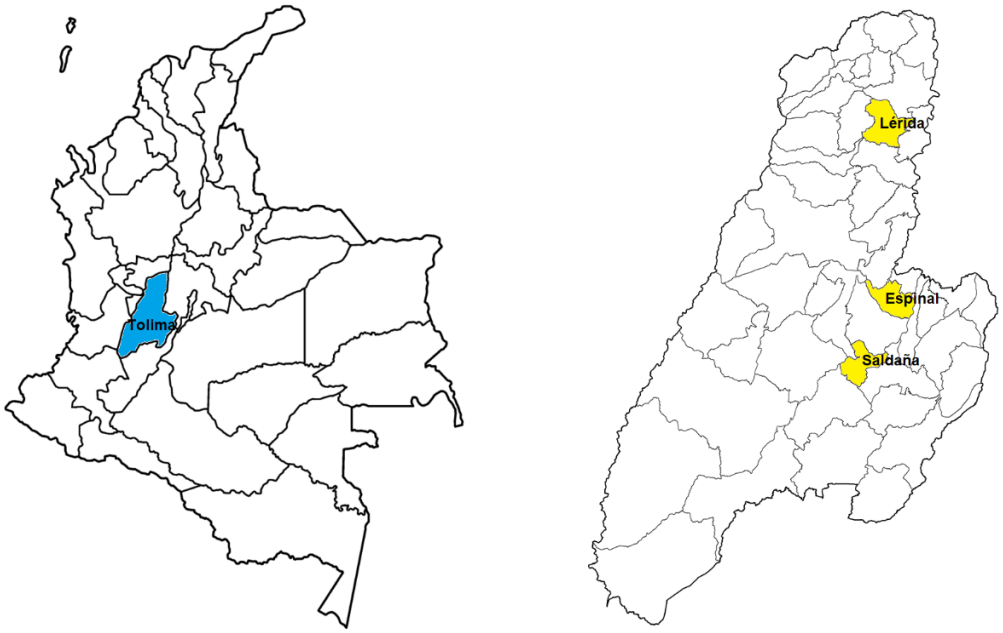
Fig. 3.3 Cámara multiespectral + DLS

Fuente: Micasense, "Altum". Disponible [online] https://www.micasense.com/es/altum

3.2. Sitio de estudio

Para este proyecto, se realizaron dos ciclos de siembra en distintas épocas del año, con el fin de evaluar el desarrollo del cultivo en diferentes condiciones climáticas, además de corregir en el segundo ciclo de siembra, procedimientos que se hubieran hecho mal o de manera inadecuada en el primero. En cada ciclo, los distritos de riego Usocoello, Usosaldaña y Asorrecio destinaron de a un lote en los municipios de Espinal, Saldaña y Lérida respectivamente, todos ubicados en el departamento del Tolima. Las coordenadas donde se encuentran ubicados los lotes del primer ciclo son: 4°11'54.2"N 74°51'32.6"W, 3°55'53.4"N 74°58'44.5"W, 4°49'45.7"N 74°55'10.8"W. Durante el segundo ciclo se destinaron tres lotes diferentes dentro de los mismos municipios, los cuales tienen las siguientes coordenadas: 4°09'50.1"N 74°53'14.4"W, 3°55'50.5"N 75°00'35.1"W, 4°49'44.8"N 74°55'09.2"W. En total se destinaron seis lotes de estudio, de 1 hectárea cada uno. Dentro de estos lotes, se destinó una parcela de aproximadamente ¼ del tamaño total del lote para el estudio de malezas. Adicionalmente la parcela se dividió en secciones, donde se aplicó diferentes cantidades de herbicida pre-emergente, para inducir las malezas de una manera heterogénea y tener variedad en los datos a estudiar.

Fig. 3.4 a) Mapa de Colombia y el departamento de Tolima en azul, b) Mapa del Tolima y los municipios de Lérida, Espinal y Saldaña en amarillo



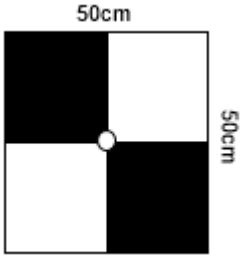
a) *b)*

Fuente: Autor

3.3. Localizacion puntos de control terrestre sobre los lotes

Antes de realizar los vuelos sobre los lotes, se deben de ubicar puntos de control terrestre para la georreferenciación. Estos deben de ser visibles desde la altura a la que se capturan las imágenes. Para ello se fabricaron seis tableros de 50cmx50cm en madera MDF[6], los cuales se situaron al final de un tubo de 1.5m de largo, para evitar que en etapas finales del cultivo, este cubra los marcadores.

Fig. 3.5 Tablero de punto de control

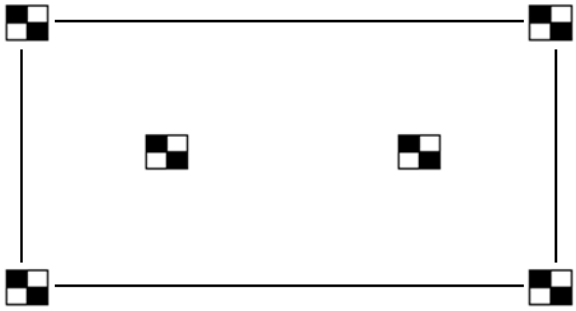


Fuente: Autor

[6] MDF o madera de fibra media, es un tipo de madera fabricado a partir de fibras de madera comprimidas.

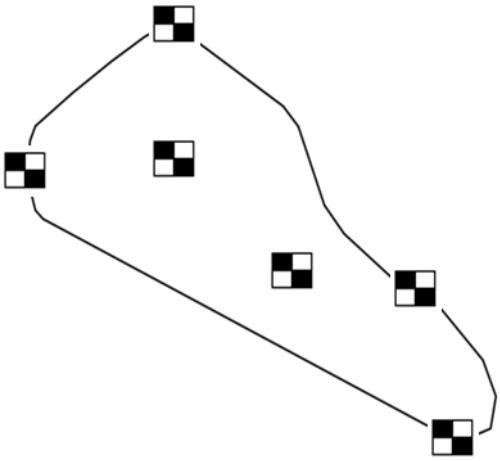
Los marcadores se situan en las esquinas de los lotes y dos en el centro si el terreno tiene forma de poligono rectangular. En caso de que el poligono del terreno sea irregular, se debe de garantizar que cuatro marcadores queden en el borde, y los otros dos en el centro

Fig. 3.6 Ubicación de los marcadores en un polígono rectangular



Fuente: Autor

Fig. 3.7 Ubicación de los marcadores en un polígono irregular



Fuente: Autor

La ubicación de estos puntos de control terrestres sobre los lotes a volar es necesario para obtener una correcta union entre las diferentes imágenes tomadas por el dron, como tambien para obtener la georreferenciacion del ortomosaico. La georreferenciacion se puede obtener tomando datos con un GPS RTK sobre dichos puntos de control, o ubicando los tableros en las imágenes capturadas, las cuales tienen los datos del GPS propio del dron.

3.4. Adquisición de las imágenes

Para determinar la mayor altura de vuelo a la cual las zonas donde hay presencia de malezas son reconocibles, se tomaron imágenes de prueba a diferentes alturas en etapas tempranas del cultivo, y se aplicaron algoritmos de procesamiento hasta la etapa de segmentación (ver sección 2.5), con el fin de observar en qué punto las plantas pequeñas empezaban a ser irreconocibles en las imágenes binarizadas. Se determinó que la mayor altura a la que eran detectables las zonas con plantas más pequeñas a la hora de procesar las imágenes es a una altura de vuelo de 40 metros. Cabe anotar que es necesario evaluar el estado fenológico de las plantas antes de realizar el vuelo, ya que si son muy pequeñas, estas no serán visibles en las imágenes y si son muy grandes, a tal punto que haya una unión entre todas las plantas de arroz, no es posible realizar la detección de malezas. Lo ideal es que las plantas alcancen a ser reconocibles en las imágenes a simple vista, y que no haya una unión entre surcos de cultivos, o de una gran cantidad de grupos de plantas.

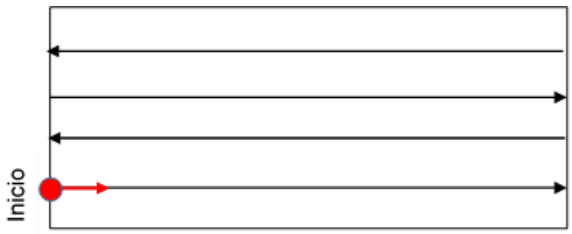
Fig. 3.8 Lote con: a) plantas muy pequeñas para hacer detección de malezas, b) lote con plantas muy grandes para hacer detección de malezas

Fuente: Autor

Después de ello, con una aplicación de vuelo de drones desde un dispositivo móvil se define el área sobre la cual se iba a realizar el vuelo, que para el caso de este proyecto se utilizó la app DJI Pilot y el dron DJI Matrice 200 V2. En la aplicación, se configuran opciones como: altura de vuelo 40m, traslape entre las imágenes del 75%, ISO de 200 y balance de blancos activado. Los vuelos se realizaron antes de las 10 de la mañana, con el fin de evitar reflejos de la luz solar sobre espejos de agua que inciden en el sensor de la cámara. Se realizaron vuelos sobre los lotes desde 15 días antes de la siembra, hasta un

mes después de emergida la planta. La frecuencia con la que se realizaban los vuelos era cada 15 días.

Fig. 3.9 Ejemplo de la planeación de una ruta de vuelo



Fuente: Autor

Para los vuelos realizados con la cámara multiespectral, es necesario realizar una corrección radiométrica de las imágenes adquiridas. La corrección radiométrica es un método para eliminar perturbación o señales de ruido que puedan ser capturadas por la cámara a la hora de tomar las imágenes. El tipo de señales más comunes de este tipo son: radiación solar reflejada por el suelo o superficies reflectantes, radiación dispersada en el ambiente y alteraciones en el movimiento del sensor. Por lo tanto la corrección radiométrica es un método usado para hacer el valor de los pixeles invariante a la hora de la captura de información, intensidad solar, y demás parámetros externos. Para este procedimiento se utilizó un panel de calibración. Este panel tiene unos valores de reflectancia conocidos y constantes para cada una de las bandas de la cámara multiespectral.

Fig. 3.10 Panel de calibración utilizado para la corrección radiométrica de las imágenes capturadas con la cámara multiespectral

Fuente: Micasense, "Calibrated Reflectance Panel". Disponible [online] https://www.micasense.com/accessories#!/Calibrated-Reflectance-Panel/p/86371681/category=25401667

4. Desarrollo

Para la detección de malezas en las imágenes capturadas durante las rutas de vuelo realizadas sobre los lotes, se deben de seguir las siguientes etapas: 1) generación ortomosaicos, 2) procesamiento de las imágenes, 3) definición descriptores propuestos, 4) elección de descriptores más importantes, 5) desarrollo sistema de clasificación, 6) generación de mapas donde solo se muestre las zonas correspondientes a maleza.

4.1. Generación Ortomosaicos

Un ortomosaico es la presentación fotográfica de una zona de la superficie terrestre, donde todos los elementos presentes en la imagen se encuentran a la misma escala, libre de errores y deformaciones. Un ortomosaico se consigue mediante la unión de conjunto de imágenes aéreas o satelitales y la corrección geométrica u ortorectifiacion de las mismas.

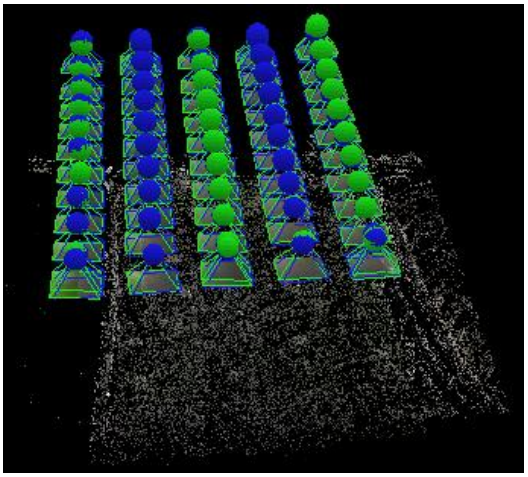
Para la generación de los ortomosaicos se deben de tener las imágenes adquiridas en la ruta de vuelo del dron con cada una de las cámaras. Es necesario dividir el proceso de la generación de los ortomosaicos en dos pasos: a) generación nube de puntos, b) generación del ortomosaico.

Fig. 4.1 Visualización de una ruta de vuelo y los puntos en los que fueron adquiridos las imágenes

Fuente: Autor

Primero se extraen los puntos clave de todas las imágenes, y debido al traslape del 75% configurado en la ruta de vuelo, se encuentran los puntos en común entre las imágenes mediante una triangulación aérea automática. Esta triangulación es posible debido a que el dron posee un GPS, y todas las imágenes capturadas tienen su respectiva georreferenciación. El resultado de esto será una nube de puntos que abarca el área del total de las imágenes capturadas. En la siguiente imagen se puede observar la proyección de una nube de puntos de todo un lote sobre las imágenes capturadas.

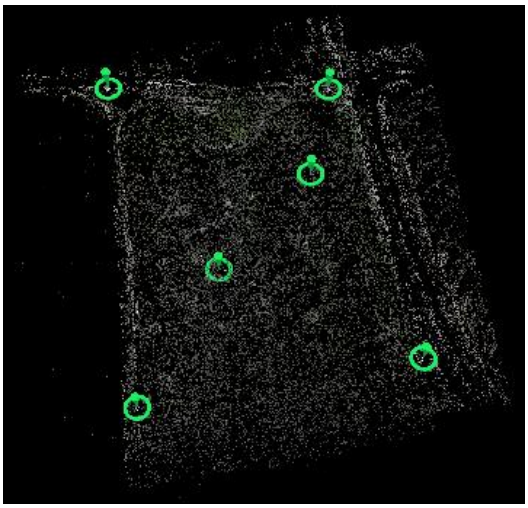
Fig. 4.2 Proyección nube de puntos de un lote a partir de las imágenes capturadas



Fuente: Autor

A continuación se deben de situar los puntos de control en la nube de puntos, para ello al ubicarse sobre un punto, se muestran todas las imágenes que contienen ese mismo punto. Se debe buscar un punto donde aparezcan imágenes de un mismo marcador y establecer como punto de control de la nube el centro de dicho marcador en todas las imágenes desplegadas, esto se debe de hacer con los seis puntos que se encuentran en el lote. El objetivo de ubicar los puntos de control en la nube de puntos es realizar la corrección geométrica del ortomosaico. Esto ayuda a tener una correcta unión entre las imágenes y obtener una georreferenciación más precisa. Como medida opcional, se puede hacer uso de un GPS RTK, para hacer la precisión de la georreferenciación milimétrica. En dicho caso se deberían de ingresar las coordenadas obtenidas con el GPS sobre la nube de puntos, en vez de hacerlo de manera manual.

Fig. 4.3 Ubicación de los puntos de control terrestres en la nube de puntos

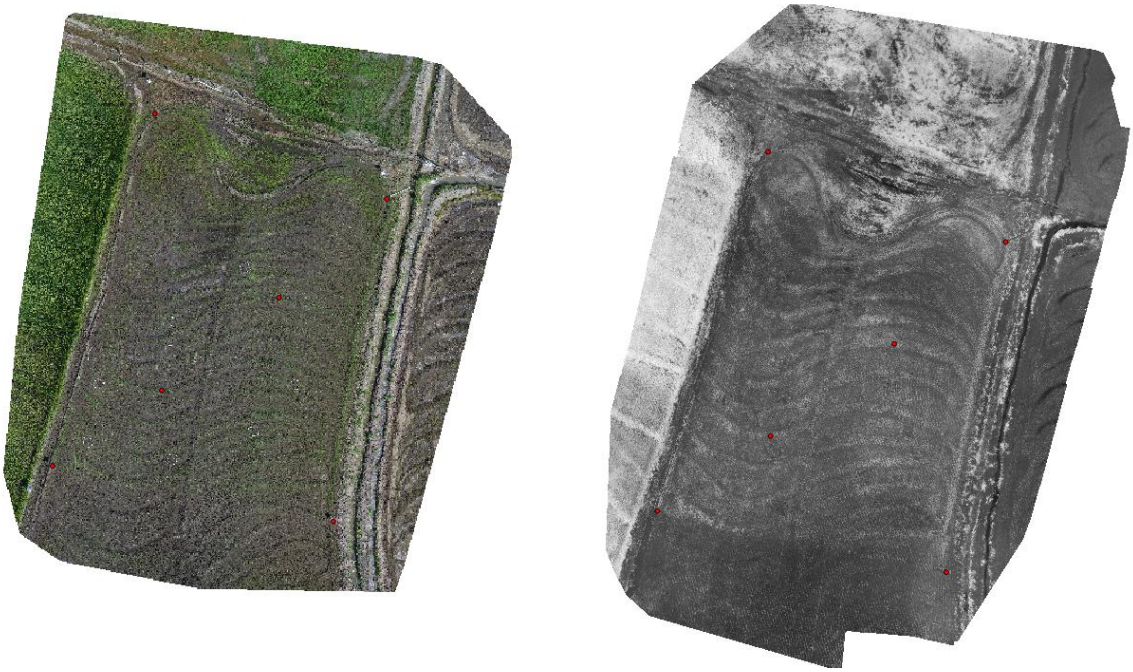


Fuente: Autor

Con base a las nubes de puntos obtenidas y la georreferenciación ingresada, se unen las imágenes, para finalmente formar los ortomosaicos. Como paso adicional, a las imágenes capturadas con la cámara multiespectral, se les debe de realizar una corrección radiométrica antes de hacer la unión de las imágenes. Para ello se le toma una fotografía al panel de calibración y dependiendo de los valores de los pixeles sobre este panel, se realiza la corrección radiométrica a todas las imágenes adquiridas.

Finalmente se unen las imágenes adquiridas por ambas cámaras, para formar los ortomosaicos. A partir de los ortomosaicos rojo e infrarrojo cercano obtenidos de la cámara multiespectral, se calcula el ortomosaico de Índice de Vegetación de Diferencia Normalizada (NDVI). Finalmente el resultado final será un ortomosaico RGB y uno NDVI. Estos archivos deben de ser guardados de tal modo que conserven su georreferenciación y puedan ser exportables a plataformas como Google Maps. La realización de la etapa de generación de ortomosaicos se realizó con ayuda del software Pix4Dmapper.

Fig. 4.4 Ortomosaico generado a partir de las nubes de puntos a) RGB, b) NDVI en escala de grises



a) b)

Fuente: Autor

Debido al margen de error en la georreferenciación de las imágenes a la hora de ser capturadas, los ortomosaicos presentan un corrimiento espacial uno con respecto al otro. Para ello se debe de generar una nueva imagen de uno de los dos ortomosaicos, esto se realiza mediante un tipo de transformación de la imagen y la georreferenciación. Para ello se utiliza la opción de georreferenciación de un software de Sistema de Información Geográfica donde se deben de ubicar en ambas imágenes sitios muy puntuales, como por

ejemplo el centro de los marcadores. Los puntos seleccionados se presentan como un punto rojos sobre los ortomosaicos, como se puede apreciar en la figura anterior. Posteriormente se deben de configurar opciones como el tipo de transformación a usar y el método de remuestreo. El tipo de transformación utilizado siempre en este proyecto fue “polinomial 2” y como método de remuestreo "vecino más cercano”. Una vez realizado esto obtenemos la misma georreferenciación para los dos ortomosaicos, como también una alineación de los elementos presentes en la imagen.

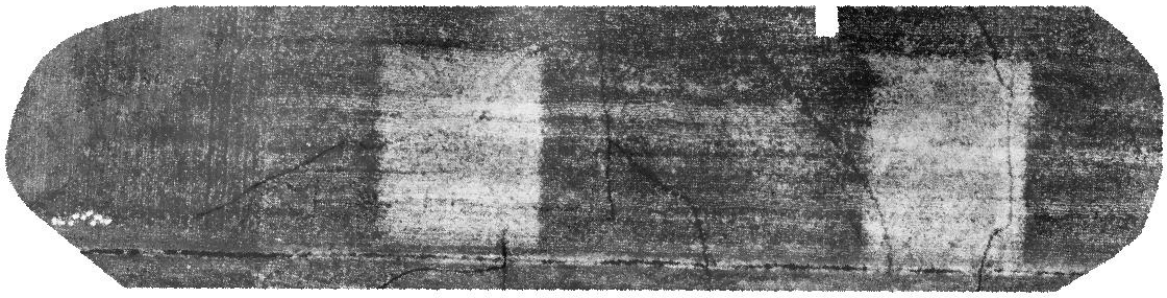
Debido a que ambos ortomosaicos abarcan áreas diferentes, además cada uno de ellos posee elementos no pertenecientes al lote de estudio como vegetación circundante, carreteras y otros cultivos cercanos, se recomienda hacer un mismo recorte en forma de polígono a ambas imágenes, para que solo abarquen el área deseada, que a su vez será exactamente la misma en ambos ortomosaicos. Debido a la diferencia de resolución entre las cámaras multiespectral y RGB, se debe de reducir la resolución del recorte del ortomosaico RGB, para que sea igual al ortomosaico NDVI. Se necesita que ambas imágenes tengan la misma resolución para la posterior extracción de características.

4.2. Procesamiento de las imágenes

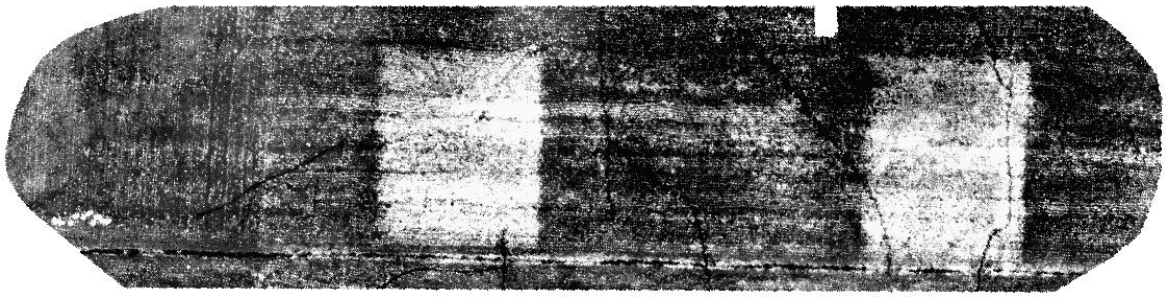
Como primera medida se deben de resaltar solo los objetos verdes en las imágenes, para ello se utiliza el ortomosaico NDVI. Como este se presenta automáticamente en escala de grises, la diferencia entre plantas y suelo o resto de objetos presente en la imagen no es tan marcada, por lo tanto se decide realzar los colores en la imagen. Para ello se establece que los pixeles que presenten un valor de índice NDVI ≥ 0.2 se muestren en una escala de blanco, donde a medida que el valor del pixel sea mayor, la intensidad del color aumente. Se realizó el mismo procedimiento para valores de NDVI ≤ 0 con la escala de color negro. En los rango 0 ≤ NDVI ≤ 0.2 se asigna automáticamente una combinación de color de ambas escalas. Este procedimiento permitió tener una diferencia más marcada entre plantas y resto de objetos, lo que hace el proceso de binarización a realizar más fácil y preciso

Fig. 4.5 a) Ortomosaico RGB, b) Ortomosaico NDVI, c) Ortomosaico NDVI con diferencia de colores más marcada

a)



b)

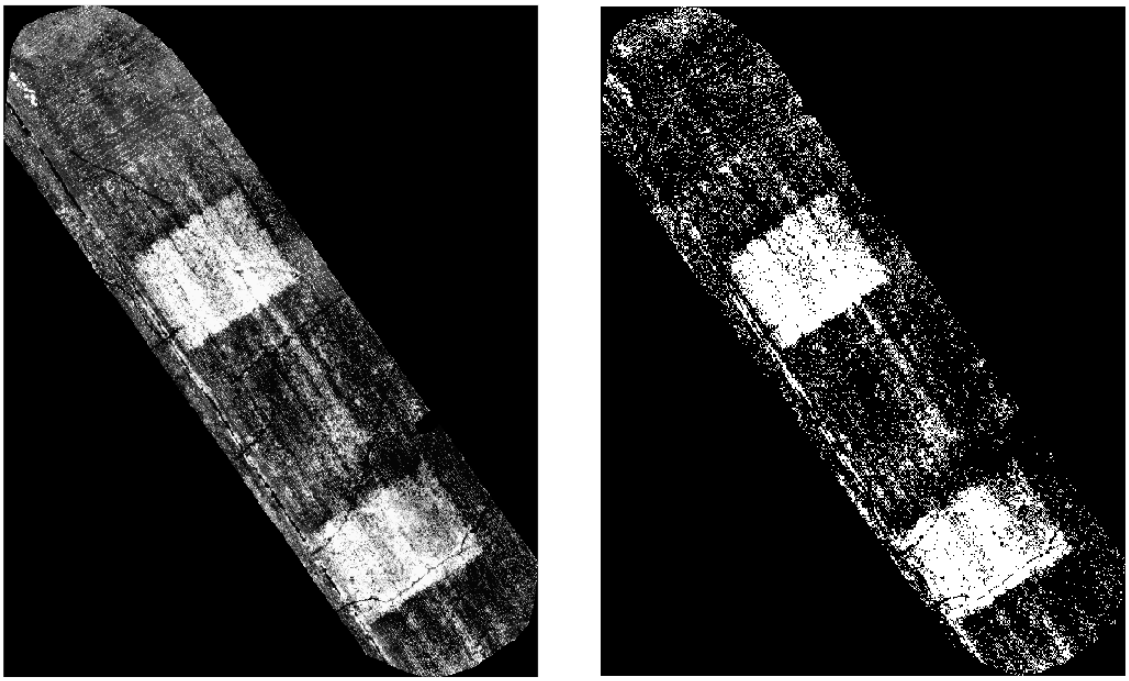


c)

Fuente: Autor

A pesar de visualmente parecer una imagen binarizada la Fig. 4.6 c, los pixeles tienen una escala continua de valores, por lo tanto es necesario realizar una binarización para obtener como resultado una imagen a blanco y negro. Donde los pixeles blancos corresponden a plantas y tienen un valor de “1”, y los pixeles negros corresponden al resto de los objetos, con un valor de pixel “0”. El valor umbral para la binarización se estableció mediante el método Otsu, aunque se observó que al utilizar este método en algunas ocasiones no se obtenía un resultado muy preciso. Esto se debe a que los entornos de programación donde se realizó la binarización como Matlab y Spyder aproximan el polígono irregular de la imagen a binarizar, a una figura rectangular que encierra la imagen, asignándoles a los pixeles no presentes en el polígono original un valor de negro. Por lo tanto se alteran los cálculos realizados por los algoritmos para establecer el valor umbral (ver sección 2.5.1). Es por ello que se debe de corroborar visualmente si los objetos conservan la forma en la imagen binarizada con respecto al ortomosaico NDVI, en caso de no ser así, el valor umbral se debe establecer manualmente. Cabe anotar que la orientación real del ortomosaico NDVI de acuerdo a su georreferenciación es como se presenta en la figura 4.7 a, en la figura 4.6 c se presenta de una manera horizontal para optimizar espacio en este documento.

Fig. 4.6 a) Orientación real del ortomosaico NDVI, b) binarización del ortomosaico NDVI



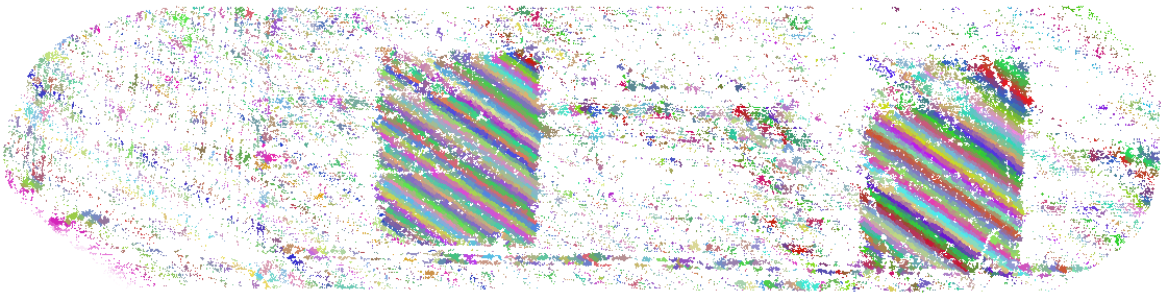
Fuente: Autor

Una vez obtenida la imagen binarizada, se procede a realizar un etiquetado de los objetos. El etiquetado en agrupar pixeles que tengan el mismo valor, en este caso “1”. Para ello se utiliza el etiquetado de componentes conectados (ver sección 2.6.1) con una vecindad de 8 pixeles. Se estableció como cantidad máxima de pixeles por objeto igual a 2000, ya que el tamaño de un pixel sobre el suelo corresponde a 1.8cm^2, por lo tanto el objeto tendría un área máxima de 0.36m^2, lo que corresponde a una zona de 60cm x 60cm. Áreas mayores corresponderían a una gran cantidad de plantas, por ello el motivo de esta limitación. Cuando el algoritmo de etiquetado cuenta 2000 pixeles y siguen habiendo vecindad de pixeles con el mismo valor, este asigna una nueva etiqueta y empieza el conteo otra vez. En la figura 4.7 se puede apreciar dos grandes zonas de plantas, donde hay varios objetos con limitación de pixeles

Los objetos menores a siete pixeles se descartan de la imagen, ya que no aportan información relevante, por lo contrario causan conflicto en etapas posteriores a la hora de calcular los descriptores. Las posiciones de todos los pixeles correspondientes a cada objeto se guardan, al igual que el valor numérico de la etiqueta del objeto, ya que se necesitaran posteriormente para la extracción de características. El archivo que tiene toda la información del etiquetado se guarda en forma de imagen, con su debida georreferenciación. A pesar de los objetos tener un valor de pixel diferente, por defecto la imagen muestra los objetos en escala de grises, lo cual los hace poco diferenciables visualmente. Por lo tanto se establece una paleta de colores aleatorios, para tener un

diferenciador visual de los objetos, aun así hay zonas donde se presentan varios objetos con colores muy parecidos y es difícil diferenciar los objetos visualmente, un ejemplo de ello es la siguiente figura, en donde hay dos zonas que parecieran estar conformadas por líneas, pero al hacer zoom a la imagen, se puede detallar la diferencia de los objetos que tienden a tener una forma cuadrada.

Fig. 4.7 Etiquetado de los objetos presentes en la imagen binarizada



Fuente: Autor

Una vez se tiene la imagen del etiquetado, esta se sobrepone sobre la imagen RGB. Debido a la alineación realizada anteriormente a las imágenes, estas poseen la misma georreferenciación, los objetos etiquetados corresponden a objetos en la imagen RGB, y estas se sobreponen de una manera bastante precisa. Por último, los pixeles blancos presentes en el etiquetado se establecen como transparentes para tener una mejor visualización de los objetos sobre la imagen RGB.

Fig. 4.8 Ortomosaico RGB + etiquetado de los objetos

Fuente: Autor

4.3. Definición descriptores propuestos

Un descriptor es una característica distintiva que sirve para describir o identificar un objeto. Las características más importantes en el análisis de objetos de imágenes son: características de textura y forma. Adicionalmente para la detección de malezas, las características de reflectancia brindan información valiosa de cada uno de los objetos. Para

la definición de los descriptores, se consultó artículos que desarrollaban este mismo trabajo o similares, con el fin de escoger que descriptores a usar. Finalmente se determinó usar algunos descriptores de Haralick como características de textura, Momentos de Hu y algunos descriptores basados en características geométricas como características de forma, y NDVI como característica de reflectancia. En la siguiente tabla se muestra la totalidad de descriptores a usar.

Tabla 4.1 Descriptores propuestos a utilizar en el clasificador

Descriptores de textura	Descriptores de forma	Descriptores de reflectancia
Contraste	Momento 1	NDVI
Correlación	Momento 2	
Energía	Momento 3	
Homogeneidad	Momento 4	
	Momento 5	
	Momento 6	
	Momento 7	
	Área	
	Longitud del eje mayor	
	Longitud del eje menor	
	Diámetro equivalente	
	Excentricidad	
	Solidez	
	Extensión	

Fuente: Autor

Los descriptores de textura y los momentos invariantes de Hu se explican en las secciones 2.8.2 y 2.7.1 respectivamente. Los descriptores basados en características geométricas y NDVI se explican a continuación:

Tabla 4.2 Definición descriptores a utilizar

Descriptor	Definición
NDVI	Promedio del valor NDVI de todos los pixeles que conforman cada objeto
Área	Número de pixeles que conforman un objeto.
Longitud del eje mayor	Longitud en pixeles del diámetro más largo de una elipse que tiene los mismos momentos centrales normalizados de segundo orden que el objeto (para saber más acerca de los momentos centrales, ver sección 2.7.1).

Longitud del eje menor	Longitud en pixeles del diámetro más corto de una elipse que tiene los mismos momentos centrales normalizados de segundo orden que el objeto.
Diámetro equivalente	Diámetro de un círculo con la misma área que el objeto.
Excentricidad	Razón entre su distancia focal[7] y la longitud del eje mayor de la elipse que tiene los mismos momentos centrales de segundo orden que el objeto
Solidez	Relación de pixeles que se encuentran en la envolvente convexa[8] y el objeto. Más específicamente es el área del objeto dividida en la envolvente convexa del mismo objeto.
Extensión	Relación de los pixeles del objeto y un cuadrado que delimita el objeto. En otras palabras es el área dividida en el área del cuadrado delimitador.

Fuente: Autor

Una vez se calcula los descriptores a cada uno de los objetos de la imagen, se construye la base de datos. Esto se hace mirando en las imágenes que objetos corresponden a maleza y arroz y obteniendo la etiqueta numérica de estos. Posteriormente de la tabla donde está el cálculo de descriptores para cada objeto de una, se extraen únicamente las reconocidas visualmente como arroz y maleza, para finalmente obtener todos los datos correspondientes a ambas clases de todas las imágenes observadas. Finalmente se construyó una base de datos con un número igual de observaciones para ambas clases

4.3.1. Elección descriptores más importantes

Antes de desarrollar el sistema de clasificación, se debe de establecer el impacto de cada uno de estos a la hora de clasificar nuevos datos, como también que tan discriminantes son entre las clases, ya que algunos pueden aportar información poco valiosa o redundante, haciendo que el algoritmo de clasificación realice una mala clasificación o que el proceso sea computacionalmente costoso. Para esta tarea de reducción de descriptores y evaluación de la información, se utilizan tres técnicas: elección de descriptores basados en la distribución normal de los datos de cada clase, análisis de componentes principales y selección secuencial de características.

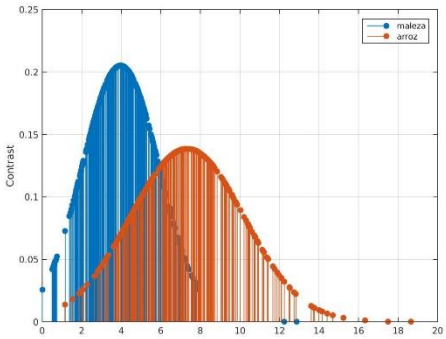
4.3.2. Distribución normal de los datos de cada clase

Un método estadístico propuesto por [27], determina que un descriptor es buen discriminante, cuando la distribución normal de los datos pertenecientes a cada clase es diferente. Para ello se grafican cada uno de los datos, teniendo en cuenta la media y la desviación estándar del conjunto de datos de cada clase. Al hacer esto, se obtiene la

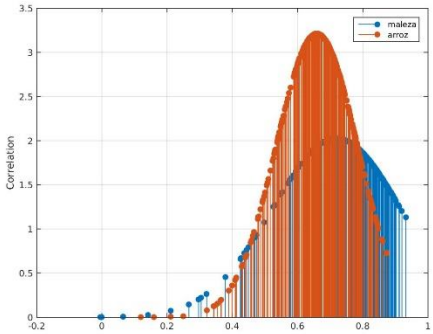
[7] Distancia focal: longitud del segmento que parte del centro de la elipse y termina en uno de sus focos.
[8] Envolvente convexa: polígono convexo cuyos vértices son algunos puntos del objeto.

gráfica de función de densidad de cada clase, donde entre más separadas se encuentren entre sí en la gráfica, el descriptor es un mejor discriminante entre clases. Se descartan los descriptores cuyas funciones de densidad se superponen entre sí. A continuación se muestran las gráficas con las funciones de densidad de cada clase de los 19 descriptores para ambas clases (arroz y maleza).

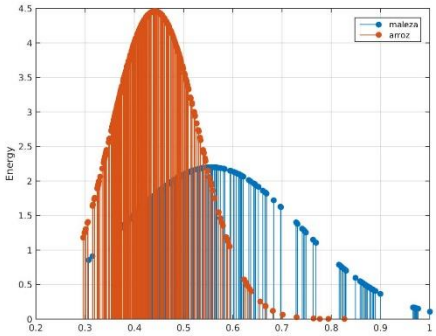
Fig. 4.9 Función de densidad de probabilidad para cada una de las clases de los siguientes descriptores a) Contraste, b) Correlación, c) Energía, d) Homogeneidad, e) Longitud del eje mayor, f) Longitud del eje menor, g) Área, h) Diámetro equivalente, i) Excentricidad, j) Solidez, k) Extensión, l) Momento 1, m) Momento 2, n) Momento 3, o) Momento 4, p) Momento 5, q) Momento 6, r) Momento 7, s) NDVI



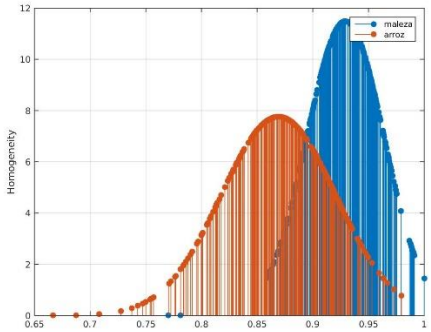
a)



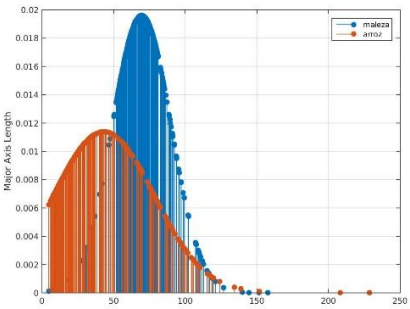
b)



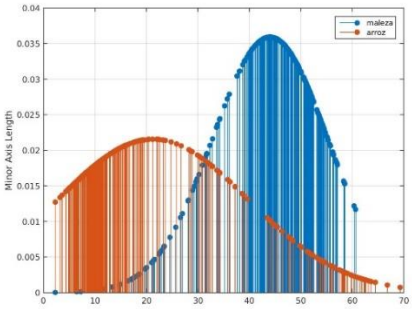
c)



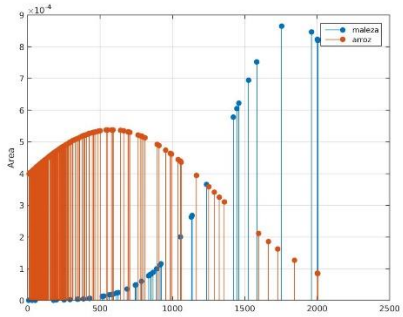
d)



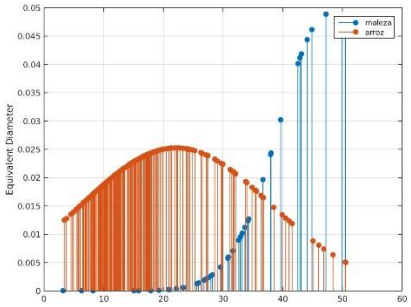
e)



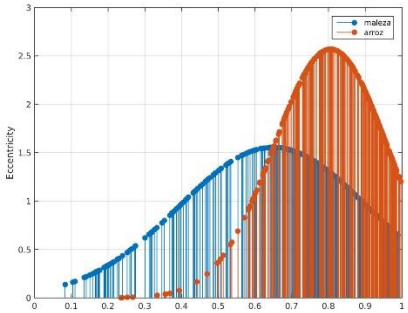
f)


Area
maleza
arroz

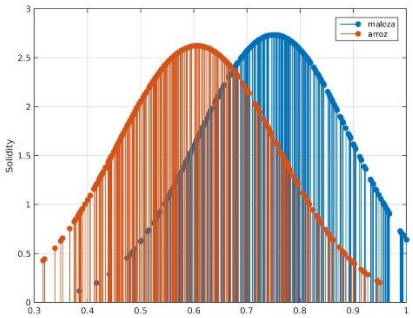
g)


Equivalent Diameter
maleza
arroz

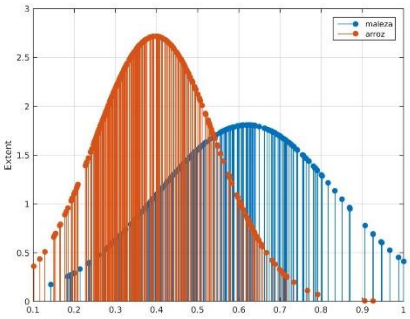
h)


Eccentricity
maleza
arroz

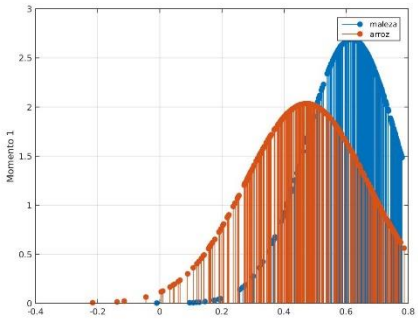
i)


Solidity
maleza
arroz

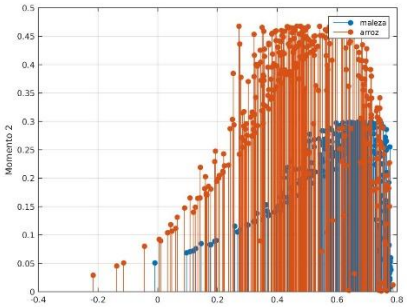
j)


Extent
maleza
arroz

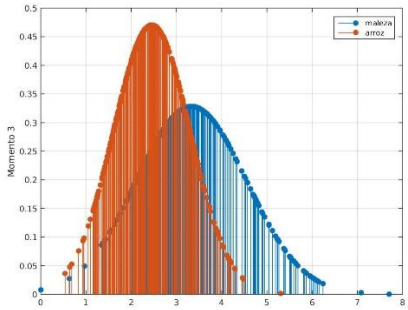
k)


Momento 1
maleza
arroz

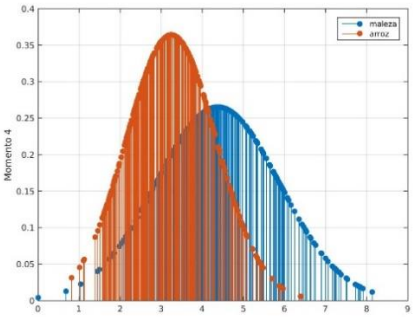
l)


Momento 2
maleza
arroz

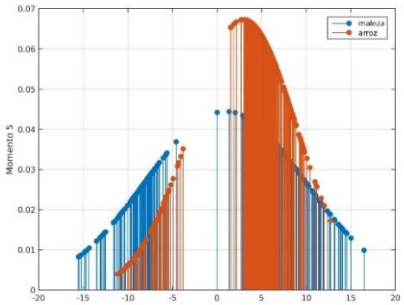
m)


Momento 3
maleza
arroz

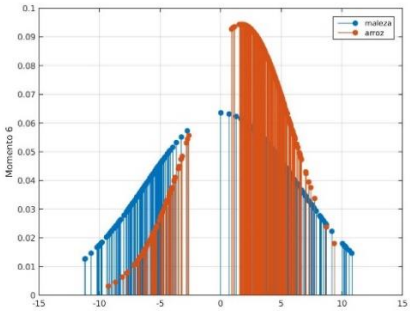
n)



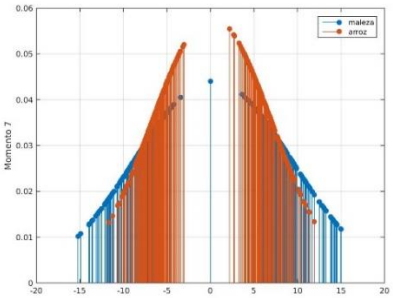
o)



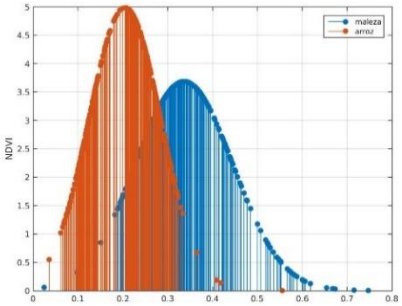
p)



q)



r)



s)

Fuente: Autor

En la imagen anterior, se puede apreciar como el descriptor correlación tiene distribución normal de los datos casi igual para ambas clases. Los descriptores momento 5, momento 6, momento 7 a pesar de tener una distribución normal en sus datos, esta no es continua, y es casi igual para ambas clases, por lo tanto no son buenos descriptores. Debido a la limitación de pixeles establecida en el etiquetado, muchos objetos se sitúan en el mismo punto de la gráfica de los descriptores área y diámetro equivalente, principalmente los objetos correspondientes a malezas. Por lo tanto esta clase en ambos descriptores no tiene una distribución normal de los datos, al igual que el descriptor momento 2, es por ello que también se descartan del algoritmo de clasificacion a crear. A continuación se muestran los descriptores a utilizar mediante este método de selección.

Tabla 4.3 Descriptores resultantes del análisis de la distribución normal de los datos de cada clase

Descriptores de textura	Descriptores de forma	Descriptores de reflectancia
Contraste	Longitud del eje mayor	NDVI
Energía	Longitud del eje menor	
Homogeneidad	Excentricidad	
	Solidez	
	Extensión	
	Momento 1	
	Momento 3	
	Momento 4	

Fuente: Autor

4.3.3. Análisis de Componentes Principales (PCA)

El análisis de componentes principales (PCA) es otro método estadístico para reducir la dimensionalidad de los datos y eliminar características que aporten información poco relevante al conjunto de datos. El resultado de este procedimiento es una matriz de *n x n*, donde *n* es el número de descriptores de la base de datos. Las columnas de esta matriz corresponden a los componentes principales, y las filas al coeficiente de cada uno de los descriptores en la componente. Para calcular los componentes principales, se debe primero normalizar la base de datos, esto quiere decir que el conjunto de datos de cada descriptor debe tener media 0 y desviación estándar 1. La primera componente principal de un grupo de descriptores ($X_1, X_2, X_3, \ldots, X_n$) es la combinación lineal normalizada de dichas variables que tienen mayor varianza [28].

$$Z_1 = \varphi_{11}X_1 + \varphi_{21}X_2 + \cdots + \varphi_{n1}X_n \tag{4.1}$$

Donde los términos $\varphi_{11}, \ldots, \varphi_{n1}$ es el peso que tiene cada variable en cada componente. Estos pesos se calculan a través vectores propios y valores propios (conocidos también como *eigenvector* y *eigenvalues*) de la matriz de covarianza. El cálculo de las posteriores componentes se hace utilizando el mismo procedimiento anteriormente descrito, con la condición de que la combinación lineal de las componentes no puede estar correlacionadas, es decir son perpendiculares entre sí. En este caso se hizo un análisis de componentes principales (PCA) a la base de datos con 19 descriptores normalizada, obteniendo el siguiente resultado.

Tabla 4.4 Matriz con los coeficientes del PCA de la base de datos normalizada

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17	PC18	PC19
D1	-0,2092	-0,3107	-0,3315	-0,0114	0,1273	-0,0951	0,0503	0,2238	-0,0807	0,1524	0,2123	-0,0278	-0,2313	0,1291	-0,1447	-0,0430	0,0549	0,0167	7,0711E-01
D2	-0,0713	0,3268	-0,2492	-0,0380	-0,3361	0,4132	-0,1873	-0,2964	0,3448	-0,0512	0,3458	-0,0067	-0,3536	0,1953	-0,0865	-0,0444	0,0498	-0,0402	2,2204E-16
D3	0,2493	0,0724	0,5277	0,0022	0,0385	-0,1647	0,2680	0,0478	-0,0599	-0,1757	0,1321	-0,0331	-0,2916	0,4445	-0,4406	-0,1329	0,0624	0,0056	6,6613E-16
D4	0,2092	0,3107	0,3315	0,0114	-0,1273	0,0951	-0,0503	-0,2238	0,0807	-0,1524	-0,2123	0,0278	0,2313	-0,1291	0,1447	0,0430	-0,0549	-0,0167	7,0711E-01
D5	0,2405	0,2952	-0,2764	0,0402	0,0933	-0,0688	0,0061	0,1732	-0,1271	-0,0838	0,0088	-0,0216	-0,0220	-0,2362	-0,1475	-0,5399	-0,3215	-0,4867	-1,6653E-16
D6	0,0610	0,4307	-0,0682	0,0329	0,1005	-0,1012	0,0374	0,3502	-0,0422	0,1173	0,4529	0,2221	0,3886	0,3240	0,1857	0,3046	-0,0024	-0,0867	1,3878E-16
D7	0,2370	0,2566	-0,3335	0,0111	0,2046	-0,1536	0,1123	-0,1033	-0,0183	-0,0262	-0,3371	-0,1523	-0,2902	-0,0468	-0,1994	0,6198	0,0860	-0,1494	2,4980E-16
D8	-0,2900	0,1491	0,0761	0,0345	-0,1541	0,1866	-0,2540	0,4524	-0,1139	-0,1624	-0,3441	-0,0976	0,0266	0,1041	-0,0735	-0,0567	0,5490	-0,2648	0,0000E+00
D9	0,2324	0,3194	-0,2462	0,0358	0,0912	-0,0597	0,0042	0,1370	-0,0567	-0,0549	-0,0493	-0,0175	-0,0191	-0,1171	-0,0125	-0,2946	0,2916	0,7459	1,8041E-16
D10	0,2668	-0,2245	0,0285	0,0642	-0,2319	0,2079	-0,1585	0,1461	-0,2928	-0,3023	0,3674	0,1505	0,0563	-0,4399	-0,3309	0,2902	0,0573	0,0407	-2,9143E-16
D11	0,3097	-0,1759	-0,0684	0,0486	-0,0672	0,0579	-0,0151	0,1056	-0,2529	-0,3660	0,0213	-0,0998	-0,3452	0,2255	0,6820	0,0067	-0,0357	-0,0450	-6,9389E-17
D12	0,2492	-0,2349	-0,3267	0,0486	-0,1087	0,1581	-0,1835	-0,1575	-0,1372	-0,0848	-0,2556	-0,0318	0,4429	0,5387	-0,2639	-0,0467	-0,1527	0,0670	1,0408E-16
D13	0,3031	-0,1797	-0,0973	-0,0110	0,0826	-0,1109	0,1543	-0,3666	0,0581	0,1149	0,1983	0,0192	0,1940	-0,0504	0,0962	-0,1688	0,6781	-0,3082	-3,6082E-16
D14	0,2894	-0,1637	0,0455	0,0691	-0,0637	0,0031	-0,0368	0,3610	0,5723	0,0558	0,0932	-0,6164	0,1353	-0,0640	-0,0012	0,0539	-0,0485	0,0065	-3,4694E-17
D15	0,2954	-0,1793	-0,0138	0,0370	-0,0329	-0,0029	-0,0935	0,2621	0,4507	0,1108	-0,2531	0,6974	-0,1795	0,0237	0,0063	-0,0153	0,0145	-0,0359	-9,0206E-17
D16	-0,0808	0,0047	-0,0739	0,6862	-0,2202	0,2757	0,6105	0,0441	-0,0204	0,0841	-0,0851	0,0325	0,0351	-0,0262	0,0145	0,0001	-0,0066	0,0027	1,0408E-17
D17	-0,0882	0,0147	0,0447	0,6773	0,0142	-0,4748	-0,5056	-0,1677	0,0489	-0,0774	0,0905	-0,0044	-0,0655	0,0271	-0,0121	-0,0069	0,0081	-0,0022	6,9389E-18
D18	0,0234	-0,0315	0,1331	0,2108	0,7778	0,5438	-0,1624	-0,0378	0,0419	-0,0430	0,0581	-0,0005	-0,0206	0,0120	-0,0114	-0,0103	0,0049	-0,0046	1,3878E-17
D19	0,2872	0,0567	0,1784	0,0619	-0,1571	0,1611	-0,2487	0,0270	-0,3506	0,7702	-0,0386	-0,1146	-0,1833	0,0091	0,0300	0,0008	-0,0085	-0,0036	-1,2143E-16

Fuente: Autor

Donde *PC1, PC2,…, PC19* son los principales componentes, y *D1, D2,…, D19* son los descriptores de la base de datos y están ordenados de la siguiente manera.

Tabla 4.5 Abreviatura de cada descriptor en la matriz de coeficientes

Abreviatura	Descriptor
D1	Contraste
D2	Correlación
D3	Energía
D4	Homogeneidad
D5	Área
D6	Longitud del eje mayor
D7	Longitud del eje menor
D8	Excentricidad
D9	Diámetro equivalente
D10	Solidez
D11	Extensión
D12	Momento 1
D13	Momento 2
D14	Momento 3
D15	Momento 4
D16	Momento 5
D17	Momento 6
D18	Momento 7
D19	NDVI

Fuente: Autor

Tabla 4.6 Varianzas de cada componente principal

	Porcentaje de varianza total explicada de cada componente	Varianza de cada componente principal
PC1	41,8322	7,9481
PC2	23,7310	4,5089
PC3	7,1377	1,3562
PC4	6,3091	1,1987
PC5	5,5814	1,0605
PC6	4,4869	0,8525
PC7	3,6149	0,6868
PC8	2,4324	0,4622
PC9	1,6356	0,3108
PC10	1,3258	0,2519
PC11	0,6457	0,1227
PC12	0,4803	0,0913
PC13	0,3305	0,0628
PC14	0,2116	0,0402
PC15	0,1322	0,0251
PC16	0,0654	0,0124
PC17	0,0270	0,0051
PC18	0,0203	0,0039
PC19	6,2159E-30	1,1810E-30

Fuente: Autor

Para saber que componentes principales escoger para realizar el análisis de la información, se decidió seguir el criterio de Kaiser [29], el cual dice que se deben de tomar en cuenta únicamente los valores propios o eigenvalues que tengan una varianza mayor a uno en la matriz de covarianza. Esto es igual a decir que solo se deben de tomar en cuenta los componentes principales cuya varianza sea mayor a 1.

Como se puede observar en la tabla anterior, solo las 5 primeras principales componentes tienen una varianza mayor a uno, por lo tanto solo se toman en cuenta esta. Cabe añadir que estas componentes representan el 84,6% de la varianza total de la información, por lo tanto es un porcentaje representativo de todo el conjunto de datos. Debido a que el signo de los coeficientes en la matriz de coeficientes solo indica el tipo de correlación (directa o inversa) entre el componente y la variable, solo se toma en cuenta la magnitud de los coeficientes. Se halla una suma ponderada de cada descriptor a través de los 5 componentes para determinar cuáles presentan mayor varianza entre sus datos y por ende son los más discriminantes entre clases. Para ello se coge el valor de las cinco componentes en cada descriptor, se multiplica por el respectivo porcentaje total de varianza explicada correspondiente al componente y se suman estos cinco valores. Finalmente se escogen los descriptores que tengan un mayor valor de coeficiente promedio como se observa en la tabla 4.7

Tabla 4.7 Suma ponderada de los descriptores a lo largo de las primeras cinco componentes

	Suma ponderada
D1	0,1927
D2	0,1463
D3	0,1614
D4	0,1927
D5	0,1981
D6	0,1403
D7	0,1960
D8	0,1729
D9	0,1979
D10	0,1839
D11	0,1830
D12	0,1924
D13	0,1817
D14	0,1711
D15	0,1713
D16	0,0958
D17	0,0871
D18	0,0835
D19	0,1590

Fuente: Autor

Debido a que la mayoría de coeficientes se concentran entre los valores 0.2 y 0.17 y no presentan una gran diferencia en valor entre ellos, se estableció este como el umbral de clasificación. Los descriptores mayores a 0.17 son tomados en cuenta (delineados con rojo), y el resto son descartados. Teniendo en cuenta las abreviaturas de la tabla 4.5, los descriptores a descartar son: correlación, energía, longitud del eje mayor, momento 5, momento 6, momento 7, NDVI

4.3.4. Selección Secuencial de Características

Por último, se utilizó un tercer método para la reducción del número de descriptores, la selección secuencial de características (en inglés: *sequential feature selection*). El objetivo de utilizar este método es eliminar características o descriptores que aporten información irrelevante o redundante. Existen diferentes maneras de realizar este método, pero el más ampliamente usado es la selección secuencial hacia adelante. Este tipo de selección evalúa todos los subconjuntos con exactamente un descriptor, y selecciona el que mejor prediga la salida. Luego se añade a este subconjunto el descriptor que produce mayor mejoría en la predicción, y así consecutivamente hasta que no ya no haya mejora alguna al añadir más descriptores [30].

El algoritmo creado para realizar esta tarea añade descriptores secuencialmente. Cada que se añade un nuevo descriptor, se realiza una validación cruzada, es decir que se divide la base de datos en 10 partes iguales aleatorias, y se utilizan diferentes porciones de los datos para entrenar y validar (con los datos de entrenamiento se predicen los datos de validación). Al realizar este tipo de validación, el algoritmo halla el promedio de datos clasificados erróneamente y en base a eso va evaluando al subconjunto de descriptores. Se van añadiendo descriptores hasta que este error no disminuya más. Se tomaron los 19 descriptores originales de la base de datos, y se aplicó el algoritmo. A continuación se muestran los resultados obtenidos.

Fig. 4.10 Descriptores resultantes del algoritmo que realiza el proceso de selección secuencial de características

```
Step 1, added column 9, criterion value 0.00261165
Step 2, added column 7, criterion value 0.00219662
Step 3, added column 6, criterion value 0.00134607
Step 4, added column 15, criterion value 0.0012423
Final columns included:  6 7 9 15

fs =

     0     0     0     0     0     1     1     0     1     0     0     0     0     0     1     0     0     0     0
```

Fuente: Autor

La figura anterior quiere decir que la mejor combinación de descriptores para predecir los datos de salida son los descriptores 6, 7, 9, 15 (ver tabla 4.5 para saber cuáles son estos descriptores). Cabe anotar que el algoritmo no añade más descriptores, ya que esto no disminuiría el error en clasificación.

4.4. Desarrollo sistemas de clasificación

Como primera medida, se creó una base de datos con 596 observaciones, donde 298 datos correspondían a maleza y los otros 298 a plantas de arroz. Cada observación correspondía a un objeto y que contiene los datos del cálculo total de los 19 descriptores, como también la clase a que pertenece dicho objeto (arroz o maleza). Se desarrollaron 3 tipos diferentes de sistemas de clasificación (Random Forest, Maquina de Soporte Vectorial, K Vecinos más Cercanos). Para cada sistema de clasificación se crearon tres modelos entrenados, uno para cada combinación de descriptores resultantes de los métodos de reducción propuestos en la sección 4.4. Finalmente se obtuvo un total de 9 sistemas entrenados

En la creación de estos modelos se escogió un 85% del total de la información aleatoriamente, donde un 60% correspondía al entrenamiento y el 25% a la validación. Se evaluó el desempeño de cada modelo y se escogió el mejor dentro de cada tipo de sistema de clasificación. Estos tres modelos finales se utilizaron para predecir el 15% restante de la información y comparar la etiqueta real de cada objeto con la predicha, para finalmente hacer una evaluación real del desempeño de cada modelo.

4.4.1. Random Forest

Para este sistema de clasificación se configuraron parámetros como el método de agregación de datos como “bag”, el tipo de aprendizaje se estableció en “árbol de decisión” y la cantidad de árboles en 101, se seleccionó un número impar para evitar la posibilidad de que ambas clases sean predichas como dominantes por una misma cantidad de árboles. Otras opciones como tasa de aprendizaje y dimensión subespacial se configuraron como “predeterminado”. Los tres diferentes modelos de este sistema de clasificación tenían estas mismas opciones configuradas, lo único cambiante en ellos eran los descriptores a utilizar.

4.4.2. Máquina de Soporte Vectorial

Se decidió usar este tipo de sistema de clasificación debido a su óptimo desempeño para clasificaciones binarias. Antes de desarrollar los modelos de este sistema de clasificación, se debe de estandarizar la base de datos. Esto quiere decir el total de los datos para cada descriptor debe de tener media cero y desviación estándar uno. También es necesario establecer algunos parámetros como función del kernel “lineal”, restricción de caja igual a uno y escala del kernel “automático”.

4.4.3. K Vecinos más Cercanos

Al igual que el sistema de clasificación Maquina de Soporte Vectorial, se debe de estandarizar primero la base de datos. Se configuran parámetros como la cantidad de vecinos cercanos a evaluar igual a 5, cabe anotar que el número de vecinos siempre debe de ser impar para evitar que el número de vecinos para ambas clases sea igual. El tipo de distancia a medir entre los objetos a predecir y los objetos de la base de datos se configuran como “distancia euclidiana” y función de ponderación de distancia se establece como “igual”.

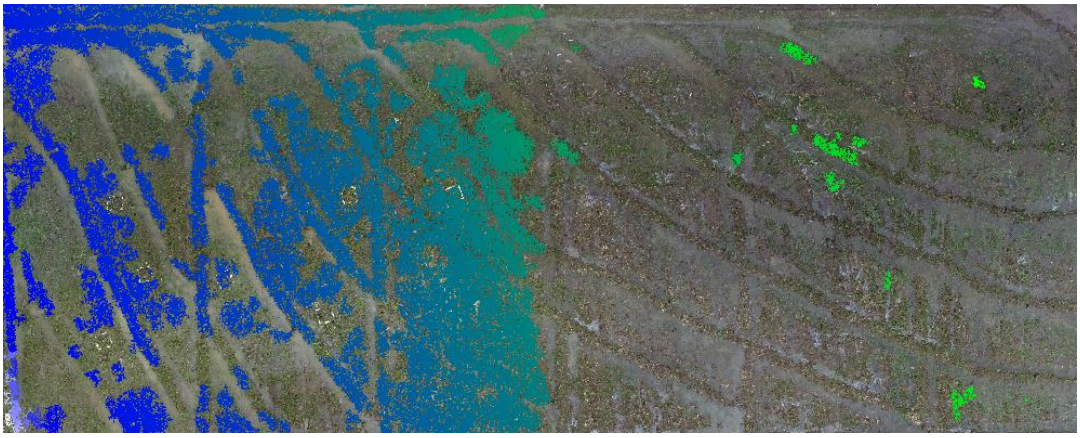
4.5. Generación mapas con zonas de malezas

Una vez se tiene el modelo entrenado de los sistemas de clasificación desarrollados, estos se utilizan para predecir nuevos datos. El modelo asignara una etiqueta numérica de “0” a los objetos clasificados como arroz y “1” a los objetos clasificados como maleza. Se extraen solo los objetos clasificados como maleza. Como en el algoritmo de etiquetado se guardaron las posiciones de los pixeles correspondientes a cada objeto, se recurre a este para saber la posición de los objetos clasificados como maleza en la imagen. Finalmente se guarda la imagen resultante con la georreferenciación del ortomosaico NDVI, ya que fue de este que se realizó el etiquetado.

Fig. 4.11 a) Ortomosaico RGB, b) Ortomosaico RGB + etiquetado de todos los objetos, c) Ortomosaico RGB + objetos etiquetados como maleza utilizando Random Forest y descriptores resultantes del PCA

a)

b)



c)

5. Resultados

Primero se evaluará cuál de los tres modelos entrenados presenta un mejor desempeño para cada tipo de sistema de clasificación. Para ello se utilizará la partición de los datos en entrenamiento y validación (60% y 25%) para obtener las matrices de confusión, a partir de estas se calcularan dos indicadores numéricos: Exactitud y F1-score. Basados en el método de elección de descriptores que presente un mejor desempeño para cada sistema de clasificación, se utilizará el 15% restante de los datos para que los tres modelos escogidos realicen su clasificación y poder comparar la etiqueta real, con la asignada por los modelos. A continuación se muestra la estructura de la matriz de confusión.

Tabla 5.1 Estructura de la matriz de confusión

		Predicción	
		Negativo 0	Positivo 1
Actual	Negativo 0	Verdadero negativo (TN)	Falso positivo (FP)
	Positivo 1	Falso negativo (FN)	Verdadero positivo (TP)

Fuente: Autor

Para los algoritmos de clasificación desarrollados, la clase negativa corresponde a arroz y la positiva a malezas. Verdadero positivo (TP) es maleza clasificada como maleza, verdadero negativo (TN) es arroz clasificado como arroz, falso positivo (FP) es arroz clasificado como maleza y falso negativo (FN) es maleza clasificada como arroz.

La exactitud de un algoritmo de clasificación es la relación entre el número de veces que clasifica correctamente, y el número total de clasificaciones. Se expresa matemáticamente como:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{5.1}$$

La precisión es la relación entre el número de observaciones clasificadas como verdaderos positivos y el número total de predicciones positivas:

$$precision = \frac{TP}{TP + FP} \tag{5.2}$$

La sensibilidad es la relación entre el número de observaciones clasificadas como verdaderas positivas, y todas las observaciones de dicha clase:

$$recall = \frac{TP}{TP + FN} \tag{5.3}$$

F1 Score es la media armónica[9] entre la precisión y la sensibilidad:

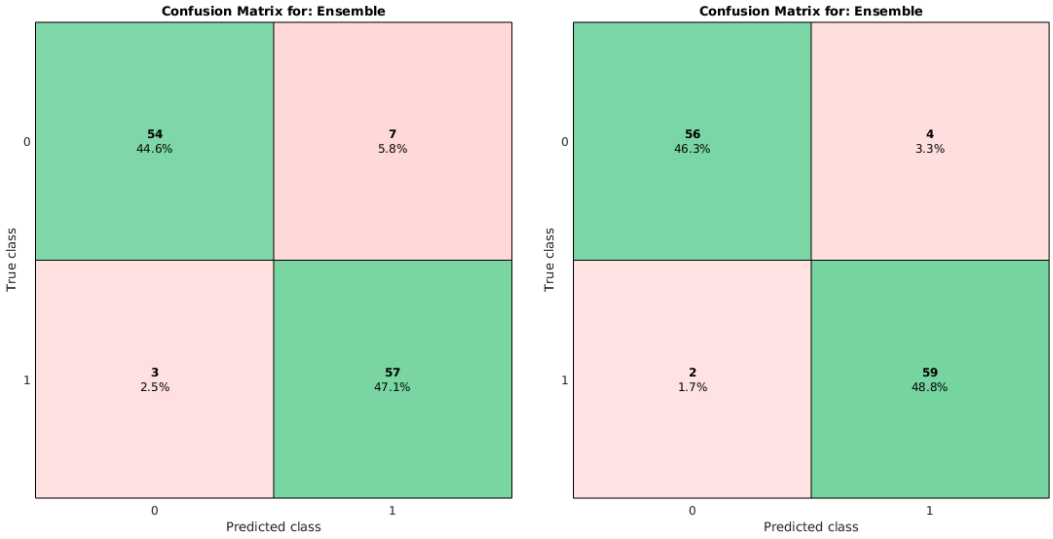
$$F1 = 2\frac{precision * recall}{precision + recall} \tag{5.4}$$

Este tipo de medida es más precisa cuando el número de observaciones por clase es desigual. El F1 score obtiene su mejor desempeño cuando el resultado es igual a 1, que quiere decir que el desempeño y la sensibilidad son perfectas [31].

5.1. Random Forest

Se desarrollaron tres modelos entrenados del tipo Random Forest, uno para cada combinación de descriptores resultantes del análisis de los 3 métodos de reducción de características. Las matrices de confusión y los resultados de desempeño se muestran a continuación.

Fig. 5.1 Matriz de confusión para los algoritmos de clasificación del tipo Random Forest utilizando los descriptores resultantes de: a) distribución normal, b) análisis de componentes principales (PCA), c) selección secuencial de características (SFS).



a) b)

[9] Media armónica (MH) es el inverso del promedio del inverso de n cantidad de datos X_1, X_2,…, Xn. $MH = \frac{1}{\sum_{i=1}^{n}\frac{1}{x_i}}$.

c)

Fuente: Autor

A partir de datos de las matrices de confusión, se tuvieron los siguientes indicadores de desempeño:

Tabla 5.2 Desempeño de los algoritmos de clasificación del tipo Random Forest.

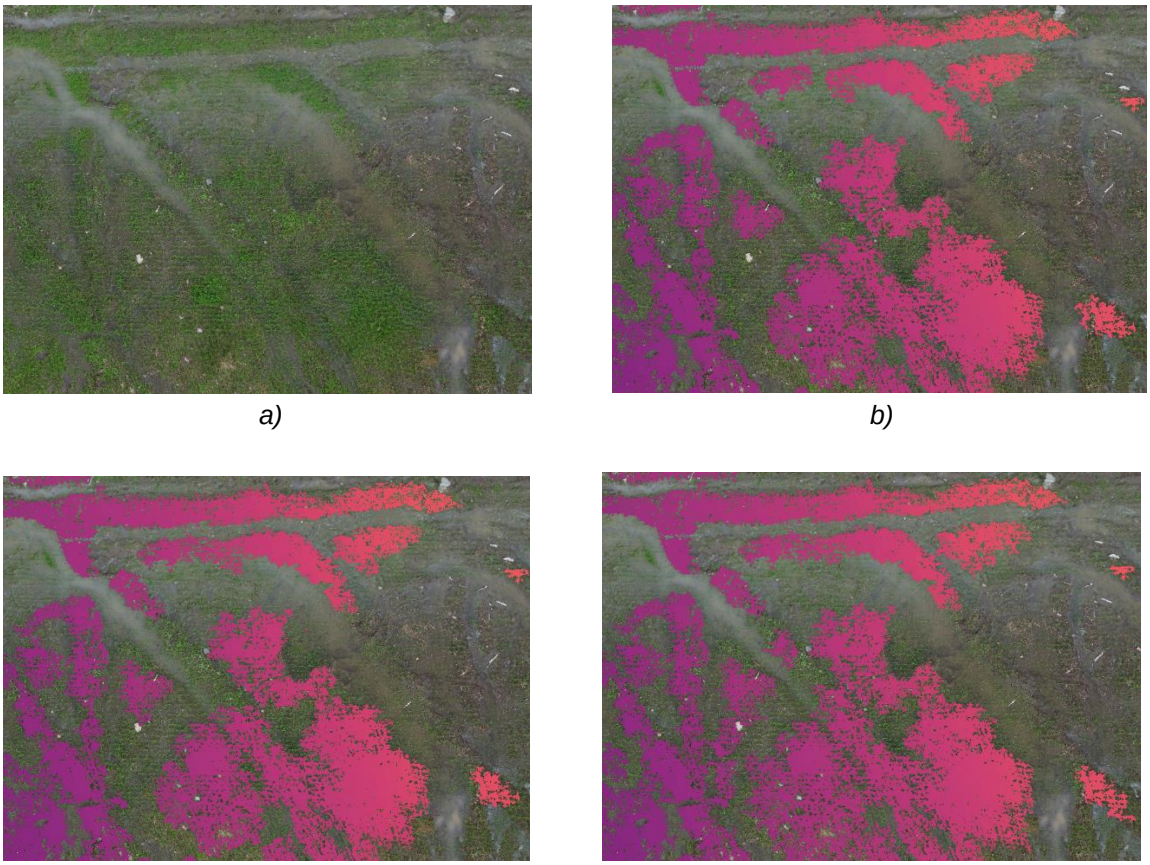
Método de elección de los descriptores	Exactitud	F1 Score
Distribución normal	91.7%	91.9%
PCA	95%	95.1%
SFS	85.1%	85.5%

Fuente: Autor

En la tabla anterior se puede observar que la exactitud del clasificador y el F1 score tienen desempeños similares, esto se debe a que la base de datos tiene un numero de observaciones igual para cada clase. Los tres métodos tienen resultados porcentuales de desempeño bastante buenos, pero el que mejores resultados logró fue en Random Forest utilizando los descriptores resultantes del análisis principal de componentes (PCA).

A continuación se muestra los objetos clasificados como maleza por los tres modelos dentro de un lote.

Fig. 5.2 a) Zona de un lote con presencia de malezas. Clasificación del tipo Random Forest utilizando los algoritmos: b) distribución normal, c) PCA, d) SFS.



a) *b)*

c) *d)*

Fuente: Autor

Tabla 5.3 Numero de objetos clasificados como maleza por los algoritmos de clasificación Random Forest

Método de clasificación	Numero de objetos detectados como malezas
Ninguno (etiquetado total)	17283
Distribución normal	2716
PCA	2706
SFS	2820

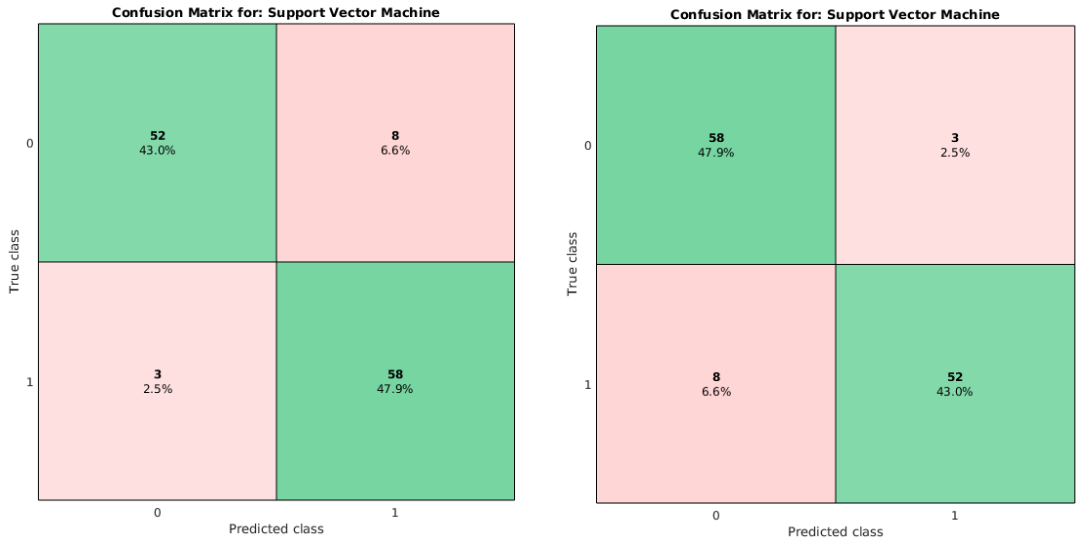
Fuente: Autor

En la figura anterior, se puede observar como los tres modelos tienen un desempeño bastante parecido, ya los tres clasifican casi los mismos objetos como maleza. El algoritmo de selección secuencial de características (SFS) detecta un mayor número de objeto como malezas, aunque algunos de estos son falsos positivos.

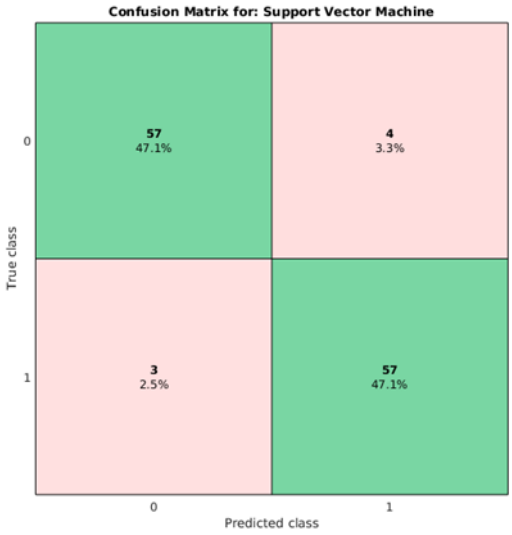
5.2. Máquina de Soporte Vectorial

Se desarrollaron tres modelos entrenados del tipo Maquina de Soporte Vectorial, uno para cada combinación de descriptores resultantes del análisis de los 3 métodos de reducción de características. Las matrices de confusión y los resultados de desempeño se muestran a continuación.

Fig. 5.3 Matriz de confusión para los algoritmos de clasificación del tipo Maquina de Soporte Vectorial utilizando los descriptores resultantes de: a) distribución normal, b) PCA, c) selección secuencial de características (SFS).



a) *b)*



c)

Fuente: Autor

Tabla 5.4 Desempeño de los algoritmos de clasificación del tipo Maquina de Soporte Vectorial

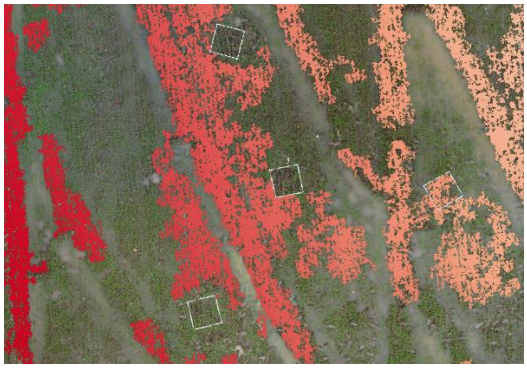
Método de elección de los descriptores	Exactitud	F1 Score
Distribución normal	90.9%	91.3%
PCA	90.9%	90.4%
SFS	94.2%	94.2%

Fuente: Autor

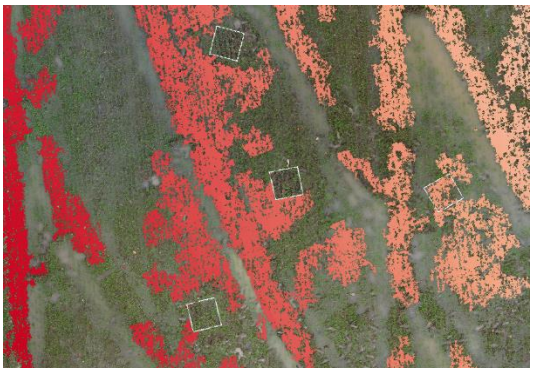
Se puede observar que para este tipo sistema de clasificación, el modelo que presentó un mejor desempeño es el que tiene en cuenta los descriptores de la selección secuencial de características. Esto es una gran ventaja debido a que este modelo tiene en cuenta un número muy pequeño de características o descriptores, lo cual hace que tenga un menor costo computacional frente a los otros modelos.

Fig. 5.4 a) Zona de un lote con presencia de malezas. Clasificación del tipo Maquina de Soporte Vectorial utilizando los algoritmos: b) distribución normal, c) PCA, d) SFS.

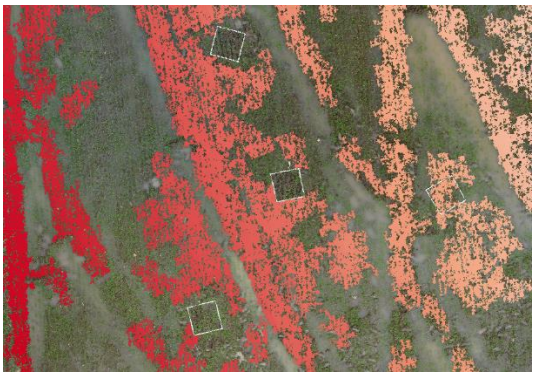
a)



b)



c)



d)

Fuente: Autor

Tabla 5.5 Numero de objetos clasificados como maleza por los algoritmos de clasificación Maquina de Soporte Vectorial

Método de clasificación	Numero de objetos detectados como malezas
Ninguno (etiquetado total)	17283
Distribución normal	2877
PCA	2802
SFS	2659

Fuente: Autor

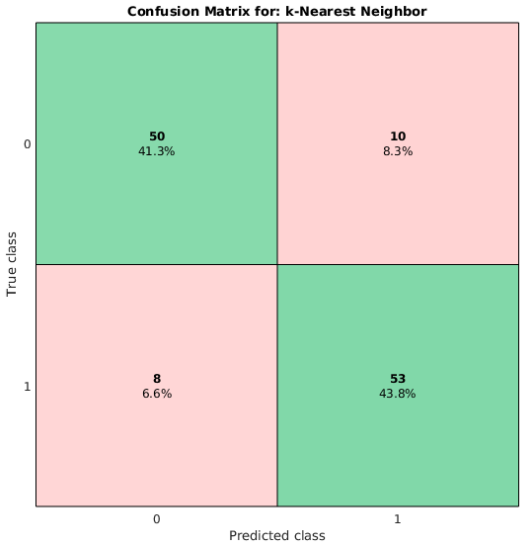
Se puede observar como los tres modelos del tipo Maquina de Soporte Vectorial tienen un desempeño bastante similar, aun así el método de selección secuencial de características detecta una mayor cantidad de objetos grandes como maleza, siendo gran parte de estos falsos positivos. Este modelo también tiende a no clasificar las malezas de menor tamaño, cosa que los otros dos modelos son capaces de hacer con mayor precisión.

5.3. K Vecinos más Cercanos

Se desarrollaron tres modelos entrenados del tipo K Vecinos más Cercanos, uno para cada combinación de descriptores resultantes del análisis de los 3 métodos de reducción de características. Las matrices de confusión y los resultados de desempeño se muestran a continuación.

Fig. 5.5 Matriz de confusión para los algoritmos de clasificación del tipo K Vecinos más Cercanos utilizando los descriptores resultantes de: a) distribución normal, b) PCA, c) selección secuencial de características (SFS).

a) *b)*



c)

Fuente: Autor

Tabla 5.6 Desempeño de los algoritmos de clasificación del tipo K Vecinos más Cercanos

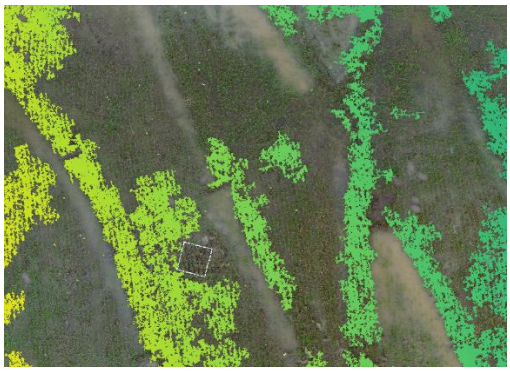
Método de elección de los descriptores	Exactitud	F1 Score
Distribución normal	88.4%	88.1%
PCA	93.4%	93.3%
SFS	85.1%	85.5%

Fuente: Autor

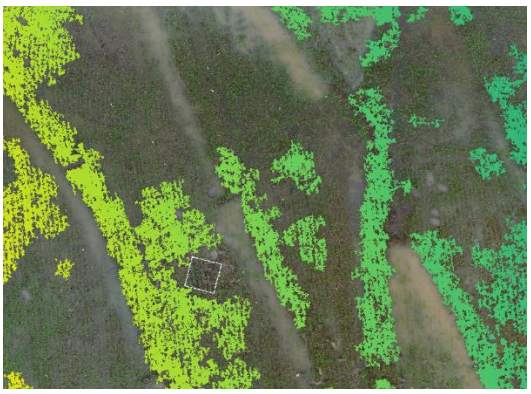
Para este tipo de sistema de clasificación, el modelo que presentó un mejor desempeño fue el que tiene en cuenta los descriptores del análisis principal de componentes (PCA).

Fig. 5.6 a) Zona de un lote con presencia de malezas. Clasificación del tipo K Vecinos más Cercanos utilizando los algoritmos: b) distribución normal, c) PCA, d) SFS.

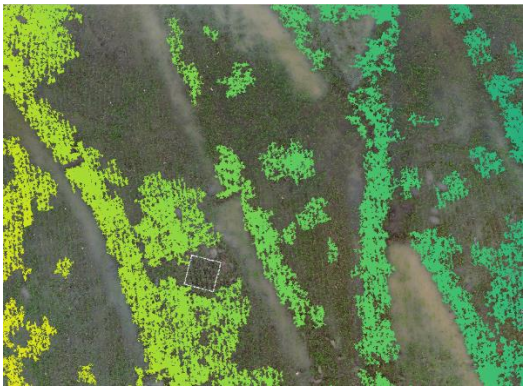
a)



b)



c)



d)

Fuente: Autor

Tabla 5.7 Numero de objetos clasificados como maleza por los algoritmos de clasificación K Vecinos más Cercanos

Método de clasificación	Numero de objetos detectados como malezas
Ninguno (etiquetado total)	17283
Distribución normal	2772
PCA	2797
SFS	2697

Fuente: Autor

Al igual que con los sistemas de clasificación anteriormente mencionados, se puede observar que en los algoritmos del tipo vecinos más cercanos, el modelo con los descriptores resultantes de la selección secuencial de características detecta un mayor número de objetos como malezas, aunque una gran parte de estos realmente sean plantas del cultivo de arroz, esto resulta ser un pro y contra a la vez. También se observa que el modelo SFS tiene inconvenientes a la hora de reconocer malezas de pequeño tamaño, mientras los otros dos modelos tienen un desempeño bastante parecido en todas sus características, por su parte PCA detecta casi los mismos objetos como maleza al igual que el modelo que toma en cuenta los descriptores basados en la distribución normal de los datos.

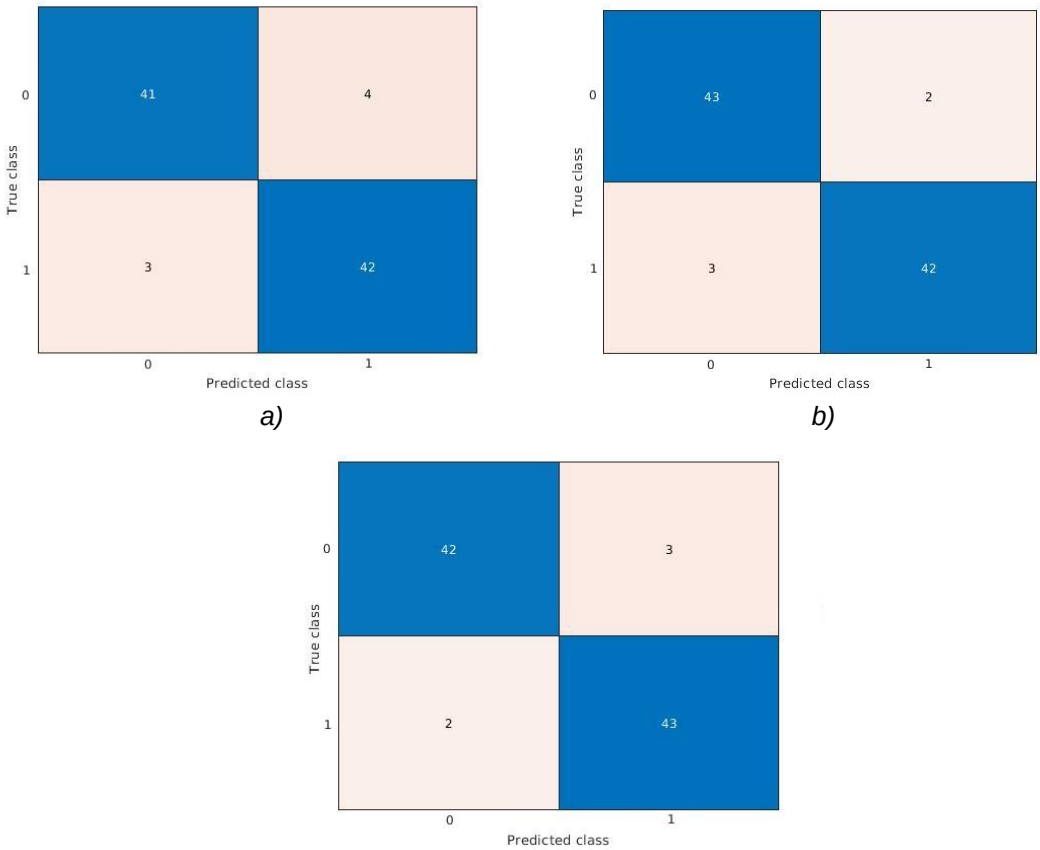
5.4. Comparación entre los tres métodos de clasificación.

En resumen, los modelos de cada tipo de sistema de clasificación que mostraron una mayor exactitud fueron:

- **Random Forest:** PCA 95%
- **Máquina de soporte vectorial:** selección secuencial de características (SFS) 94.2%
- **K vecinos más próximos:** PCA 93.4%

Con estos tres modelos finales, se realizó una clasificación sobre el 15% restante de la base de datos total, que corresponde a un total de 90 datos. En base al dato de la etiqueta actual y predicha para cada objeto por los sistemas de clasificación, se halla nuevamente la matriz de confusión e indicadores de desempeño como exactitud y F1 score.

Fig. 5.7 Matriz de confusión sobre el 15% del total de los datos con los sistemas de clasificación: a) Random Forest, b) Maquina de Soporte Vectorial, c) K Vecinos más Cercanos.



a) *b)* *c)*

Fuente: Autor

Tabla 5.8 Desempeño de los algoritmos resultantes de cada sistema de clasificación

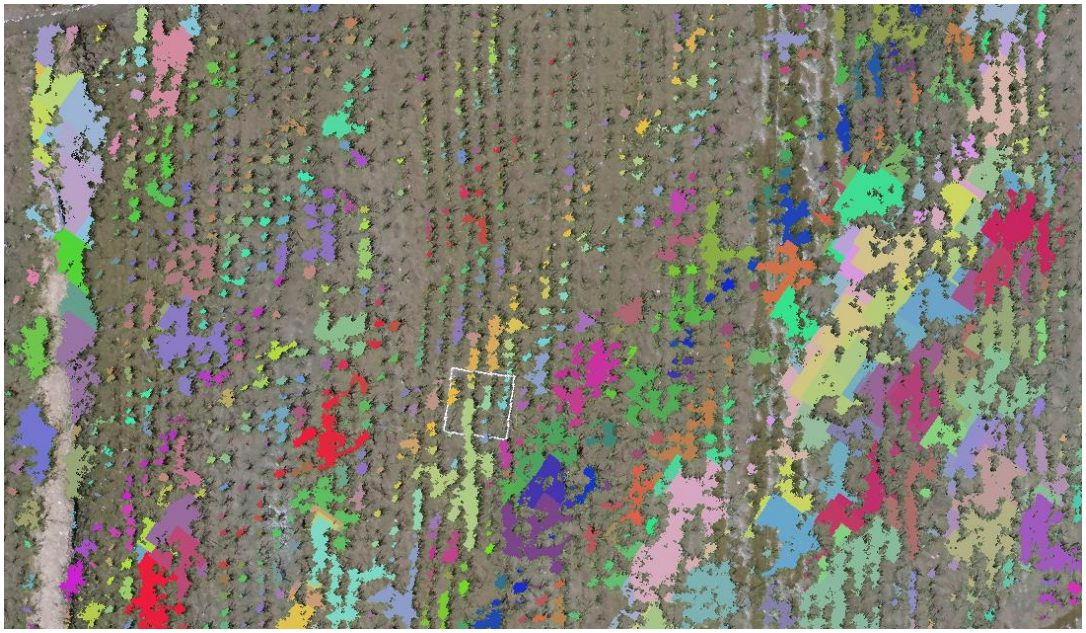
Sistema de Clasificación	Exactitud	F1 Score
Random Forest	92.2%	92.9%
Máquina de Soporte Vectorial	94.4%	94.4%
K Vecinos más Cercanos	94.4%	94.5%

Fuente: Autor

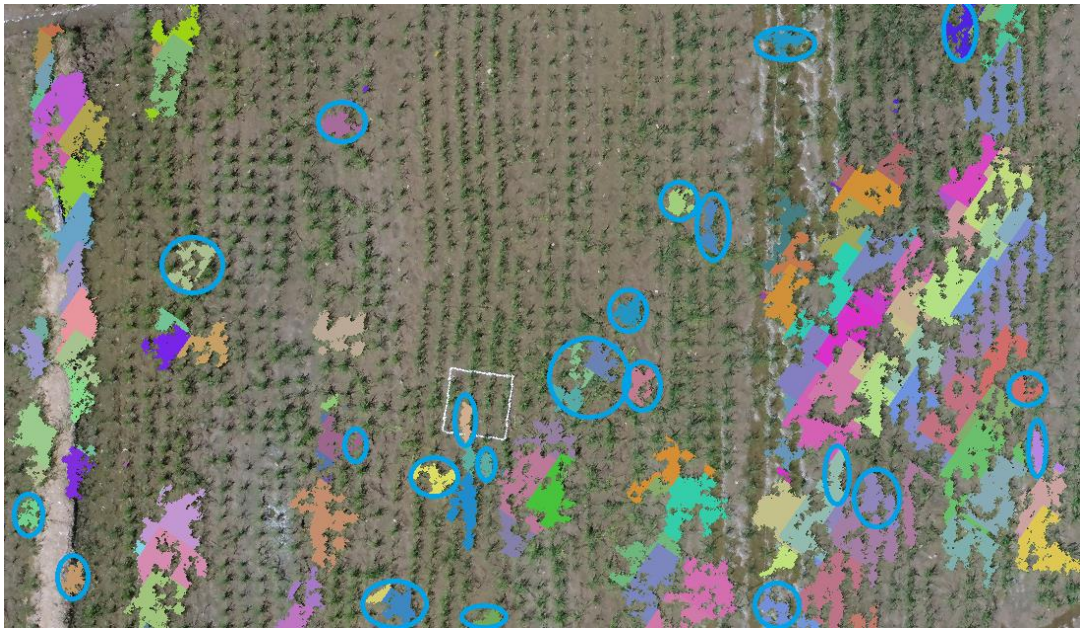
Se puede observar como en la fase de prueba con el 15% de los datos, el desempeño de los modelos varia un poco con respecto a la fase anterior de validación de los datos, aun así esta variación no es muy grande. Cabe anotar que el algoritmo de clasificación Maquina de Soporte Vectorial y K Vecinos más Cercanos tienen un desempeño casi igual entre ellos, mientras Random Forest se encuentra un poco más abajo. A continuación se comparará el desempeño de estos tres algoritmos resultantes a la hora de clasificar nuevos objetos como malezas en una zona de un ortomosaico.

Fig. 5.8 a) Imagen RGB de una zona de un lote, b) etiquetado de todos los objetos. Clasificación: c) Random Forest, d) Maquina de Soporte Vectorial, e) K Vecinos más Cercanos

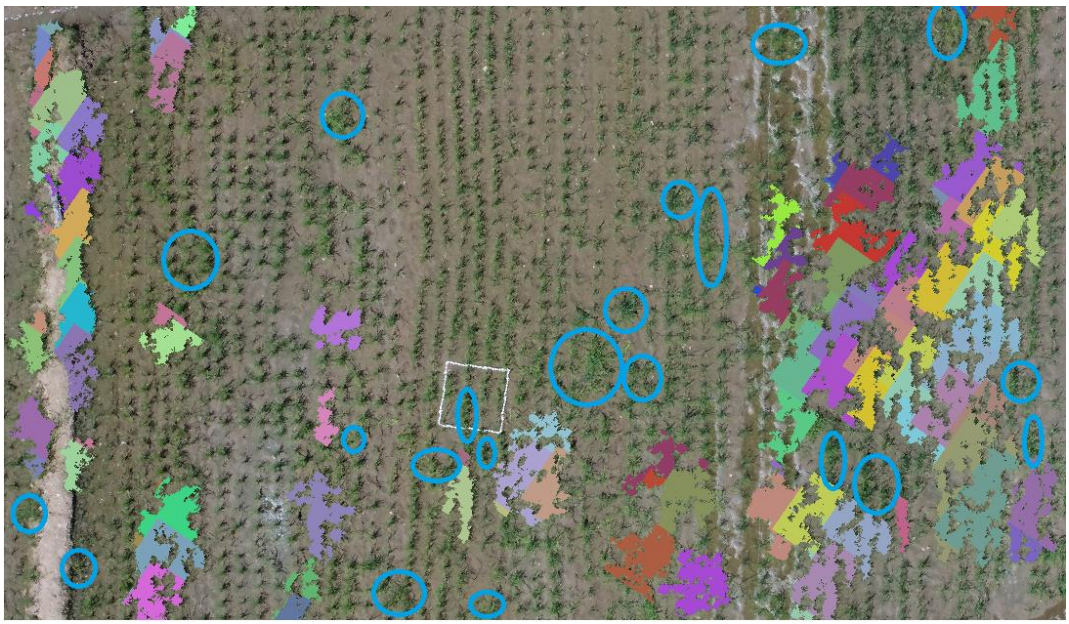
a)



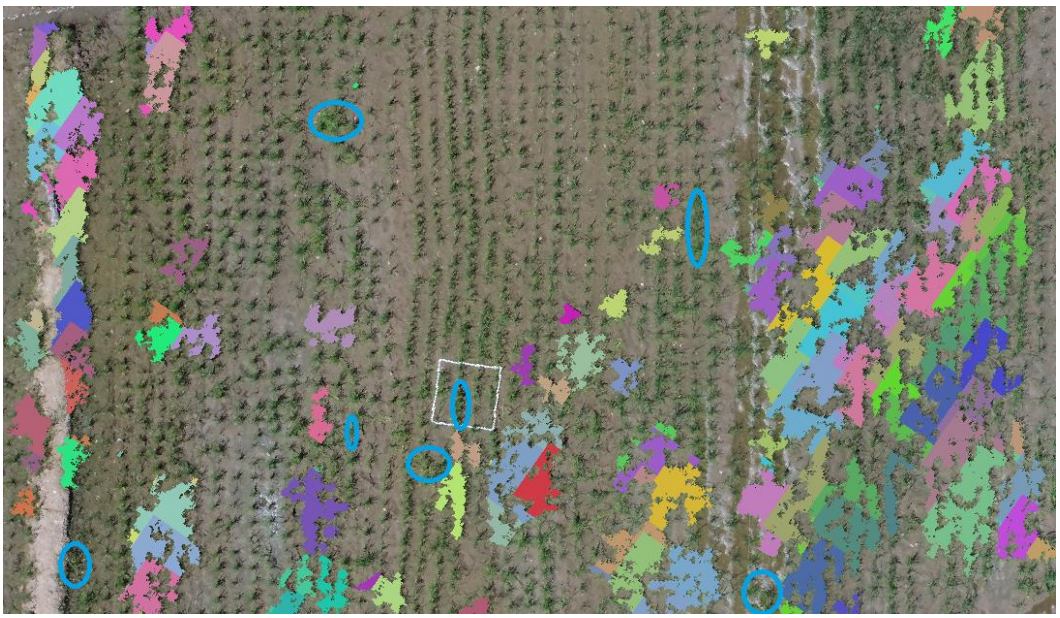
b)



c)



d)



e)

Fuente: Autor

Tabla 5.9 Numero de objetos clasificados como maleza en el ortomosaico

Método de clasificación	Numero de objetos detectados como malezas
Ninguno (etiquetado total)	16399
Random Forest	2535
Máquina de Soporte Vectorial	2064
K Vecinos más Cercanos	2500

Fuente: Autor

Al hacer una comparación visual entre los tres algoritmos, se puede observar como el modelo del tipo Random Forest detecta un número mayor de malezas que los otros dos sistemas de clasificación. El algoritmo de K vecinos más cercanos a pesar de clasificar como maleza un número menor de malezas, la diferencia con Random Forest es muy pequeña y el desempeño de ambos algoritmos es bastante similar, esto se debe a que ambos modelos utilizan la misma combinación de descriptores. Por su parte el modelo del tipo máquina de soporte vectorial detectó un número mucho menor de malezas que los otros dos sistemas de clasificación, esto se debe en parte a la reducida cantidad de descriptores o características que se tienen en cuenta a la hora de crear el algoritmo.

En la fig. 5.8 se puede observar como dentro de los objetos detectados como maleza por el algoritmo Random Forest en la zona del lote de la imagen, se encuentran todos los objetos detectados por los otros dos algoritmos. Esto quiere decir que estos dos algoritmos no detectaron un solo objeto perteneciente a maleza diferente a Random Forest, de hecho en la imagen se puede observar encerrados en círculos de color azul, objetos que solo fueron detectados por este algoritmo y al hacer una corroboración visual con la imagen RGB se puede apreciar que en efecto son malezas.

Por todo lo mencionado anteriormente y después de hacer un análisis visual sobre diferentes imágenes, se puede corroborar que el algoritmo que presenta el mejor desempeño es el Random Forest. Esto puede diferir un poco a los indicadores de desempeños explicados anteriormente mediante la matriz de confusión, pero debido a que la base de datos cuenta con 596 observaciones y en general los ortomosaicos de un lote de tamaño de una hectárea tienen en promedio más de 15mil objetos, la base de datos no es lo suficientemente representativa como para decidir que algoritmo de clasificación presenta un mejor desempeño en base a los indicadores sacados a partir de la matriz de confusión y se hace necesario realizar análisis visual. Además de que todos los lotes tienen características diferentes como el estado vegetativo de la planta, cantidad de agua en el cultivo, el comportamiento de los algoritmos puede variar ligeramente en diferentes ortomosaicos.

6. Conclusiones y recomendaciones

6.1. Conclusiones

1. Se desarrollaron un conjunto de algoritmos con el fin de realizar detección de zonas con presencia de malezas en cultivos de arroz a partir de imágenes RGB y de infrarrojo cercano tomadas desde drones.

2. Es necesario tener en cuenta el estado fenológico de las plantas y las condiciones actuales del lote a la hora de realizar el vuelo para obtener información fiable y de calidad. También se debe de tener en cuenta la hora del día y las condiciones meteorológicas.

3. Tanto la georreferenciación de los ortomosaicos entre ellos, como procesamiento de imágenes (binarización y etiquetado) son pasos fundamentales para la extracción de características por objetos. Una incorrecta georreferenciación entre ortomosaicos provocaría que los pixeles en ambas imágenes no correspondan, lo cual podría ocasionar la extracción de características que no pertenezcan realmente al objeto deseado. También es necesario guardar las nuevas imágenes generadas por los algoritmos con la georreferenciación de los ortomosaicos RGB y NDVI.

4. En general los diferentes métodos de clasificación utilizados en este trabajo tienen un desempeño similar ya que demostraron tener una exactitud cerca al 90%. El éxito de obtener un alto porcentaje en la exactitud de los datos clasificados es como primera medida consultar la literatura que descriptores pueden servir para la tarea que se desea realizar. Después se deben de utilizar métodos de reducción de características como los analizados en este trabajo para obtener los descriptores que mejor discriminan entre clase en base a la base de datos obtenida. Como se pudo observar, cada método de reducción arrojó diferentes sub conjuntos de descriptores como óptimos para los sistemas de clasificación, es por ello que es necesario realizar un análisis visual de los resultados obtenidos por cada uno, para determinar cuál es el que presenta mejores resultados.

5. Se requiere tener una gran variedad de datos de las plantas a lo largo de todo el periodo de detección de malezas (antes de la siembra, 15 y 30 días después de emergidas las plantas de arroz), para tener heterogeneidad en la base de datos y hacer el proceso de clasificación realizado por los algoritmos más preciso. También se debe de tener clara la diferencia visual entre las plantas de arroz y las malezas a la hora de crear la base de datos. Algunas clases de malezas tienen una apariencia visual muy similar al arroz, ya que si se confunden las clases de los objetos al realizar la base de datos, los resultados de los sistemas de clasificación no serán precisos.

6.2. Recomendaciones

- Evaluar el estado fenológico de las plantas de arroz antes de realizar el vuelo. El tamaño óptimo de las plantas para hacer detección de malezas es cuando hay una separación visible entre los surcos de las plantas, usualmente esto ocurre hasta los 30 días. Las plantas deben de serlo suficientemente grandes como para ser visibles en las imágenes capturada por el dron a la altura de la ruta de vuelo. Se recomienda que el lote no este inundado al momento de realizar el vuelo, ya que desmejora la calidad de la información obtenida.

- Se debe de tener en cuenta la hora a la que se va a realizar el vuelo. Horas cercanas al medio día producen reflejos de la luz solar en el lente de las cámaras, si el lote a volar no ha sido drenado. La hora recomendada para volar es antes de las 10 de la mañana. También se debe de tener en cuenta que se debe de configurar un valor de ISO fijo para la cámara RGB, lo recomendable es de 200, activar el balance de blancos y configurar un traslape entre imágenes del 75%.

- Para la cámara multiespectral, siempre se debe de tomar una foto al panel de calibración justo antes, o después de haber realizado el vuelo. Se debe evitar volar el dron con esta cámara cuando la condición de luz solar es muy cambiante (nubes que tapan el sol por momentos mientras el dron realiza la ruta de vuelo).

- Debido a la gran cantidad de información a procesar, es necesario tener un computador con muy buena capacidad de procesamiento y memoria RAM. Desde el proceso de generación de ortomosaicos, hasta la clasificación de nuevos datos son procesos computacionalmente muy costosos.

- Ampliar la base de datos para tener resultados más precisos.

Referencias bibliográficas

[1] Dane, «Encuesta Nacional de Arroz Mecanizado,» 09 08 2019. [En línea]. Available: http://fedearroz.com.co/new/documentos/2019/encuesta_arroz_mecanizado.pdf.

[2] Q. Zhang, Precision Agriculture Technology for Crop Farming, Boca Raton: CRC Press, 2016.

[3] D. Mulla, «Twenty five years of remote sensing in precision agriculture:,» *Biosystems Engineering,* vol. 114, nº 4, pp. 358-371, 2013.

[4] D. Slaughter, D. Giles y D. Downey, «Autonomous robotic weed control systems: A review,» *Computers and Electronics in Agriculture,* vol. 61, nº 1, pp. 63-78, 2008.

[5] S. Huang, S. Wu, C. Sun, X. Ma, Y. Jiang y L. Qi, «Deep localization model for intra-row crop detection in paddy field,» *Computers and Electronics in Agriculture,* vol. 169, 2020.

[6] M. Costamagna, J. Panero, G. Peretti, S. Felissia, G. Lurgo y M. Cicioli, «Detección en tiempo real de malezas a través de técnicas de visión artificial,» de *XIX Workshop de Investigadores en Ciencias de la Computación*, Buenos Aires, 2017.

[7] A. Pantoja, A. Fischer, F. Correa, R. Sanint, A. Ramirez, E. Tascon y E. Garcia, MIP en arroz: Manejo integrado de plagas; artrópodos, enfermedades y malezas, Caracas, 1997.

[8] A. Kwesi y S. K, A Handbook for Weed Control in Rice, Manila: International Rice Reasearch Institue, 1991.

[9] M. Weiss, F. Jacob y G. Duveiller, «Remote sensing for agricultural applications: A meta-review,» *Remote Sensing of Environment,* vol. 236, 2020.

[10] M. Gilabert, J. Gonzalez y J. Garcia, «Acerca de los Indices de Vegetación,» *Teledeteccion,* vol. 8, 1997.

[11] A. Likas, N. Vlassis y V. Jakob, «The global k-means clustering algorithm,» *Pattern Recognition,* vol. 36, nº 2, pp. 451-461, 2003.

[12] L. He, Y. Chao, K. Suzuki y K. Wu, «Fast connected-component labeling,» *Pattern Recognition,* vol. 42, nº 9, pp. 1977-1987, 2009.

[13] M. Safar y C. Shahabi, Shape Analysis and Retrieval of Multimedia Objects. Multimedia Systems and Applications, vol. 23, Nueva York: Springer Science + Bussiness Media, LLC, 2003, pp. 13-18.

[14] M. Hu, «Visual Pattern Recognition by Moment Invariants,» *IRE TRANSACTIONS ON INFORMATION THEORY,* vol. 8, nº 2, pp. 179-187, 1962.

[15] Z. Huang y J. Leng, «Analysis of Hu's moment invariants on image scaling and rotation,» de *2010 2nd International Conference on Computer Engineering and Technology*, Chengdu, China, 2010.

[16] M. Bharati, J. Lui y J. MacGregor, «Image texture analysis: methods and comparisons,» *Chemometrics and Intelligent Laboratory Systems,* vol. 72, nº 1, pp. 57-71, 2004.

[17] M. Presutti, «La matriz de co-ocurrencia en la clasificacion multiespectral,» de *4ª Jornada de Educação em Sensoriamento Remoto no Âmbito do Mercosul*, São Leopoldo, 2004.

[18] R. Haralick, K. Shanmugam y I. Dinstein, «Textural Features for Image Clasification,» *IEEE Transactions on Systems, Man, and Cybernetics* , Vols. %1 de %2SMC-3, nº 6, pp. 610-621, 1973.

[19] D. Calvo, «Diego Calvo,» [En línea]. Available: http://www.diegocalvo.es/aprendizaje-supervisado/.

[20] T. Yiu, «Towards Data Science,» Medium, [En línea]. Available: https://towardsdatascience.com/understanding-random-forest-58381e0602d2.

[21] J. Shubham, «Becoming Human: Artificial Intelligence Magazine,» Media, [En línea]. Available: https://becominghuman.ai/ensemble-learning-bagging-and-boosting-d20f38be9b1e.

[22] B. Stecanella, «Monkey Learn,» [En línea]. Available: https://monkeylearn.com/blog/introduction-to-support-vector-machines-svm/.

[23] O. Harrison, «Towards Data Science,» Medium, [En línea]. Available: https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761.

[24] DJI, «DJI,» [En línea]. Available: https://www.dji.com/matrice-200-series.

[25] DJI, «Zenmuse X4S, DJI,» [En línea]. Available: https://www.dji.com/zenmuse-x4s.

[26] Micasense, «Micasense,» [En línea]. Available: https://www.micasense.com/es/altum.

[27] M. Forero, F. Sroubek y G. Cristobal, «Identification of tuberculosis bacteria based on shape and color,» *Real-Time Imaging,* vol. 10, nº 4, pp. 251-262, 2004.

[28] J. Amarat, «Ciencia de Datos,» [En línea]. Available: https://www.cienciadedatos.net/documentos/35_principal_component_analysis#ejemplo_c%C3%A1lculo_eigenvectors_y_eigenvalues.

[29] H. Kaiser, «A NOTE ON GUTTMAN'S LOWER BOUND FOR THE NUMBER OF COMMON FACTORS,» *British Journal of Statistical Psychology,* vol. 14, pp. 1-2, 1961.

[30] D. Aha y R. Bankert, «A Comparative Evaluation of Sequential Feature,» *Learning Notes in Statistics,* vol. 112, pp. 199-206, 1996.

[31] B. Shmueli, «towards data science,» Medium, [En línea]. Available: https://towardsdatascience.com/multi-class-metrics-made-simple-part-ii-the-f1-score-ebe8b2c2ca1.