

Tesina de Grado

Conteo de plantas de girasol en cultivos en hileras a partir de imágenes UAV

Licenciatura en Ciencias de la Computación



Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Universidad Nacional de Rosario

Fernando Javier Camussi

Dirigida por: Mónica G. Larese

Marzo 2023

Resumen

Las imágenes capturadas por UAV o drones se han vuelto muy útiles durante la última década en la agricultura de precisión, dada su practicidad y su bajo costo. Sin embargo, todavía queda mucho por explorar, actualmente el INTA realiza el conteo de plantas de girasol de forma manual sobre imágenes UAV pero no cuentan con una tecnología que les permita hacerlo de forma automática. En éste trabajo se propone una metodología compuesta por un conjunto de algoritmos para el conteo automático de plantas de girasol en cultivos en hileras a partir de imágenes UAV. Primero se realiza la segmentación de los objetos de plantas de girasol para separarlos del fondo utilizando el índice de vegetación ExG y el método de umbralamiento de Otsu, seguido de mejoras mediante operaciones morfológicas a la máscara obtenida. Luego, se detecta la orientación de las hileras utilizando la imagen de autocorrelación de la máscara y se rota la imagen para que las hileras de los cultivos queden en forma horizontal. El tercer paso consiste en la detección de las hileras y el etiquetado de objetos de plantas de girasol que componen las mismas. Por último se estima la cantidad de plantas de girasol mediante regresión lineal. Para ello se entrenaron modelos utilizando las características de tamaño y forma de los objetos, tales como el área, el perímetro y la compacidad, que son calculadas a partir de un conjunto de coordenadas de plantas generado manualmente sobre un ortomosaico con plantas de girasol en el estadio 19 DDE. Para la detección de objeto de plantas de girasol se obtuvo una *precision* y *recall* de 0.95 y 0.94, respectivamente. En la etapa de testeo se obtuvieron valores de R^2 de 0.96 para el conjunto de coordenadas de plantas y de 0.86 para el conteo realizado por el INTA. También se testeó en el estadio de crecimiento 27 DDE, cuando las plantas son más grandes y complejas, y se obtuvieron valores de R^2 de hasta 0.80. La metodología propuesta resultó ser eficiente para el conteo de plantas de girasol, por lo que podemos imaginar un futuro prometedor, en el campo de la agricultura de precisión, para las tecnologías que combinen imágenes UAV con técnicas de procesamiento digital e inteligencia artificial.

Palabras clave: conteo, plantas, girasol, imágenes, UAV

Índice general

Lista de abreviaturas	III
1. Introducción	1
2. Estado del arte	3
3. Materiales	7
3.1. Imágenes a analizar	7
3.2. Conteo del INTA	13
3.3. Píxeles por metro	16
3.4. Parcelas para entrenamiento y testeo	18
3.5. Conjuntos de coordenadas de plantas y de parcelas	18
4. Metodología	20
4.1. Segmentación	20
4.1.1. Detección de vegetación	20
4.1.2. Mejora de la máscara	23
4.2. Detección de la orientación de las hileras	26
4.3. Enfoque para el conteo de plantas de girasol	29
4.4. Detección de hileras y etiquetado de objetos	32
4.5. Descriptores de características	37
4.6. Conjuntos de muestras	41
4.7. Regresión	42
4.7.1. Modelo lineal multivariado con mínimos cuadrados	42
4.7.2. Regresión Ridge	43
4.7.3. Lasso	43
4.7.4. Regresión de vectores soporte lineal	44
4.7.5. Estandarización y análisis de las variables	44
4.7.6. Hiperparámetro λ para Ridge, Lasso, y SVR	45
4.7.7. Validación cruzada repetida	48

5. Resultados	50
5.1. Segmentación	50
5.2. Detección de la orientación de las hileras	53
5.3. Detección de hileras y etiquetado de objetos	53
5.4. Regresión	54
5.4.1. Entrenamiento	54
5.4.2. Testeo	54
5.5. Metodología completa	60
6. Conclusiones	65

Lista de abreviaturas

CNN Red Neuronal Convolutacional. 6

DDE Días Después de la Emergencia. 7

ExG *Excess Green*. 3

HOG Histograma de Gradientes Orientados. 4

INTA Instituto Nacional de Tecnología Agropecuaria. 1

PPM Píxeles Por Metro. 16

SVM Máquinas de Vectores Soporte. 4

SVR Regresión de Vectores Soporte. 42

UAV Vehículo Aéreo no Tripulado. 1

Capítulo 1

Introducción

En los cultivos de cereales como el girasol existe una estrecha relación entre la densidad de las plantas y la productividad. La densidad deseada u óptima de un cultivo es el mínimo número de plantas que permite alcanzar los máximos rendimientos [1].

Los esfuerzos de mejora basada en genética y en el manejo del cultivo a menudo se centran en identificar la densidad óptima de las plantas para un ambiente determinado. Esto requiere que la densidad de las mismas se determinen tanto en parcelas de investigación como en grandes campos de cultivo [2].

Históricamente, los agricultores han tenido que contar las plantas a mano dentro de una sub-sección de su campo y luego extrapolar el número proporcionalmente al resto del campo. Este método es impreciso porque asume un grado de homogeneidad que es difícil de lograr incluso para las grandes granjas industriales [3].

Los vehículos aéreos no tripulados (UAV) o drones son instrumentos muy prometedores en las ciencias agrícolas. Los vuelos son bastante independientes de las condiciones climáticas y el tiempo, las imágenes se pueden capturar en días nublados, mientras que la captura de imágenes en condiciones similares por satélite no es posible. Además, la tecnología UAV ofrece un ahorro de tiempo y una captura de imágenes más económica y permite un procesamiento de imágenes flexible e inmediato [4], con altura de vuelo variable y mejor resolución de imagen.

Actualmente el Instituto Nacional de Tecnología Agropecuaria (INTA) realiza el conteo de las plantas de girasol de forma manual, a partir de imágenes UAV, lo cual es una tarea tediosa que requiere mucho tiempo. Por lo tanto, existe una clara necesidad de contar con tecnologías que permitan un conteo preciso y eficiente de las plantas de girasol en diferentes estadios de crecimiento.

Las técnicas de análisis para el conteo de plantas basadas en imágenes UAV no se ha publicado ampliamente. Además, cada especie de planta presenta un reto distinto y las técnicas aplicadas para una especie no funcionan de manera directa en otra. Incluso para la misma especie el problema varía dependiendo de su estadio, o incluso del método y la

distancia de sembrado.

El objetivo de ésta tesina es desarrollar un método automático, basado en procesamiento de imágenes e inteligencia artificial, para el conteo de plantas de girasol en cultivos en hileras a partir de imágenes UAV. Dicho método está compuesto por los siguientes algoritmos:

- Segmentación de la imagen, para separar las plantas de girasol del fondo.
- Detección de la orientación de las hileras, para rotar la imagen de manera que las hileras queden de forma horizontal.
- Detección de hileras y etiquetado de objetos de plantas de girasol a las mismas.
- Estimación de la cantidad de plantas de girasol mediante métodos de regresión.

Como parte de éste trabajo se desarrolló en el lenguaje Python (junto a las librerías PyQt5, OpenCV, NumPy, pandas, scikit-learn, PIL y SciPy) la aplicación “Conteo de plantas de girasol” ¹ que implementa los algoritmos propuestos en ésta tesina y permite, mediante una interfaz gráfica, generar también los conjuntos de entrenamiento.

En el capítulo 2 se discute el estado del arte. El *ground truth* se detalla en el capítulo 3. En el capítulo 4 se detalla la metodología propuesta. Los resultados obtenidos y su comparación con los valores reales se encuentran en el capítulo 5. Finalmente, en el capítulo 6 se describen las conclusiones y los trabajos a futuro.

¹<https://repositorio.cifasis-conicet.gov.ar/fernando/girasol>

Capítulo 2

Estado del arte

Liu *et al.* [5] desarrollaron un método de conteo automático de plántulas de trigo utilizando 50 imágenes aéreas adquiridas en la etapa de una hoja en cinco cultivares con cinco densidades de siembra distintas.

Las plántulas de trigo fueron extraídas del fondo utilizando el índice de vegetación ExG (*Excess Green*) y el método de umbralización de Otsu.

Cada plántula de trigo se transformó en un segmento de línea según las características del esqueleto de las plántulas de trigo superpuestas. Con base en las imágenes de esqueleto resultantes y el número de puntos de las esquinas del mismo, se desarrolló una ecuación de conteo automático de plántulas de trigo.

Los resultados mostraron que el método desarrollado puede contar eficazmente las plántulas de trigo con una tasa de precisión promedio del 89.94% y entre las cinco densidades de plántulas probadas la precisión del conteo más alta fue de 97.14%.

En los cinco cultivares probados la precisión del conteo disminuyó a medida que aumentaba la densidad de las plántulas y el conteo no puede continuar si la densidad de las plántulas es demasiado alta.

Dado que las hojas de las plantas de girasol son considerablemente más anchas que las de trigo, con forma más o menos oval y superpuestas entre sí, dicho método no sería factible de llevar a cabo en éste trabajo.

Gnädinger y Schmidhalter [4] llevaron a cabo un análisis en cuatro cultivares de maíz en parcelas con sistemas de siembra en hilera y triangular mediante imágenes registradas por UAV entre BBCH 13-15 (etapa de desarrollo de 3 a 5 hojas) que es cuando el color de las hojas jóvenes son de un color verde más claro que el de las hojas más viejas.

El análisis consistió en calcular la cantidad de plantas de maíz y el grado de cobertura del suelo, para ello se adoptó un proceso de mejora del contraste mediante un estiramiento de descorrelación a la imagen RGB original.

Para obtener la cantidad de plantas de maíz se realizó una segmentación a partir de los

histogramas en el espacio de color HSV seleccionando los umbrales manualmente para que correspondan al color verde claro y llevando a cabo un conteo de la cantidad de objetos obtenidos. Para la detección del grado de cobertura se realizó una segmentación a partir de los histogramas en el espacio de color L*a*b* seleccionando los umbrales manualmente para que correspondan al color verde.

Se encontraron correlaciones con un valor de R^2 cercano a 0.89 entre plantas contadas *in situ* y contadas a partir de imágenes que adaptaron un procedimiento de mejora de contraste. En promedio, el error entre las plantas contadas visualmente y digitalmente fue menor al 5 %. La cobertura del suelo osciló entre el 76 % y el 83 % en estas etapas. En contraste, los autores destacan, que la correlación entre la cobertura del suelo y la cantidad de plantas contadas digitalmente fue muy baja (R^2 de 0.023).

En el caso de las plantas de soja no se visualizó un cambio entre el color verde de las hojas más viejas y las más nuevas, lo cual es fundamental para poder llevar a cabo el análisis propuesto.

Wang *et al.* [6] propusieron un método automático para la detección y enumeración de palmeras de aceite utilizando imágenes UAV. Éste método consta de tres pasos: Primero, las imágenes se clasifican en clases de vegetación y no-vegetación a nivel de píxel utilizando un clasificador SVM (Máquinas de Vectores Soporte) donde las muestras de entrenamiento se seleccionaron mediante interpretación visual. Segundo, se extraen características a nivel de parche de las áreas clasificadas como vegetación mediante un descriptor de características para palmeras basado en HOG (Histograma de Gradientes Orientados) y utilizando ventanas móviles adaptativas. Tercero, se aplica un segundo clasificador SVM para detectar palmeras de aceite individuales donde las muestras de entrenamiento fueron parches seleccionados manualmente de manera que en el caso positivo contengan una palmera de aceite completa y en el caso negativo contengan una palmera incompleta u otra vegetación.

Dicho método se entrenó en un sitio y se validó de forma independiente en otros cuatro sitios con diferentes situaciones. La precisión general de la detección de palmeras fue del 99.21 % en el sitio de entrenamiento y del 99.39 %, 99.06 %, 99.90 % y 94.63 % en los cuatro sitios de validación.

Zhao *et al.* [7] desarrollaron un método para el conteo de plantas de colza a partir de imágenes RGB de ultra alta resolución tomadas mediante UAV en dos etapas tempranas de crecimiento.

Un total de 18 parcelas de muestra fueron seleccionadas para el estudio y estimación de conteo de plántulas de colza. Las mismas se contaron y registraron manualmente después de la recolección de imágenes el mismo día.

Las plantas de colza se segmentaron utilizando índices de vegetación y el método

de umbralización de Otsu, luego se extrajeron características de forma tales como área, relación largo-ancho y excentricidad de los objetos obtenidos tras la segmentación y finalmente se entrenaron modelos de regresión escalonada múltiple para predecir la cantidad de plantas.

Los resultados demostraron que las características de forma tenían fuertes correlaciones con el número de rodales de plántulas medido en el suelo. Los modelos de regresión alcanzaron valores de R^2 de 0.845 y 0.867, respectivamente, para las dos etapas de crecimiento. Los errores absolutos medios del conteo total de rodales fueron 9.79 % y 5.11 % para las dos etapas respectivas. Un solo modelo sobre estas dos etapas tuvo un valor R^2 de 0.846, y el número total de plantas de colza también se estimó con precisión con un error relativo promedio de 6.83 %.

Guo *et al.* [8] propusieron un método de dos pasos basado en procesamiento de imágenes para detectar y contar el número de panojas de sorgo en imágenes de alta resolución capturadas por UAV.

Primero se realizó una segmentación basada en píxeles con un enfoque de aprendizaje automático supervisado basado en árboles de decisión para detectar las regiones candidatas a ser panojas de sorgo utilizando como predictores las componentes correspondientes a los espacios de color RGB, HSV, y $L^*a^*b^*$.

Luego, a las regiones candidatas, se les extrajeron características morfológicas como área, excentricidad y *extent*, entre otras, y se utilizaron junto a los números de panojas correspondientes como respuesta en una SVM cuadrática para predecir la cantidad.

Para demostrar el rendimiento del método propuesto, se etiquetaron manualmente 52 imágenes: los valores de *precision* y *recall* para la detección de la panoja fueron 0.87 y 0.98 respectivamente, y el coeficiente de determinación R^2 entre el método de conteo manual y el nuevo fue de 0.84.

Koh *et al.* [9] desarrollaron y describieron un algoritmo basado en *template matching* para la estimación del conteo de plántulas en las primeras etapas de crecimiento mixto de plántulas de cártamo utilizando imágenes capturadas mediante UAV en dos experimentos de campo distintos.

Las plantas se separaron del fondo mediante un proceso de segmentación *chessboard* y la proporción de verde en los píxeles.

Luego se utilizó el canal verde de la imagen RGB correspondiente a las plantas extraídas junto con una plantilla agrupada que fue generada con 1000 parches de imagen que representan plántulas de cártamo para detectar las mismas mediante la técnica de *template matching*.

El algoritmo fue validado en 300 parcelas y los conteos estimados de plántulas se correlacionaron estrechamente con el conteo manual. Obtuvieron valores de R^2 de 0.87,

MAE de 8.18 y RSME de 9.38 para el primer experimento de campo y R^2 de 0.86, MAE de 9.16, RSME de 10.51 para el segundo.

Los autores señalaron que el rendimiento del algoritmo se degrada levemente (sobreestimación o subestimación) para áreas con grupos apretados de plántulas de cártamo, probablemente debido a la situación compleja causada por altos solapamientos entre plantas.

Algunos autores adoptaron un enfoque de *deep learning* para el conteo de plantas. Wu *et al.* [10] propusieron un método eficiente que utiliza la visión por computadora para contar con precisión las plántulas de arroz en una imagen digital utilizando una red de regresión inspirada en una red neuronal profunda y totalmente convolucional. Osco *et al.* [11] abordaron el desafío de estimar el número de árboles de cítricos en huertos muy densos a partir de imágenes multiespectrales de UAV utilizando redes neuronales convolucionales (CNN). Valente *et al.* [12] abordaron un método automatizado para contar plantas de espinaca de 10 semanas de edad a partir de imágenes de UAV de muy alta resolución utilizando visión artificial y transferencia de aprendizaje mediante CNN. Pang *et al.* [13] diseñaron un sistema que utiliza información de descriptores geométricos con redes neuronales profundas para determinar rodales de maíz de temporada temprana a partir de datos aéreos de resolución espacial relativamente baja.

También existen algunas aplicaciones comerciales *on-line* como solución orientada a los productores. Podemos mencionar la aplicación de conteo de plantas en hileras de PrecisionHawk ¹. Ésta trabaja con imágenes de alta resolución considerando los modelos específicos del dron y de la cámara, junto con información de georreferenciación. También requiere asistencia por parte del usuario para identificar partes de su campo como tierra, plantas e hileras. Otro requisito es que las hileras sean uniformes, es decir, que se haya plantado la misma cantidad de semillas en cada una de forma equidistante y que se le especifique el número de semillas plantadas y el espacio aproximado entre las mismas. Se remarca que la herramienta es más útil en las etapas tempranas de emergencia de la planta.

¹<https://www.precisionhawk.com>

Capítulo 3

Materiales

3.1. Imágenes a analizar

Para el desarrollo del presente trabajo se cuenta con 9 imágenes UAV conteniendo cultivos en hileras de plantas de girasol (*Helianthus annuus* L.), de distintos genotipos, y una planilla con un conteo manual de plantas, realizado sobre imagen, de 150 parcelas. Dicho material fue proporcionado por el Instituto Nacional de Tecnología Agropecuaria (INTA), estación experimental Balcarce.

El objetivo de los experimentos fue comparar los distintos genotipos entre sí, a fin de identificar si algunos de ellos se comportaban mejor con una menor disponibilidad de agua, para en un futuro llegar a obtener cultivares con mejor tolerancia a situaciones de sequía.

Las imágenes corresponden a 2 cultivos en diferentes lugares, cada uno, en 2 fechas distintas. Uno de ellos fue sembrado el día 10/11/15 en Estación Camet, en las afueras de Mar del Plata, provincia de Buenos Aires (cultivo A), mientras que el otro fue sembrado el día 17/11/15 cercano a la localidad de Quemú, provincia de La Pampa (cultivo B). En el primer caso las plantas emergieron (se hicieron visibles por encima de la superficie) el día 18/12/15 y en el segundo caso emergieron el día 27/11/15.

Una de las imágenes, *camet1.tif*, es un ortomosaico completo en la cual se realizó el conteo sobre parcelas con riego en la fecha 19 DDE (Días Después de la Emergencia). Las 8 imágenes restantes son recortes de ortomosaicos rotados manualmente. Dentro de éstas 8 imágenes, tenemos dos imágenes, *camet1cr.jpg* y *camet1sr.jpg*, que son recortes del ortomosaico completo, y una imagen, *camet2cr.jpg*, que contiene 40 parcelas correspondientes al conteo, en la fecha 27 DDE, cuando las plantas ya son más grandes.

En la tabla 3.1 se muestran los detalles de cada imagen: a que cultivo pertenecen, el estadio de crecimiento al momento de capturar las imágenes medido en DDE, si el tiempo estaba nublado, si el cultivo contenía riego, la altura a la que se tomaron las imágenes,

Imagen	Cultivo	DDE	Nublado	Riego	Altura	Tamaño	Geo
camet1.tif	A	19	Sí	Una parte	30m	19586×20000	Sí
camet1cr.jpg	A	19	Sí	Sí	30m	5312×4800	No
camet1sr.jpg	A	19	Sí	No	30m	4896×4272	No
quemu1cr.jpg	B	14	Sí	Sí	20m	6816×5904	No
quemu1sr.jpg	B	14	Sí	No	20m	6544×5312	No
camet2cr.jpg	A	27	No	Sí	30m	4304×3920	No
camet2sr.jpg	A	27	No	No	30m	5632×5280	No
quemu2cr.jpg	B	32	Sí	Sí	20m	6512×5760	No
quemu2sr.jpg	B	32	Sí	No	20m	6384×5184	No

Tabla 3.1: Detalles de las imágenes

el tamaño de la imagen en píxeles, y si la misma posee información de georreferenciación. Todas las imágenes que componen los ortomosaicos originales fueron capturadas utilizando un drone 3D Robotics X8 con una cámara Canon PowerShot ELPH 100 HS sobrevolando a una velocidad de 2 m/s.

A continuación se muestran las imágenes. En el caso de las imágenes camet1.tif y camet2cr.jpg, se marcaron en rojo las parcelas contenidas en el conteo realizado por el INTA.

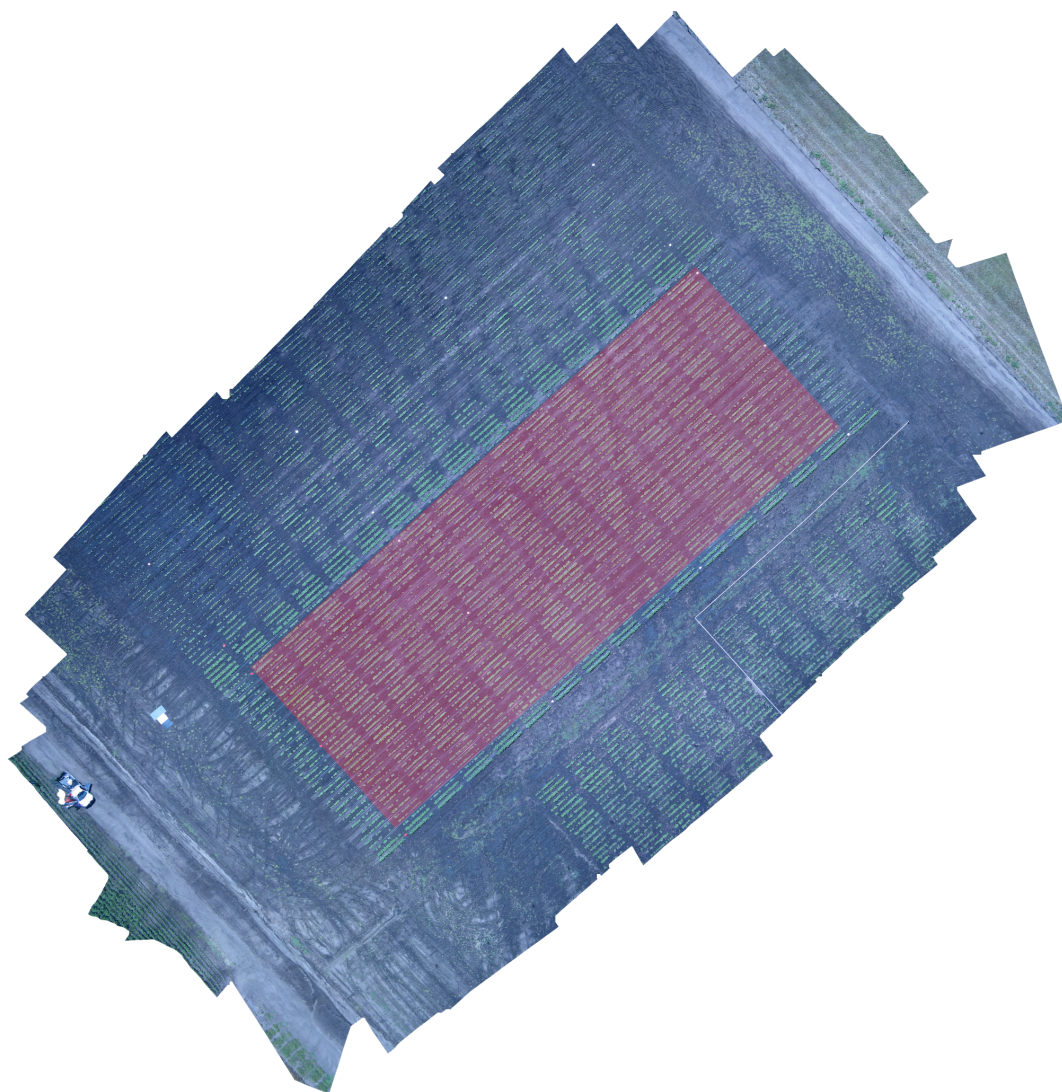


Figura 3.1: Imagen camet1.tif con las 150 parcelas del conteo marcadas en rojo

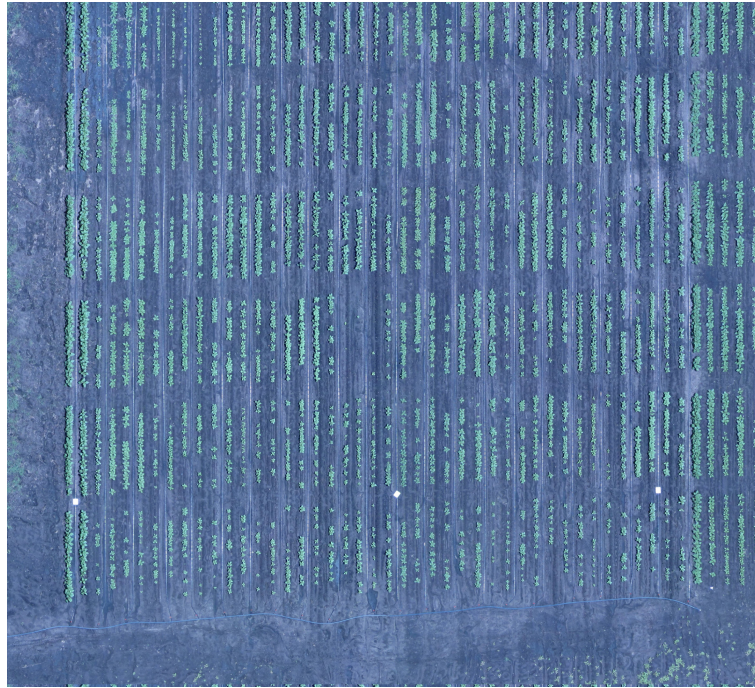


Figura 3.2: Imagen camet1cr.jpg



Figura 3.3: Imagen camet1sr.jpg

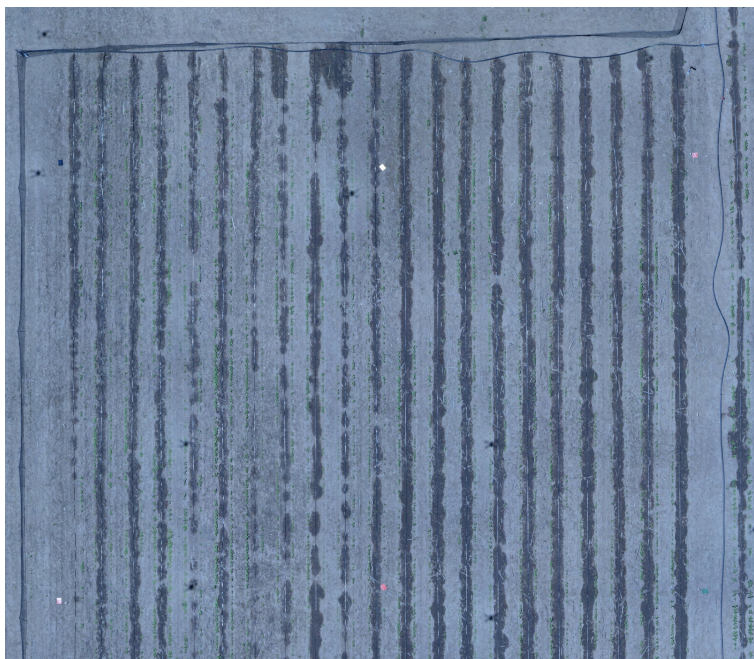


Figura 3.4: Imagen quemulcr.jpg



Figura 3.5: Imagen quemulsr.jpg



Figura 3.6: Imagen camet2cr.jpg con 40 parcelas del conteo marcadas en rojo

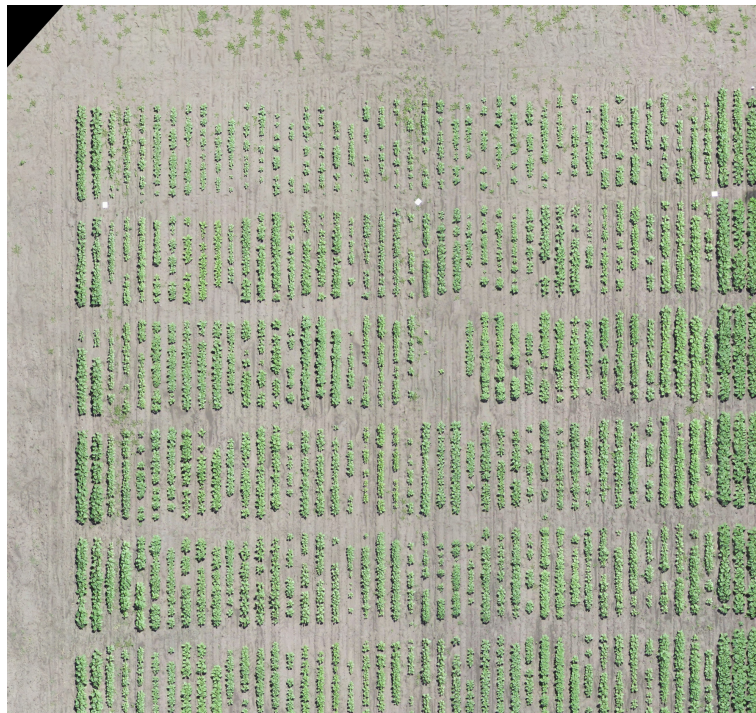


Figura 3.7: Imagen camet2sr.jpg



Figura 3.8: Imagen quemu2cr.jpg

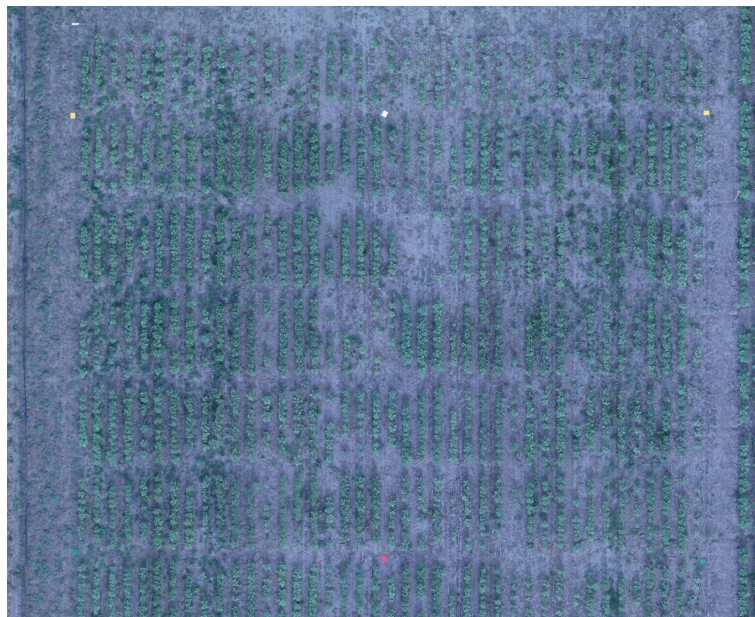


Figura 3.9: Imagen quemu2sr.jpg

3.2. Conteo del INTA

Cada parcela se compone de 3 surcos, 2 laterales y uno central. El conteo del INTA contiene la cantidad de plantas por parcela, divididas en surcos laterales y en surco central.

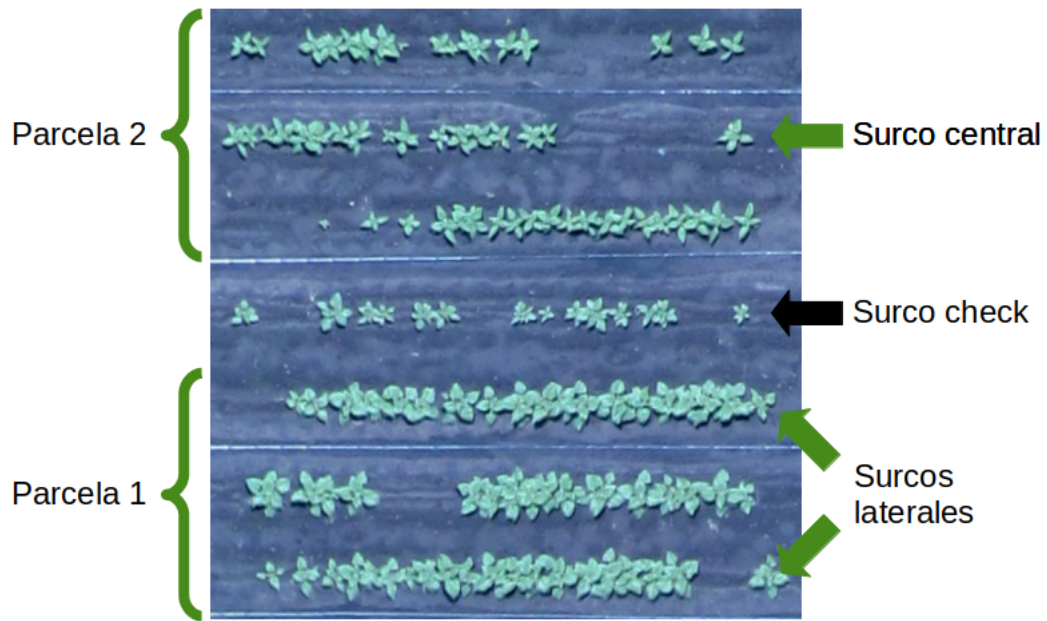


Figura 3.10: Partes de la parcela y surco check

A la vez cada parcela se encuentra separada por un surco de control o *check*. Dicho surco tiene como objetivo evitar el efecto de competencia entre las plantas de distintos genotipos. En la figura 3.10 se muestran los surcos laterales, el surco central, y el surco *check*.

Parcela 150	Parcela 131	Parcela 130	...	Parcela 30	Parcela 11	Parcela 10
hilera check						
Parcela 149	Parcela 132	Parcela 129	...	Parcela 29	Parcela 12	Parcela 9
hilera check						
Parcela 148	Parcela 133	Parcela 128	...	Parcela 28	Parcela 13	Parcela 8
hilera check						
⋮						
Parcela 143	Parcela 138	Parcela 123	...	Parcela 23	Parcela 18	Parcela 3
hilera check						
Parcela 142	Parcela 139	Parcela 122	...	Parcela 22	Parcela 19	Parcela 2
hilera check						
Parcela 141	Parcela 140	Parcela 121	...	Parcela 21	Parcela 20	Parcela 1

Figura 3.11: Enumeración de las parcelas del conteo

En la figura 3.11 se puede apreciar la forma en que se enumera cada parcela y como están separadas entre sí de forma vertical por una hilera check (compuesta por surcos



Figura 3.12: Imagen camet1x.tif obtenida a partir de la imagen camet1.tif

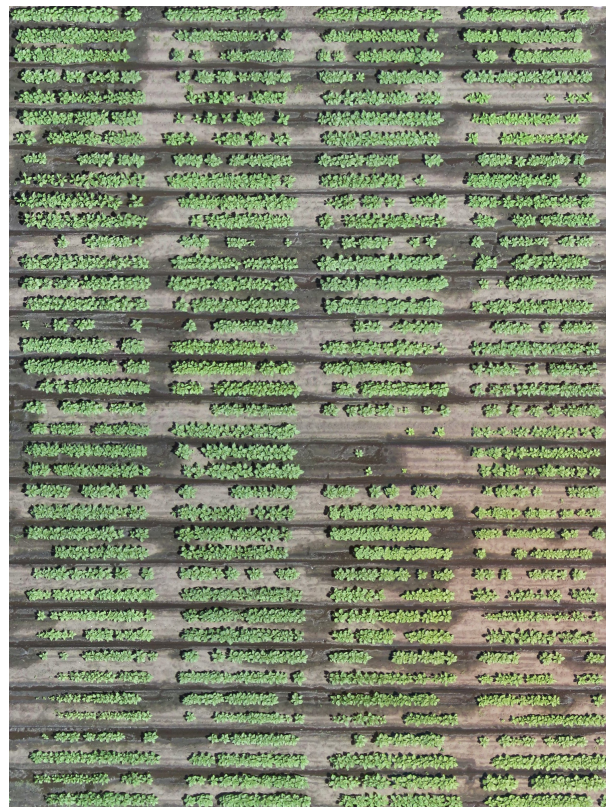


Figura 3.13: Imagen camet2crx.tif obtenida a partir de la imagen camet2cr.jpg

check).

Con el fin de aislar la región del conteo en la imagen camet1.tif, se rotó la imagen en sentido antihorario a partir del ángulo obtenido con el algoritmo que se describirá en la sección 4.2 para que las hileras queden orientadas horizontalmente y luego se recortó para que contenga solamente, las 150 parcelas y las hileras *check* que intercalan dichas parcelas,

en concordancia con la figura 3.11. Ésta nueva imagen que se muestra en la figura 3.12 tiene un tamaño de 11062×4000 y la llamaremos *camet1x.tif*.

De la misma manera se rotó y recortó la imagen *camet2cr.jpg* para aislar las parcelas contenidas en el conteo del INTA, dando como resultado la imagen que llamaremos *camet2crx.tif*, que tiene un tamaño de 2501×3322 , y que se muestra en la figura 3.13. Las mismas corresponden a las parcelas número 1 a 40 del conteo.

3.3. Píxeles por metro

Con el software *QLandkarte GT*¹ se extrajo la cantidad de PPM (Píxeles Por Metro) de la imagen *camet1.tif*, la única imagen que contiene información de georreferenciación. El valor obtenido fue de 143 PPM.

Todas las imágenes del cultivo A fueron capturadas a una altura de 30m, mientras que las del cultivo B fueron capturadas a una altura de 20m. En todas las imágenes se encuentran marcadores con forma rectangular, utilizados para generar los ortomosaicos, que tienen un tamaño de 21cm de alto y 25cm de ancho.

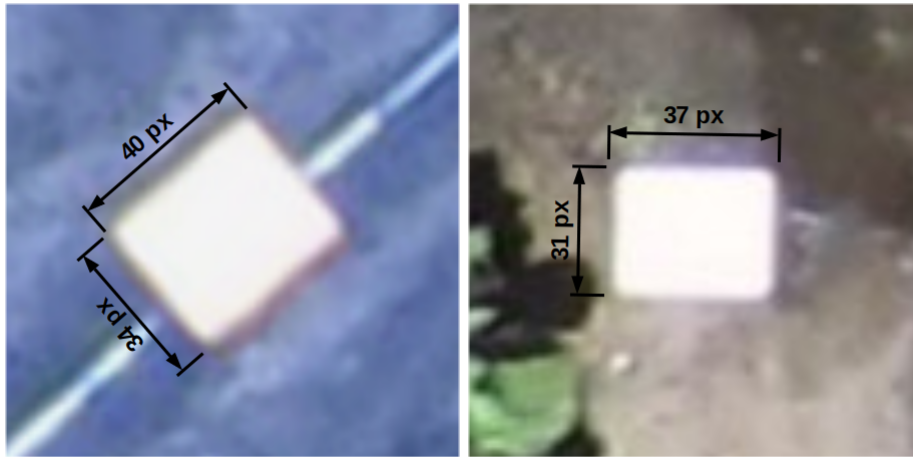
A partir del tamaño en píxeles de dichos marcadores se determinó que todas las imágenes del cultivo A tienen 143 PPM, excepto la imagen *camet2cr.jpg* que fue escalada.

Para determinar los PPM de la imagen *camet2cr.jpg*, se midieron el alto y el ancho en píxeles de uno de los marcadores de la imagen *camet1.tif* y de uno de los marcadores de la imagen *camet2cr.jpg*. En las figuras 3.14a y 3.14b se muestran éstas mediciones. Como los marcadores de ambas imágenes tienen el mismo tamaño, se calcularon proporcionalmente los PPM de cada dimensión del marcador en *camet2cr.jpg* y se promediaron ambos valores. Para el alto del marcador tenemos $\frac{31}{34} \times 143 = 130.38$ y para el ancho tenemos $\frac{37}{40} \times 143 = 132.28$, que en promedio nos da 131 PPM.

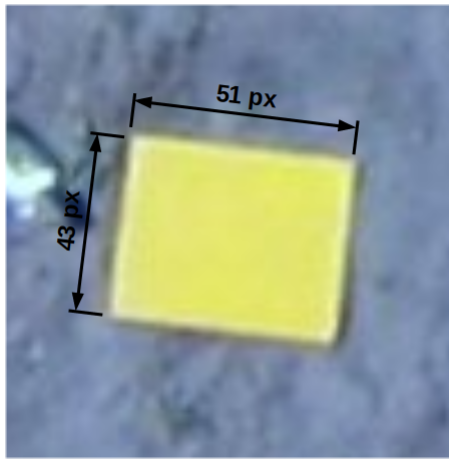
De la misma manera se calcularon los PPM para las imágenes del cultivo B. Dado que todas tienen la misma cantidad de PPM, ya que los marcadores tienen el mismo tamaño en píxeles, se realizó el cálculo sólo sobre la imagen *quemulsr.jpg*. En la figura 3.14c se muestra el tamaño del marcador en la imagen *quemulsr.jpg*. Para el alto del marcador tenemos $\frac{43}{34} \times 143 = 180.85$ y para el ancho tenemos $\frac{51}{40} \times 143 = 182.33$, que en promedio nos da 182 PPM.

En la tabla 3.2 se resumen los PPM para cada imagen.

¹<http://www.qlandkarte.org>



(a) Marcador en camet1.tif (cultivo A) (b) Marcador en camet2cr.jpg (cultivo A)



(c) Marcador en quemulsr.jpg (cultivo B)

Figura 3.14: Marcadores en cultivos A y B

Imagen	PPM
camet1.tif	143
camet1cr.jpg	143
camet1sr.jpg	143
quemulcr.jpg	182
quemulsr.jpg	182
camet2cr.jpg	131
camet2sr.jpg	143
quemu2cr.jpg	182
quemu2sr.jpg	182

Tabla 3.2: PPM de las imágenes

3.4. Parcelas para entrenamiento y testeo

A partir del conteo manual realizado por el INTA de las 150 parcelas sobre la imagen *camet.tif*, dividimos las parcelas del conteo de forma aleatoria en parcelas para entrenamiento y para testeo. Usaremos el 90 %, 135 parcelas, para entrenamiento, y el 10 %, 15 parcelas, para testeo. En la figura 3.15 se muestra la división por parcela.

150	131	130	111	110	91	90	71	70	51	50	31	30	11	10
149	132	129	112	109	92	89	72	69	52	49	32	29	12	9
148	133	128	113	108	93	88	73	68	53	48	33	28	13	8
147	134	127	114	107	94	87	74	67	54	47	34	27	14	7
146	135	126	115	106	95	86	75	66	55	46	35	26	15	6
145	136	125	116	105	96	85	76	65	56	45	36	25	16	5
144	137	124	117	104	97	84	77	64	57	44	37	24	17	4
143	138	123	118	103	98	83	78	63	58	43	38	23	18	3
142	139	122	119	102	99	82	79	62	59	42	39	22	19	2
141	140	121	120	101	100	81	80	61	60	41	40	21	20	1

Figura 3.15: Parcelas para entrenamiento en verde y para testeo en amarillo

Las 40 parcelas de la imagen *camet2crx.tif* se utilizarán también para testeo.

3.5. Conjuntos de coordenadas de plantas y de parcelas

Para éste trabajo fue necesario marcar manualmente las coordenadas de cada planta contenida tanto en las parcelas del conteo como también en las hileras *check* en la imagen *camet1x.tif*. En la imagen de la figura 3.16 se muestra el marcado de plantas con círculos azules.

A partir de las coordenadas marcadas, se crearon 3 conjuntos de coordenadas de plantas como se describe a continuación:

- *camet1x-completo*: Conteniendo todas las coordenadas de las plantas, incluyendo las de las hileras *check*.
- *camet1x-entrenamiento*: Conteniendo todas las coordenadas de las plantas correspondientes a las parcelas de entrenamiento.
- *camet1x-testeo*: Conteniendo todas las coordenadas de las plantas correspondientes a las parcelas de testeo.

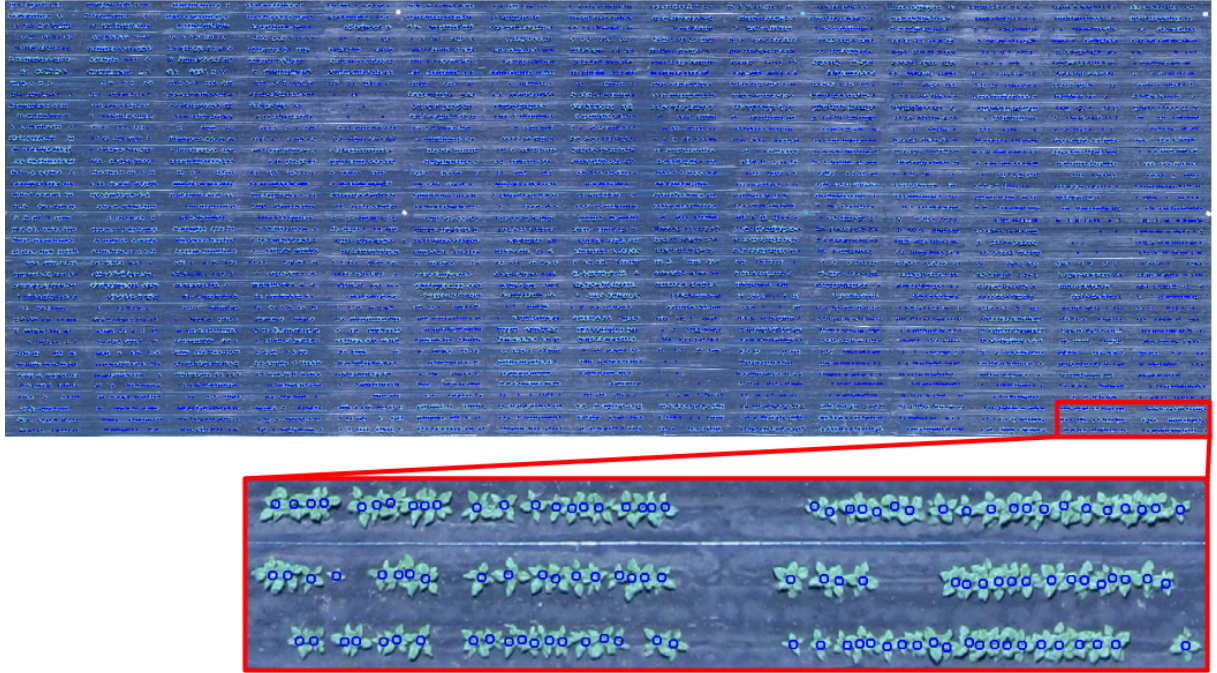


Figura 3.16: Marcado de plantas con círculos azules sobre la imagen camet1x.tif

El conjunto camet1x-completo contiene las coordenadas de 9063 plantas, camet1x-entrenamiento contiene las coordenadas de 6673 plantas, mientras que camet1x-testeo contiene las coordenadas de 685 plantas.

También fue necesario delimitar las parcelas que se utilizarán para el testeo. Para ello se seleccionaron manualmente sobre la imagen las parcelas y se guardaron las coordenadas x e y correspondientes a la esquina superior-izquierda, el ancho, y el alto de la parcela. Dando a lugar los conjuntos de coordenadas siguientes:

- camet1x-parcelas-testeo: Conteniendo las coordenadas de las 15 parcelas de testeo de la imagen camet1x.tif.
- camet2crx-parcelas: Conteniendo las coordenadas de las 40 parcelas de la imagen camet2crx.tif

Capítulo 4

Metodología

4.1. Segmentación

El objetivo de la etapa de segmentación es clasificar los píxeles (i, j) de la imagen I en dos clases. Una clase correspondiente a las plantas y otra correspondiente al fondo, es decir, otros objetos que no son plantas tales como el suelo, el riego, etc. Esta clasificación da como resultado la máscara M descrita en la ecuación 4.1.

$$M(i, j) = \begin{cases} 255 & \text{si } I(i, j) \in \text{plantas} \\ 0 & \text{sino} \end{cases} \quad (4.1)$$

4.1.1. Detección de vegetación

Dentro de la literatura existen diferentes enfoques para la identificación y clasificación de cultivos de plantas a partir de imágenes aéreas. En particular, los índices de vegetación han sido ampliamente utilizados.

Los índices de vegetación están diseñados para maximizar la sensibilidad a las características de la vegetación y minimizar los factores de confusión como la reflectancia de fondo del suelo, los efectos direccionales o atmosféricos. [14]

Algunos índices de vegetación relevantes basados en el espectro RGB visible utilizados en agricultura y que fueron probados para el presente trabajo son: NDI (*Normalized Difference Index*) [15], ExG (*Excess Green*) [16], CIVE (*Colour Index of Vegetation Extraction*) [17], ExR (*Excess Red*) [18], ExGR (*Excess Green minus Excess Red Index*) [19], VEG (*Vegetative Index*) [20] y GLI (*Green Leaf Index*) [21]

Woebbecke *et al.* [16] testearon cuatro índices de vegetación utilizando coordenadas

cromáticas y valores RGB normalizados. Además testearon un *hue* modificado del espacio de color HSI. Los autores encontraron que los mejores resultados de contraste entre las superficies vegetales y no vegetales se obtenían con el índice referido como ExG (ecuación 4.2).

$$ExG = 2g - r - b \quad (4.2)$$

donde r , g y b son las coordenadas cromáticas:

$$r = \frac{R_n}{R_n + G_n + B_n}, \quad g = \frac{G_n}{R_n + G_n + B_n}, \quad b = \frac{B_n}{R_n + G_n + B_n}$$

y R_n , G_n y B_n son los valores RGB normalizados:

$$R_n = \frac{R}{R_m}, \quad G_n = \frac{G}{G_m}, \quad B_n = \frac{B}{B_m}$$

con $R_m = G_m = B_m = 255$ que son el valor máximo para cada color primario en una imagen de 24 bits (3 canales de 8 bits cada uno).

Algunos autores, sin embargo, no utilizan coordenadas cromáticas ni valores RGB normalizados para el índice ExG (Meyer y Neto [19], Zhao *et al.* [7], Rico-Fernández *et al.* [22], Diao *et al.* [23]), en su lugar computan el índice directamente con los valores de los canales RGB.

En los trabajos de Rico-Fernández *et al.* [22] y Diao *et al.* [23] los valores obtenidos al calcular el índice ExG que son menores a 0 se reemplazan por 0 y los que son mayores a 255 se reemplazan por 255. De ésta manera los valores permanecen dentro del rango 0-255. La ecuación 4.3 describe ésta forma de calcular ExG expresada con subíndices y nos referimos a ella con el sufijo RGB para diferenciarla de la original.

$$ExG_{RGB}(i, j) = \begin{cases} 0 & si \quad ExG'_{RGB}(i, j) < 0 \\ 255 & si \quad ExG'_{RGB}(i, j) > 255 \\ ExG'_{RGB} & sino \end{cases} \quad (4.3)$$

donde

$$ExG'_{RGB} = 2G - R - B$$

El índice ExG_{RGB} de la ecuación 4.3 discriminó con mas precisión entre plantas y otras regiones que los otros índices probados incluyendo el método ExG original propuesto por Woebbecke *et al.* [16]. Por lo tanto, se utilizó como método para extraer de la imagen la información de las plantas.

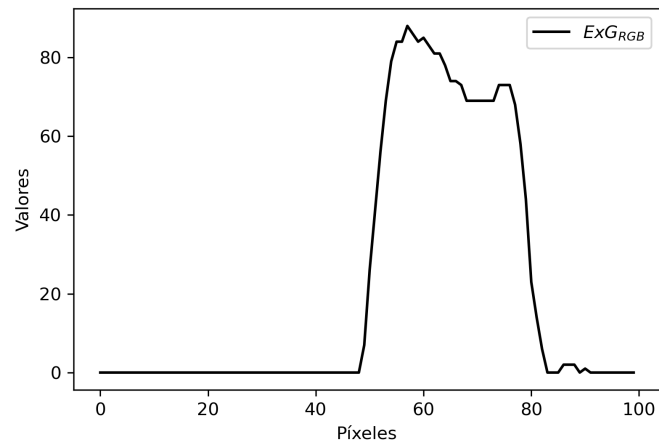
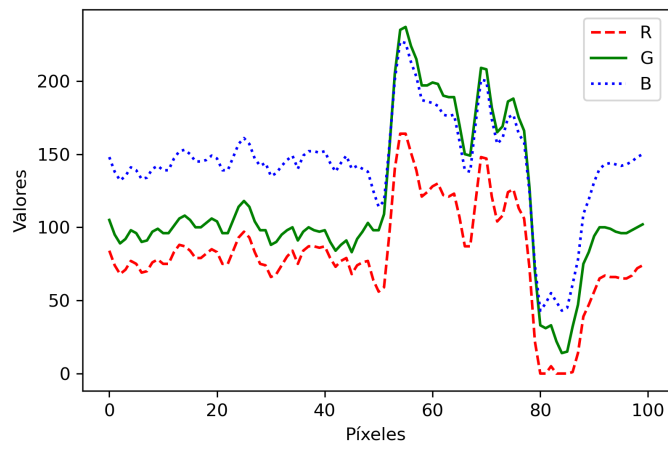
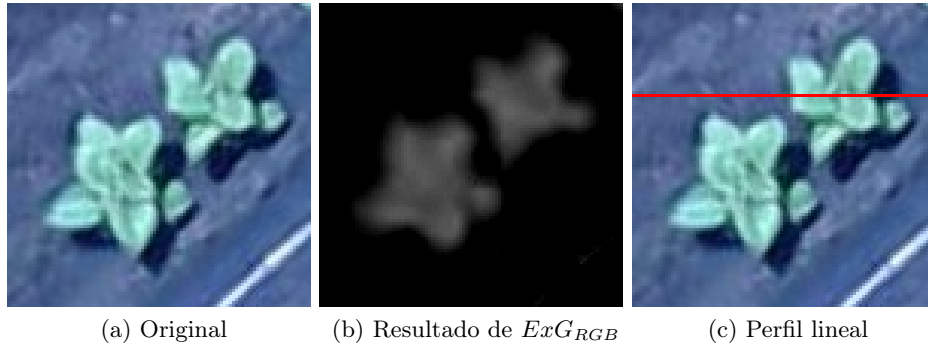


Figura 4.1: Perfil lineal para R, G, B y ExG_{RGB}

En la imagen de la figura 4.1b se muestra el resultado de computar el índice ExG_{RGB}

a la imagen original de la figura 4.1a, el resultado es una imagen en escala de grises donde se contrastan los objetos de color verde (plantas) del fondo. En la imagen de la figura 4.1c se muestra la selección en rojo de un perfil lineal. En la figura 4.1d se observa el perfil lineal para los canales R, G, y B de la imagen original y en la figura 4.1e el perfil lineal de la imagen resultante de computarse el índice ExG_{RGB} .

Una vez que se obtiene la imagen en escala de grises luego de computar el índice ExG_{RGB} , se requiere un paso crítico para seleccionar el valor de umbral para binarizar la imagen resultante. La técnica de umbralización automática más utilizada en combinación con los índices de vegetación es el método de Otsu [24], el cual se utiliza en éste trabajo en combinación con la ecuación 4.3

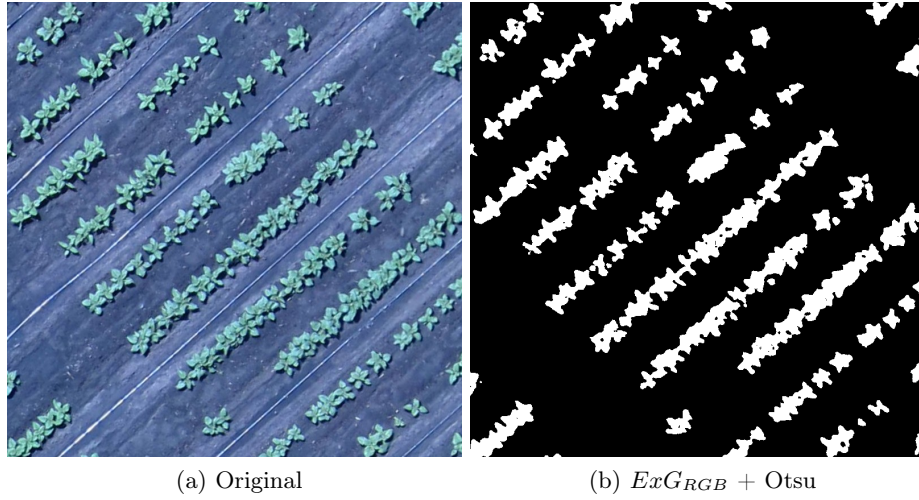


Figura 4.2: Resultado de computar ExG_{RGB} y aplicar umbralización Otsu

En la imagen de la figura 4.2b se puede apreciar la viabilidad de computar el índice ExG_{RGB} junto con la técnica de umbralización de Otsu a la imagen de la figura 4.2a.

4.1.2. Mejora de la máscara

Para mejorar la máscara obtenida, se lleva a cabo una operación morfológica binaria de apertura [25] con un elemento estructurante circular de radio igual a 3 píxeles para suavizar los contornos, desconectar componentes poco conectadas y eliminar protuberancias pequeñas.

Para introducir la definición de apertura morfológica es necesario definir antes las operaciones de erosión y dilatación [25].

Dados A y B conjuntos en $\mathbb{Z}^2$, la erosión de A por B, notada $A \ominus B$, se define como

$$A \ominus B = \{z \mid (B)_z \subseteq A\} \quad (4.4)$$

donde

$$(B)_z = \{c \mid c = b + z, b \in B\}$$

Es decir, la erosión de A por B, son los z tales que B desplazado por z está contenido en A.

Y la dilatación de A por B, notada $A \oplus B$, se define como

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\} \quad (4.5)$$

donde

$$\hat{B} = \{w \mid w = -b, b \in B\}$$

Es decir, la dilatación de A por B, son los z tales que la reflexión de B, desplazada por z, se intersectan en al menos un elemento.

La apertura morfológica, notada $A \circ B$ se define como

$$A \circ B = (A \ominus B) \oplus B \quad (4.6)$$

En otras palabras, la apertura de A por B es el resultado de aplicar la erosión y luego la dilatación usando el mismo elemento estructurante B.

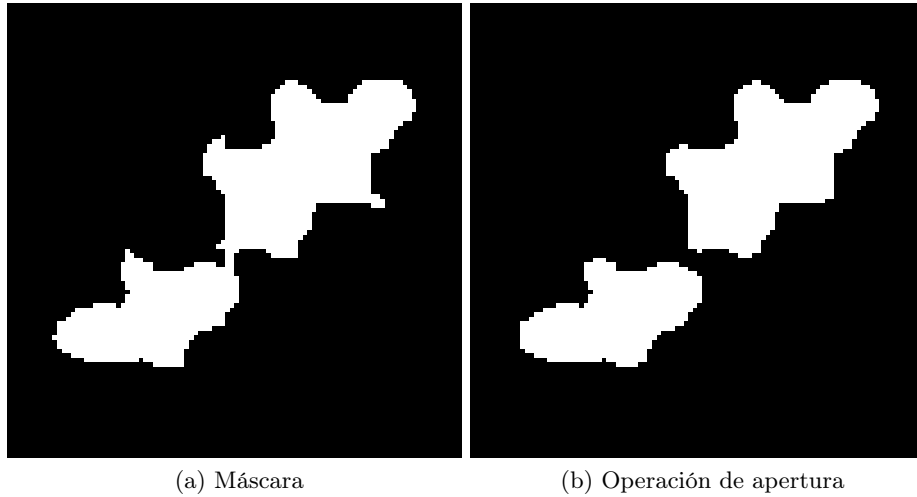


Figura 4.3: Operación de apertura con elemento estructurante circular de radio 3

En la figura 4.3b se muestra el resultado de aplicar la operación de apertura a la máscara de la figura 4.3a. Nótese como los bordes se suavizan y se eliminan protuberancias, a la vez que se separan los objetos poco conectados.

Otra mejora de la máscara que se lleva a cabo es el rellenado de huecos, para ello se extraen los contornos externos de todos los objetos y se etiquetan los píxeles internos de los mismos como plantas.

Por último, se remueven los objetos pequeños, para ello se calcula el área [25] de cada objeto, es decir, la cantidad de píxeles que componen a cada uno, y para los cuales el área es menor al umbral U_{area} , sus píxeles son marcados como no-plantas, es decir, como fondo.

Si el umbral es muy pequeño se tomarían objetos de fondo como plantas, y si es muy grande se estarían eliminando plantas pequeñas de la máscara. Uno de los problemas a los que se hizo frente en éste trabajo es que hay plantas que tienen un tamaño menor a una sola hoja de una planta grande.

El umbral está dado por la fórmula

$$U_{area} = \left(\frac{PPM}{C}\right)^2 \quad (4.7)$$

donde PPM es la cantidad de píxeles por metro en la imagen y $C = 20$ es una constante que se seleccionó a partir de la observación de los tamaños de los objetos de fondo que se segmentaron como plantas, como el caso de hojas muertas, y el tamaño de las plantas más pequeñas.

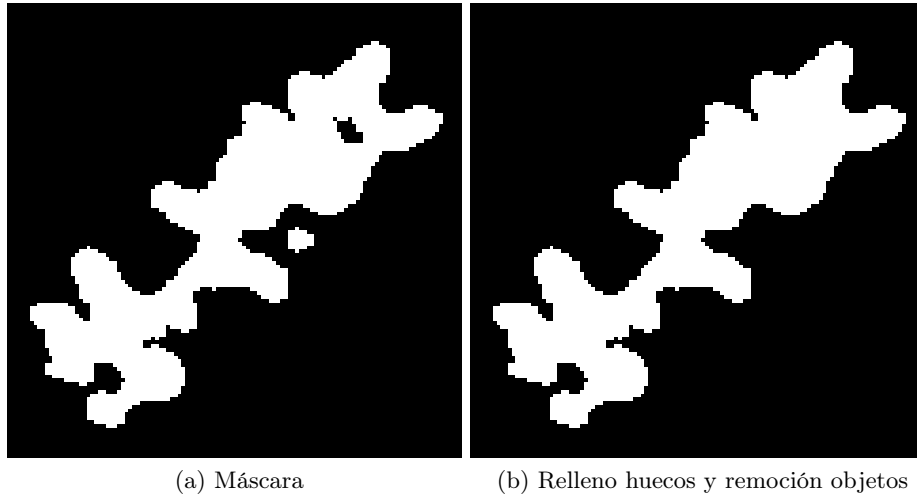


Figura 4.4: Rellenado de huecos y remoción de objetos con área menor al umbral

En la imagen de la figura 4.4b se muestra el resultado de rellenar los huecos y remover los objetos pequeños de la máscara de la figura 4.4a.

4.2. Detección de la orientación de las hileras

En esta sección se propone utilizar la imagen de autocorrelación de la máscara obtenida anteriormente para determinar la orientación de las hileras de cultivo respecto del eje horizontal de la imagen. De manera tal que la imagen pueda rotarse posteriormente para que las hileras queden orientadas horizontalmente, haciéndose más simple su visualización y simplificando el algoritmo para el conteo de las plantas.

La autocorrelación se utiliza para combinar todas las partes de la imagen, con el fin de encontrar estructuras repetitivas. Para imágenes en las que existe una disposición más o menos regular de patrones, la imagen de autocorrelación encuentra este patrón y también facilita su medición. [26]

En su tesis doctoral, Larese [27] propuso utilizar la imagen de autocorrelación sobre las imágenes de microarreglos de ADNc para obtener los vectores de desplazamiento y a partir de ellos el ángulo de rotación.

La correlación cruzada [25] de un kernel w con una imagen f , notada $w \star f$, está dada por la ecuación

$$(w \star f)(x, y) = \sum_s \sum_t w(s, t) f(x + s, y + t) \quad (4.8)$$

donde los límites de las sumatorias se toman sobre la región compartida por w y f .

La correlación tiene sus valores más altos en las regiones donde f y w son iguales o casi iguales. En otras palabras, busca ubicaciones donde w coincide con una región de f .

La autocorrelación es un caso particular de la correlación cruzada, donde f y w son iguales. Si f es una imagen de dimensión $H \times W$, su autocorrelación está dada por la ecuación

$$R(x, y) = (f \star f)(x, y) = \sum_{s=0}^{H-1} \sum_{t=0}^{W-1} f(s, t) f(x + s, y + t) \quad (4.9)$$

donde la dimensión de la imagen resultante es de $2H - 1 \times 2W - 1$.

Los perfiles de intensidad de la región circundante al pico de la autocorrelación, a lo largo de sus ejes mayor y menor, miden cuantitativamente la orientación predominante [26]

En la imagen de la figura 4.5a vemos la máscara de la imagen de la figura 4.2b con las mejoras propuestas en la sección 4.1.2, y en la figura 4.5b se muestra el resultado de aplicar la autocorrelación a dicha máscara.

Para obtener la orientación primero se binariza la imagen de autocorrelación, para ello

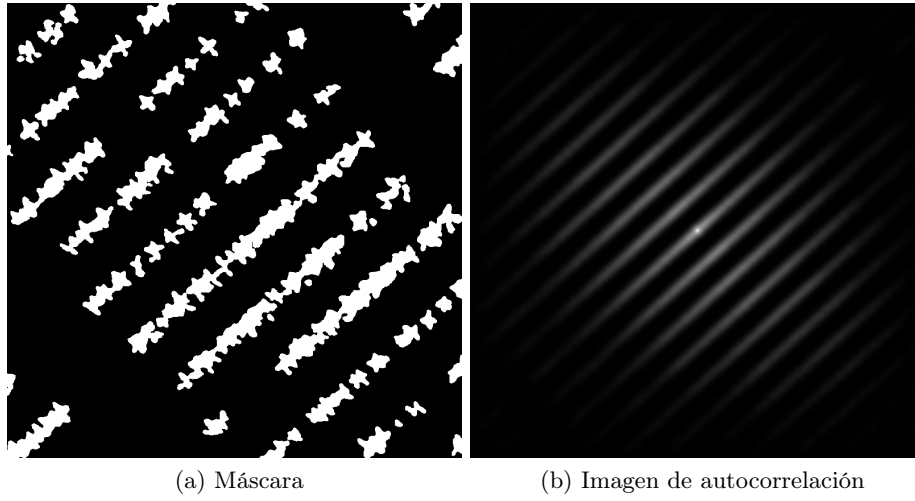


Figura 4.5: Imagen de autocorrelaci3n de la m3scara

aplicamos el m3todo de Otsu ya utilizado en la secci3n 4.1.1. Luego se extrae el contorno del objeto central de la imagen binarizada y se calcula la orientaci3n del mismo.

Dado que para obtener la orientaci3n de la imagen no es necesario calcular la autocorrelaci3n de toda la m3scara, especialmente si esta es muy grande, se opt3 por aplicarla solo a una porci3n central de la misma, de manera que no supere los 4000×4000 p3xeles de tama1o. Si alguna de las dimensiones de la m3scara supera los 4000 p3xeles entonces se recorta para no superar ese tama1o. De 3sta manera el tiempo de computo de la autocorrelaci3n es menor y tambi3n se obtiene un resultado m3s preciso de la orientaci3n, ya que las im3genes pueden contener otra vegetaci3n fuera de la regi3n central.

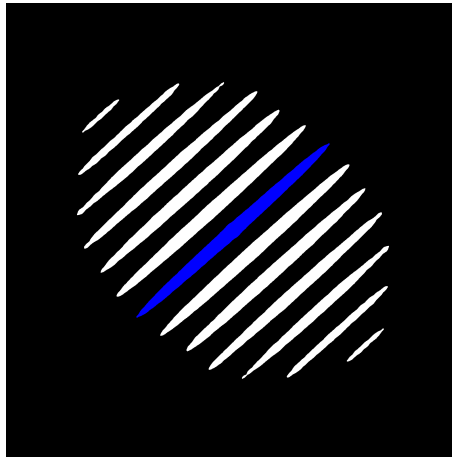


Figura 4.6: Autocorrelaci3n binarizada con el objeto central en azul

En la imagen de la figura 4.6 se muestra la imagen de la figura 4.5b ya binarizada con el objeto central marcado en azul.

Para calcular la orientación del objeto central se aproxima el contorno del mismo con una elipse que tiene los mismos segundos momentos centrales que el contorno, es decir, que sus ejes están alineados con los ejes principales del contorno del objeto central.

Como el contorno está compuesto por dos dimensiones, tenemos que considerar la varianza de cada variable, así como también la covarianza entre ellas. Éstas componentes conforman la matriz de covarianza C [25] que se estima utilizando la ecuación

$$C = \frac{1}{K-1} \sum_{k=1}^K (z_k - \bar{z})(z_k - \bar{z})^T \quad (4.10)$$

donde los z_k son vectores con las coordenadas x e y de cada punto del contorno, K es la cantidad total de puntos en el mismo, y $\bar{z}$ es el vector promedio

$$\bar{z} = \frac{1}{K} \sum_{k=1}^K z_k$$

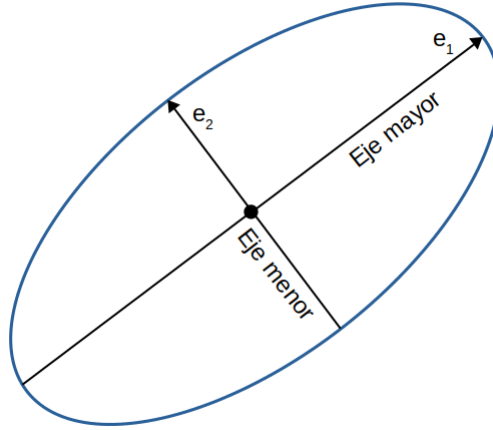


Figura 4.7: Dirección de los ejes mayor y menor de una elipse dados por los vectores e_1 y e_2

La orientación de los ejes principales de la elipse están dados por los autovectores de la matriz de covarianza C . En la figura 4.7 podemos ver un ejemplo ilustrativo de una elipse con sus ejes mayor y menor donde su direcciones están dadas por los vectores e_1 y e_2 respectivamente.

La dirección de máxima varianza de los puntos que conforman el contorno del objeto central está dada por el autovector asociado al mayor autovalor (autovector dominante) de la matriz de covarianza C . Sea $e_1 = (x, y)$ dicho autovector, el ángulo de orientación del objeto central y por tanto la orientación de la imagen de hileras de cultivo se obtiene

calculando el ángulo

$$\theta = \text{atan2}(y, x) \quad (4.11)$$

donde la función atan2 equivale a calcular $\text{atan}(y/x)$ y corregir el ángulo al cuadrante correcto.

En este caso los autovectores de C son $e_1 = \langle 0.74, 0.67 \rangle$ y $e_2 = \langle -0.67, 0.74 \rangle$, y los autovalores respectivos son $\lambda_1 = 71719.46$ y $\lambda_2 = 371.09$. Como el autovector asociado al mayor autovalor es e_1 , se tiene que la orientación de la imagen es $\theta = 42.10^\circ$



Figura 4.8: Dirección de la orientación de las hileras

En la imagen de la figura 4.8 se muestra la orientación del vector e_1 sobre la imagen de la figura 4.2a.

4.3. Enfoque para el conteo de plantas de girasol

En éste estudio se analizó sobre la imagen de plantas de girasol la viabilidad del enfoque propuesto por Wang *et al.* [6] para el conteo de palmeras de aceite utilizando el descriptor de característica HOG (histograma de gradientes orientados) propuesto por Dalal y Triggs [28] inicialmente para la detección de peatones. También se analizó la viabilidad del enfoque propuesto por Koh *et al.* [9] para el conteo de plántulas de cártamo utilizando la técnica de *template matching* sobre el canal G de la imagen RGB

Para el primer enfoque se convirtió la imagen de la figura 4.9a a escala de grises y se le aplicó la máscara a la misma obteniendo el resultado que se muestra en la imagen de la figura 4.9b.

Luego se calculó y visualizó el descriptor de HOG utilizando diferentes combinaciones de parámetros. Los mismos se detallan en la tabla 4.1. En ninguno de los casos se observó



(a) Original



(b) Escala de grises + máscara



(c) HOG

Figura 4.9: Visualización de los histogramas de gradientes orientados

algún patrón que se repitiera para cada planta como en el caso de las palmeras aceiteras donde se observa un patrón obvio de círculos concéntricos.

Orientaciones	Píxeles por celda	Celdas por bloque
9	4×4	2×2
9	4×4	4×4
9	8×8	2×2
9	8×8	4×4
9	12×12	2×2
9	12×12	4×4
9	16×16	2×2
9	16×16	4×4

Tabla 4.1: Parámetros utilizados para el descriptor HOG

En la imagen de la figura 4.9c se muestra el resultado visual del descriptor HOG sobre una imagen de 239×64 píxeles utilizando 9 orientaciones, celdas de 8×8 píxeles y bloques de 2×2 celdas, que son los parámetros utilizados por Wang *et al.* [6].

Dada a la variedad de formas y posturas que adoptan las hojas de las plantas de girasol, sumado al solapamiento de las mismas, los histogramas de gradiente orientados muestran una distribución muy irregular y por tanto se concluye que éste enfoque no es viable para el conteo de las mismas.

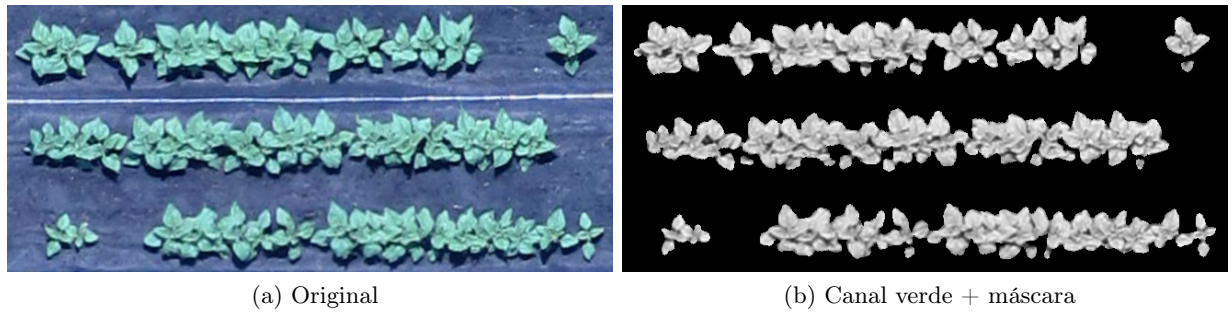


Figura 4.10: Aplicación de la máscara sobre el canal G

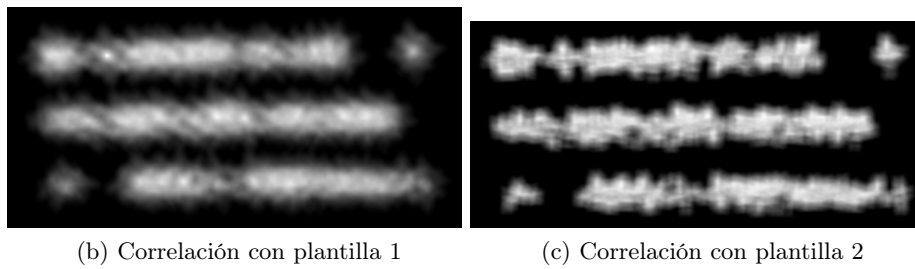
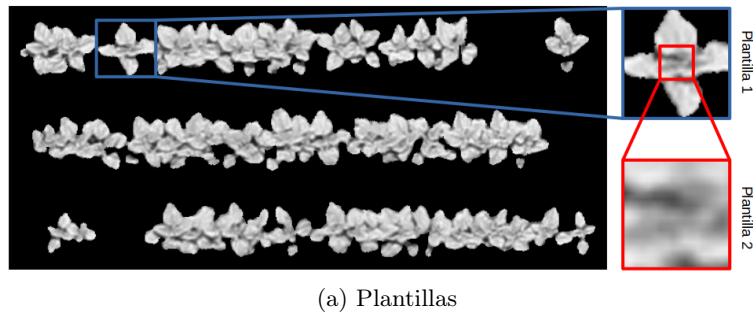


Figura 4.11: Correlación cruzada con dos plantillas

Para el segundo enfoque se extrajo el canal G de la imagen de la figura 4.10a y se le aplicó la máscara al mismo, obteniendo el resultado que se muestra en la imagen de la figura 4.10b

Luego, de la imagen obtenida, se extrajeron porciones correspondientes a plantas completas y a centros de las mismas, que fueron utilizadas como plantillas para calcular la correlación cruzada con la imagen de la figura 4.10b. En la figura 4.11a se muestra la extracción de dos plantillas que se utilizaron para evaluar éste enfoque.

Los resultados obtenidos se muestran en 4.11b para la correlación con la plantilla 1 y en 4.11c para la correlación con la plantilla 2. Para el primer caso se observan pocos picos máximos de correlación con claridad y para el segundo caso ninguno. Esto se debe a que las plantas no solo varían en forma, posturas y solapamiento como ya mencionamos sino que la textura de los centros de las plantas no difieren lo suficiente de las texturas de las hojas debido a la perspectiva y a la iluminación. Por tanto se concluye que éste enfoque

tampoco es viable para el conteo.

Dada la complejidad de identificar las plantas de forma individual debido a los problemas ya descritos, se optó por un enfoque de análisis del tamaño y la forma de los objetos de plantas de girasol como propusieron Zhao *et al.* [7] para el conteo de plantas de colza y Guo *et al.* [8] para la detección del número de panojas de sorgo. La idea de éste enfoque es tratar a los objetos segmentados como grupos de plantas y estimar la cantidad correspondientes de plantas en cada grupo, es decir, para cada objeto, mediante regresión a partir de características de tamaño y forma de los mismos.

4.4. Detección de hileras y etiquetado de objetos

Lottes *et al.* [29] propusieron utilizar la transformada de Hough sobre la máscara de la imagen de plantas de remolacha azucarera para detectar las líneas de las hileras de cultivos, mientras que Ji y Qi [30] utilizaron la transformada aleatoria de Hough aplicada a los centroides de los objetos extraídos de la máscara de la imagen de plantas de soja para detectar las líneas de las mismas.

En el trabajo de Zhao *et al.* [7] los objetos de colza segmentados se convirtieron en puntos y se etiquetaron de acuerdo con el número de hilera y el eje x . Luego las líneas de las hileras se generaron conectando los puntos con las mismas etiquetas a lo largo del eje y .

En éste trabajo se propone un algoritmo para detectar las hileras, y por ende las líneas de las hileras, para de ésta manera estimar la cantidad de plantas sólo en los objetos que pertenecen a las mismas, descartando el resto. Para llevar a cabo la detección se hace uso de perfiles acumulados aplicados a la máscara, para calcular la cantidad de hileras y los centros aproximados de las mismas, en combinación con un enfoque basado en objetos como propusieron Zhao *et al.* [7]

Definimos la función de perfiles acumulados (horizontales) p como

$$p(y) = \sum_{x=0}^{W-1} (M(x, y)) \quad (4.12)$$

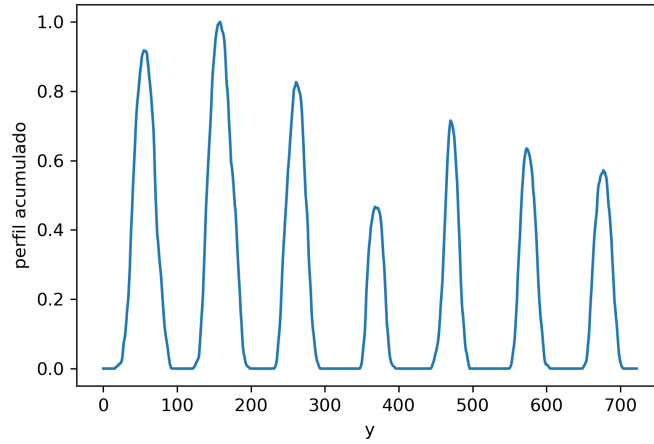
con $0 \leq y \leq H$, donde H y W son el alto y el ancho de la máscara M respectivamente.

Una vez calculada la función p , la normalizamos entre 0 y 1 dividiendo por P_{max} que corresponde con el valor máximo de p , y obtenemos la función

$$p_{norm}(y) = p(y)/P_{max} \quad (4.13)$$



(a) Máscara



(b) Perfiles acumulados

Figura 4.12: Función de perfiles acumulados aplicada a la máscara

En la figura 4.12b se muestra la función de perfiles acumulados p_{norm} . Se pueden ver 7 picos correspondientes a las 7 hileras de la máscara.

Entonces para contabilizar dichos picos se buscan los máximos locales de la función p_{norm} que tengan una prominencia de al menos 0.1 de altura y una distancia horizontal mínima entre picos de $\frac{PPM}{10}$, el equivalente a 10cm. Nos referiremos a los valores de éstos máximos locales como centros de hileras y CH_i al centro de hilera i , con $1 \leq i \leq N$ donde 1 es el primer centro de hilera y N la cantidad total.

También, junto a los máximos locales de la función p_{norm} se calculan los anchos de los picos en la mitad de sus prominencias y nos referiremos a éstos valores como AH_i para el ancho de hilera i , de similar forma a como lo hacemos con los centros de hileras.

En la figura 4.13 se muestra la función de perfiles acumulados p_{norm} con los máximos

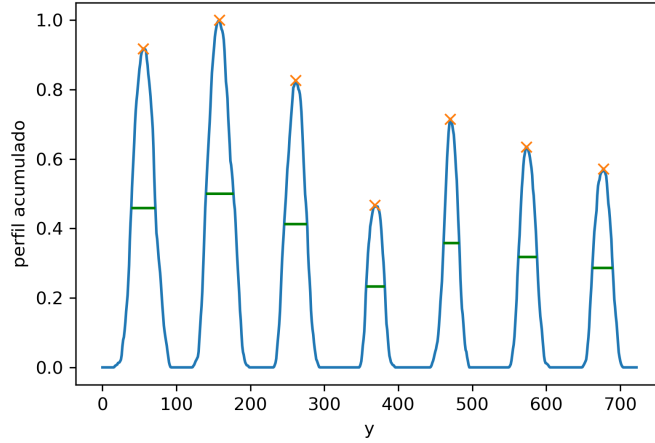


Figura 4.13: Máximos locales y anchos de los picos

locales de la máscara de la figura 4.12a marcados con una equis y los anchos de los picos en la mitad de sus prominencias con una línea en verde.

Una vez detectada la cantidad de hileras, y calculados los valores CH_i y AH_i para cada hilera i , extraemos los objetos de la máscara. Nos referiremos más adelante al conjunto de objetos de la máscara como *OBJETOS*. También calculamos para cada uno de los mismos el centroide y el *bounding box*.

Para el cálculo del centroide [26] de un objeto utilizamos la ecuación

$$\begin{cases} \text{centroide}_x(\text{objeto}) = \frac{\sum_i x_i}{N} \\ \text{centroide}_y(\text{objeto}) = \frac{\sum_i y_i}{N} \end{cases} \quad (4.14)$$

donde x_i e y_i son las coordenadas de los píxeles del objeto, y N es la cantidad de píxeles (área) del mismo.

El *bounding box* [26] de un objeto es un rectángulo donde su esquina superior-izquierda se corresponde con el píxel de menor coordenadas y su esquina inferior-derecha se corresponde con el píxel con mayor coordenadas, en las direcciones horizontal y vertical. Notamos $BB_x(\text{objeto})$ a las coordenadas x del *bounding box* de un objeto.

A partir de los datos calculados hasta ahora, se aplican 3 pasos para la detección de las hileras, los cuales se describen a continuación.

El primer paso consiste en etiquetar con $HILERA_i$ para cada $i = 1...N$, los objetos de manera inicial, utilizando como umbral a AH_i con respecto al centro de hilera CH_i .

```

para cada  $i = 1...N$  hacer
|   para cada  $objeto \in OBJETOS$  hacer
|   |   si  $CH_i - AH_i \leq \text{centroide}_y(objeto) \leq CH_i + AH_i$  entonces
|   |   |   Agregar objeto a  $HILERA_i$ 
|   |   fin
|   fin
fin

```

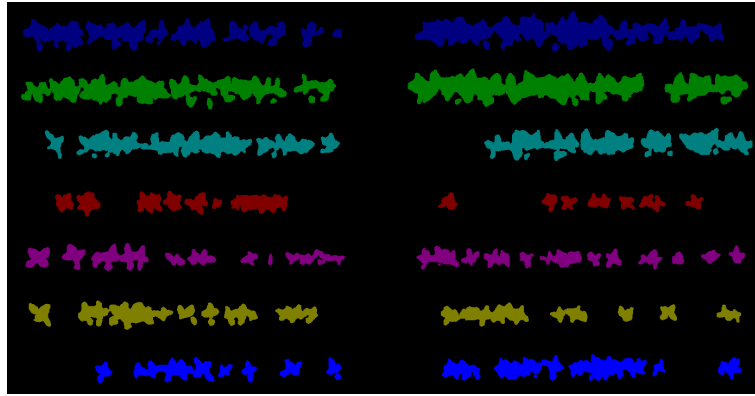


Figura 4.14: Etiquetado de los objetos según paso 1

En la imagen de la figura 4.14 se muestran los objetos con distintos colores según la hilera en la que se etiquetó.

El segundo paso consiste en descartar los objetos que no deberían ser parte de alguna hilera, éste es el caso de dos objetos que fueron etiquetados dentro de una misma hilera y uno de ellos contiene las coordenadas x de su *bounding box* completamente dentro de las coordenadas x del *bounding box* del otro objeto.

```

para cada  $i = 1...N$  hacer
|   para cada  $objeto_j, objeto_k \in HILERA_i, k \neq j$  hacer
|   |   si  $BB_x(objeto_j) \subseteq BB_x(objeto_k)$  entonces
|   |   |   Remover  $objeto_j$  de  $HILERA_i$ 
|   |   fin
|   fin
fin

```

Un ejemplo frecuente de éste caso se da cuando una planta no se segmenta correctamente debido a sombras que se producen en las hojas inferiores a causa de las superiores,

quedando una hoja desconectada del resto de la planta. Otro ejemplo frecuente es el de vegetación fuera de las hileras por un error en la máquina de siembra o algún tipo de maleza.

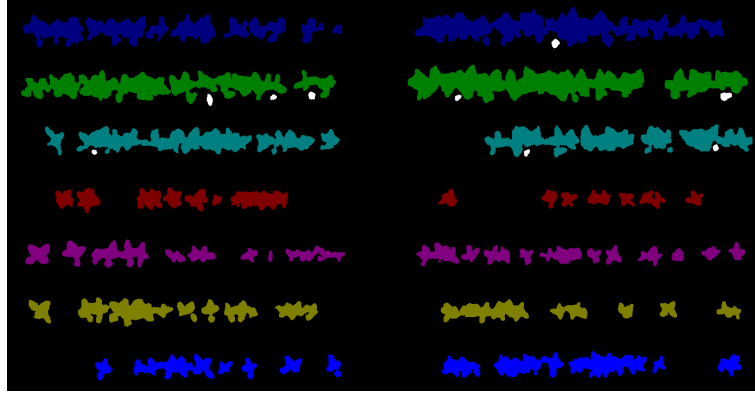


Figura 4.15: Desetiquetado de objetos según paso 2

En la imagen de la figura 4.15 se muestra el resultado del paso 2, donde los elementos sin etiquetar se aprecian en blanco.

El tercer paso consiste en refinar el cálculo de las hileras achicando el umbral de AH_i a $AH_i/2$ con respecto al centro de hilera CH_i para que un objeto pertenezca a la misma. Además, como las hileras no son completamente paralelas entre sí, sino que existe una leve desviación entre las mismas debido a la forma de siembra o a las características del terreno, los centros de las mismas se van ajustando a medida que los objetos se van etiquetando con $HILERA2_i$. Para el ajuste de los centros, los objetos se van etiquetando de izquierda a derecha y el valor del centro de hilera se va actualizando como el promedio de los valores $centroide_y$ de los últimos L objetos etiquetados, donde se toma $L = 3$.


```

para cada  $i = 1 \dots N$  hacer
  Ordenar  $HILERA_i$  de menor a mayor según el valor de  $centroide_x$ 
  para cada  $objeto \in HILERA_i$  hacer
    si  $|HILERA2_i| < L$  entonces
      | Agregar objeto a  $HILERA2_i$ 
    en otro caso
      |  $CH_i =$  Promediar el valor  $centroide_y$  de los últimos  $L$  objetos en
      |  $HILERA2_i$ 
      si  $CH_i - AH_i/2 \leq centroide_y(objeto) \leq CH_i + AH_i/2$  entonces
        | Agregar objeto a  $HILERA2_i$ 
      fin
    fin
  fin
fin

```

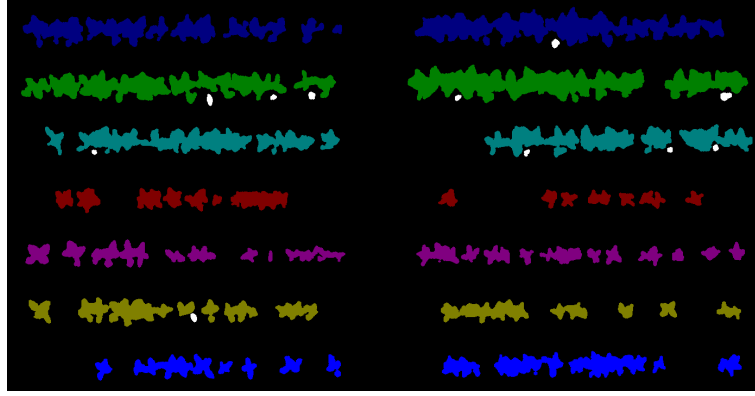


Figura 4.16: Desetiquetado de objetos según paso 3

En la imagen de la figura 4.16 se muestra el resultado del tercer paso, se puede apreciar que se desetiquetaron 2 objetos más respecto del paso 2.

Finalmente obtenemos los objetos que conforman cada una de las N hileras etiquetados con $HILERA2_i$ para cada $i = 1 \dots N$. En la imagen de la figura 4.17 se muestran las líneas de las hileras de la máscara de la figura 4.12a formadas a partir de los centroides de los objetos correspondientes a cada hilera.

4.5. Descriptores de características

En ésta sección se definen los descriptores de características de forma y tamaño [25, 26, 31] que se calculan para cada objeto y que son utilizados posteriormente para predecir la cantidad de plantas mediante regresión.

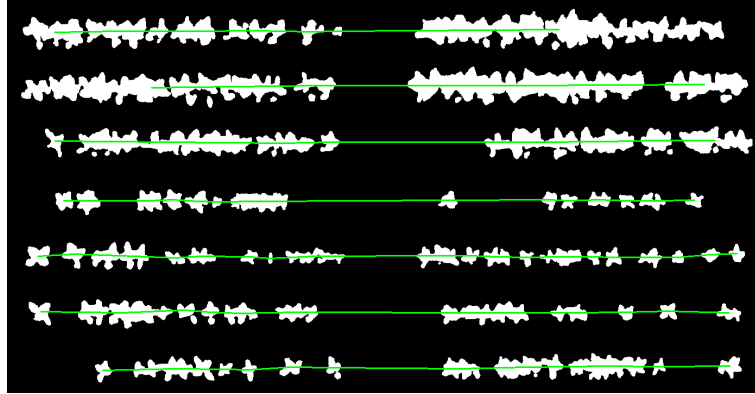


Figura 4.17: Líneas de las hileras

Nos referiremos con *AREA* al área del objeto como ya fue descrita en la sección 4.1.2 y con *PERIMETRO* al perímetro del objeto, es decir, la longitud del contorno del mismo, en píxeles.

Se utiliza la definición de *bounding box* de un objeto tal cual se hizo en la sección 4.4. Ésto se debe a que nos interesa que los ejes del *bounding box* estén alineados con los ejes de coordenadas, a diferencia del mínimo *bounding box* que es un rectángulo rotado.



(a) Objeto



(b) Cápsula convexa

Figura 4.18: Cápsula convexa de un objeto

Por otra parte definimos a la cápsula convexa de un objeto como la menor región convexa que incluye a éste. En la imagen de la figura 4.18b se muestra un ejemplo de la cápsula convexa del objeto de la figura 4.18a.

Los primeros cuatro descriptores se expresan en unidad métrica en base a los píxeles por metro (*PPM*) de la imagen, mientras que el resto son valores escalares.

- Área en metros cuadrados

$$AREA_M = \frac{AREA}{PPM^2} \quad (4.15)$$

■ **Área en metros cuadrados del *bounding box***

$$AREA_BB_M = \frac{H \cdot W}{PPM^2} \quad (4.16)$$

donde H y W son el alto y el ancho del *bounding box* del objeto.

■ **Perímetro en metros**

$$PERIMETRO_M = \frac{PERIMETRO}{PPM} \quad (4.17)$$

■ **Perímetro en metros del *bounding box***

$$PERIMETRO_BB_M = \frac{2 \cdot (H + W)}{PPM} \quad (4.18)$$

donde H y W son el alto y el ancho del *bounding box* del objeto.

■ **Compacidad (*compactness*)**

La relación de compacidad expresa que tan similar es un objeto a un círculo.

$$COMPACIDAD = \frac{PERIMETRO^2}{AREA} \quad (4.19)$$

Éste descriptor tiene un valor de 4π (su mínimo valor) cuando el objeto es un círculo y de 16 cuando el objeto es un cuadrado.

■ **Excentricidad (*eccentricity*)**

La excentricidad mide la relación de aspecto de un objeto. Es la relación entre la longitud del eje mayor y la longitud del eje menor.

$$EXCENTRICIDAD = \sqrt{1 - \left(\frac{\lambda_2}{\lambda_1}\right)^2} \quad (4.20)$$

donde $\lambda_1 \geq \lambda_2$ son los autovalores de la matriz de covarianza C de la ecuación 4.10 definida en la sección 4.2.

Para el caso de un objeto circular tenemos que $\lambda_1 = \lambda_2$ y la excentricidad es 0, mientras que si el objeto es una línea tenemos que $\lambda_2 = 0$, y la excentricidad del objeto es 1.

- **Relación de aspecto del *bounding box***

$$RELACION_ASPECTO_BB = \frac{W}{H} \quad (4.21)$$

donde H y W son el alto y el ancho del *bounding box* del objeto.

- ***Extent***

Es la relación entre el área del objeto y el área del *bounding box* del mismo.

$$EXTENT = \frac{AREA}{H \cdot W} \quad (4.22)$$

donde H y W son el alto y el ancho del *bounding box* del objeto.

- **Convexidad (*convexity*)**

La convexidad es la relación entre el perímetro de la cápsula convexa del objeto y el perímetro del mismo.

$$CONVEXIDAD = \frac{PERIMETRO_CH}{PERIMETRO} \quad (4.23)$$

donde $PERIMETRO_CH$ es el perímetro de la cápsula convexa del objeto.

- **Solidez (*solidity*)**

La solidez describe el grado de convexidad de un objeto.

$$SOLIDEZ = \frac{AREA}{AREA_CH} \quad (4.24)$$

donde $AREA_CH$ es el área de la cápsula convexa del objeto.

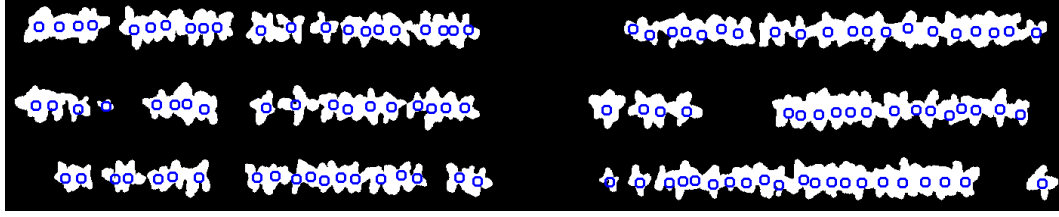


Figura 4.19: Imagen segmentada con plantas marcadas con un círculo

4.6. Conjuntos de muestras

A partir de los conjuntos de coordenadas de plantas, `camet1x-entrenamiento` y `camet1x-testeo`, se crean conjuntos de muestras de entrenamiento y testeo, de la siguiente manera: Para cada objeto segmentado, que contiene alguna de las coordenadas de los conjuntos de coordenadas, se les calcula los descriptores de características definidos en la sección 4.5 y se contabiliza la cantidad de coordenadas que contienen, es decir, la cantidad de plantas que fueron marcadas.

En la figura 4.19 se muestran las plantas marcadas sobre la máscara correspondiente a dos parcelas donde se puede apreciar la cantidad de plantas que contiene cada objeto.

			
<i>AREA_M</i>	0.039	0.118	0.232
<i>AREA_BB_M</i>	0.065	0.193	0.359
<i>PERIMETRO_M</i>	1.279	2.076	3.559
<i>PERIMETRO_BB_M</i>	1.021	1.832	2.699
<i>COMPACIDAD</i>	42.004	36.523	54.561
<i>EXCENTRICIDAD</i>	0.473	0.951	0.988
<i>RELACION_ASPECTO_BB</i>	1.028	1.787	2.712
<i>EXTENT</i>	0.598	0.611	0.648
<i>CONVEXIDAD</i>	0.710	0.730	0.653
<i>SOLIDEZ</i>	0.708	0.781	0.778
<i>CANTIDAD_PLANTAS</i>	1	3	5

Tabla 4.2: Muestras para objetos con 1, 3, y 5 plantas contabilizadas

Los conjuntos de muestras generados cuentan con 2055 muestras para el conjunto de entrenamiento, correspondiente a las 135 parcelas de entrenamiento, y 249 para el conjunto de testeo, correspondiente a las 15 parcelas de testeo. Cada muestra tiene 11 campos, donde los campos son, los valores obtenidos a partir de los descriptores y la cantidad de plantas en cada objeto, que denominaremos *CANTIDAD_PLANTAS*. En la tabla 4.2 se dan como

ejemplo, 3 muestras que son calculadas en objetos donde se contabilizaron, 1, 3, y 5 plantas.

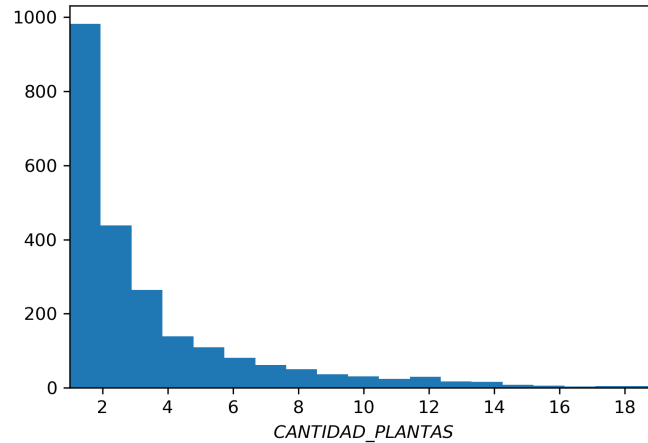


Figura 4.20: Histograma de *CANTIDAD_PLANTAS*

En la figura 4.20 se muestra el histograma de *CANTIDAD_PLANTAS* para las 2304 muestras que conforman el total de muestras. Se puede observar que los objetos con pocas plantas son mucho más frecuentes y que los objetos se van volviendo menos frecuentes a medida que contienen más plantas.

4.7. Regresión

A partir del conjunto de muestras de entrenamiento se entrenaron modelos de regresión y luego se validaron mediante validación cruzada repetida.

Los métodos propuestos en éste trabajo para estimar la cantidad de plantas son: Modelo lineal multivariado con mínimos cuadrados, regresión Ridge, Lasso, y SVR (Regresión de Vectores Soporte) lineal.

4.7.1. Modelo lineal multivariado con mínimos cuadrados

El modelo de regresión lineal [32, 33] ha sido un pilar de las estadísticas durante los últimos 40 años y sigue siendo una de las herramientas más importantes.

El mismo utiliza la relación entre dos o más variables para que una de las variables, llamada variable de respuesta, pueda predecirse a partir de las restantes, llamadas variables predictoras.

Dado un vector de entrada $X^T = \langle x_1, \dots, x_p \rangle$ donde sus componente son las variables predictoras, queremos predecir el valor real de y , denominada variable de respuesta, como salida.

El modelo de regresión lineal tiene la forma

$$y = \beta_0 + \sum_{j=1}^p x_j \beta_j + \epsilon \quad (4.25)$$

donde $\beta_0, \dots, \beta_p$ son coeficientes desconocidos y ϵ es el error o residuo. El término β_0 se denomina intersección y también se conoce como sesgo.

El modelo lineal asume que la esperanza condicional $E(y|X)$, también conocida como función de regresión, es lineal en sus variables de entrada, ó que el modelo lineal es una aproximación razonable para las mismas.

Dado un conjunto de entrenamiento $\{m_1, \dots, m_N\}$ donde $m_i = \langle x_{i1}, \dots, x_{ip}, y_i \rangle$ es la i -ésima muestra del conjunto, estimamos los coeficientes $\beta = \langle \beta_0, \dots, \beta_p \rangle^T$ mediante el método de mínimos cuadrados, que consiste en elegir β tal que minimice la suma de cuadrados residual, RSS

$$RSS(\beta) = \sum_{i=1}^N \epsilon_i^2 = \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 \quad (4.26)$$

4.7.2. Regresión Ridge

Cuando hay muchas variables correlacionadas en un modelo de regresión lineal, sus coeficientes pueden estar mal determinados y exhibir una alta varianza. La regresión Ridge [33] reduce los coeficientes $\beta_1, \dots, \beta_p$ imponiendo una penalización a sus tamaños. Ésto se logra minimizando una suma de cuadrados residual penalizada, como se muestra en la siguiente ecuación

$$\hat{\beta}_{ridge} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^p \beta_j^2 \right\} \quad (4.27)$$

donde $\lambda \geq 0$ es un parámetro de ajuste que controla la cantidad de contracción, cuanto mayor es el valor de λ , mayor es la contracción y los coeficientes se reducen hacia cero y entre sí, disminuyendo así, la varianza.

4.7.3. Lasso

Lasso [33] es un método similar al de la regresión Ridge, con la diferencia de que se reemplaza la norma L_2 : $\sum_{j=1}^p \beta_j^2$, por la norma L_1 : $\sum_{j=1}^p |\beta_j|$, en el término de penalización, como se aprecia a continuación

$$\hat{\beta}_{lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (4.28)$$

Debido a la naturaleza de la norma L_1 , hacer que λ sea lo suficientemente grande, hará que algunos coeficientes β sean estrictamente 0. Por tanto, el método Lasso hace una suerte de selección de variables, produciendo un modelo más interpretativo.

4.7.4. Regresión de vectores soporte lineal

La SVR [33] es un método de regresión que se deriva a partir del concepto de SVM (Máquina de Vectores Soporte) empleada para la clasificación, donde la salida es un valor real en lugar de una etiqueta de clase.

Para estimar los coeficientes β de un modelo lineal $f(x) = \beta_0 + \sum_{j=1}^p x_j \beta_j$, se minimiza la ecuación

$$H(\beta, \beta_0) = \sum_{i=1}^N V_{\epsilon}(y_i - f(x_i)) + \frac{\lambda}{2} \|\beta\|^2 \quad (4.29)$$

donde

$$V_{\epsilon}(r) = \begin{cases} 0 & \text{si } |r| < \epsilon \\ |r| - \epsilon & \text{si } |r| \geq \epsilon \end{cases} \quad (4.30)$$

es una medida de error denominada función de pérdida, que ignora los errores menores a ϵ de forma similar, a como, la SVM ignora los puntos en el lado correcto del límite de decisión y alejados de él, en la optimización. En la regresión, estos puntos de bajo error son los que tienen residuos pequeños y se encuentran dentro del margen máximo determinado por dos bandas paralelas al hiperplano como se muestra en la imagen 4.21.

Los puntos fuera de las bandas son los denominados vectores soporte, ya que a partir de ellos se optimiza el modelo.

El parámetro λ es un parámetro de regularización y controla la penalización. Si λ toma un valor pequeño, aumenta la penalización, mientras que si el valor de λ es grande, la penalización será menos estricta.

4.7.5. Estandarización y análisis de las variables

Antes de avanzar con los modelos de regresión es necesario estandarizar (normalizar) las variables del conjunto de muestras de entrenamiento. La estandarización centra en 0 los datos restando la media μ , y los escala para que la varianza sea igual a 1 dividiéndolos por el desvío estandar σ .

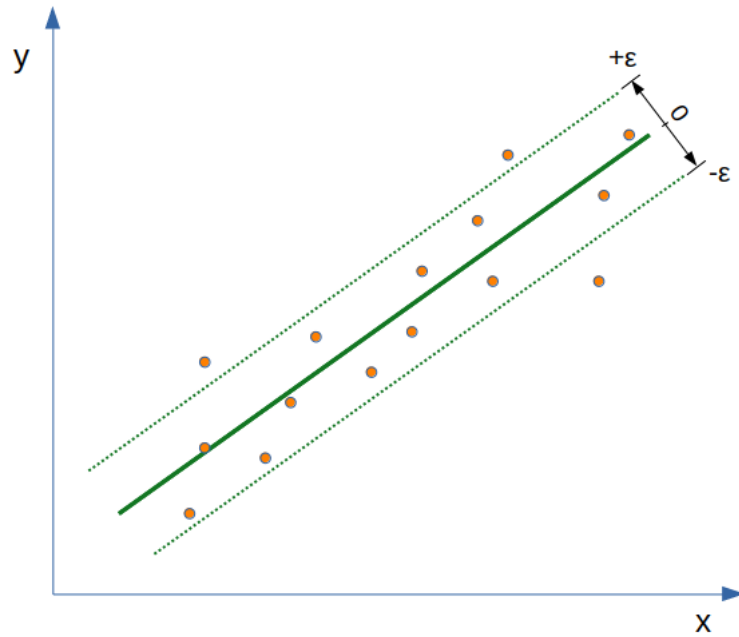


Figura 4.21: Hiperplano y bandas en la regresión de vectores soporte

Una vez obtenidos los modelos, se estandariza el conjunto de muestras de testeo utilizando los mismos parámetros, μ y σ , con los que se estandarizó el conjunto de muestras de entrenamiento y se obtiene la predicción de la variable *CANTIDAD_PLANTAS*, a la cual se le realiza la transformación inversa de la estandarización para volver a la escala original de dicha variable.

Como el objetivo es utilizar las variables asociadas a los descriptores de características como variables predictoras y *CANTIDAD_PLANTAS* como variable de respuesta, se analiza la relación de cada variable predictora con la variable de respuesta en el conjunto de muestras de entrenamiento para evaluar su comportamiento. En la figura 4.22 se muestran dichas relaciones.

En la figura 4.23 se muestra el *heatmap* con los coeficientes de correlación de Pearson de las variables y sus valores absolutos. Puede observarse que las variables *AREA_M*, *AREA_BB_M*, *PERIMETRO_M*, *PERIMETRO_BB_M*, *COMPACIDAD* y *RELACION_ASPECTO_BB* presentan una alta correlación lineal con la variable de respuesta *CANTIDAD_PLANTAS*, y a la vez existe una alta correlación entre las mismas.

4.7.6. Hiperparámetro λ para Ridge, Lasso, y SVR

Para implementar la regresión de Ridge, Lasso y SVR se requiere de un método para seleccionar un valor para el hiperparámetro de ajuste λ de las ecuaciones 4.27, 4.28 y 4.29.

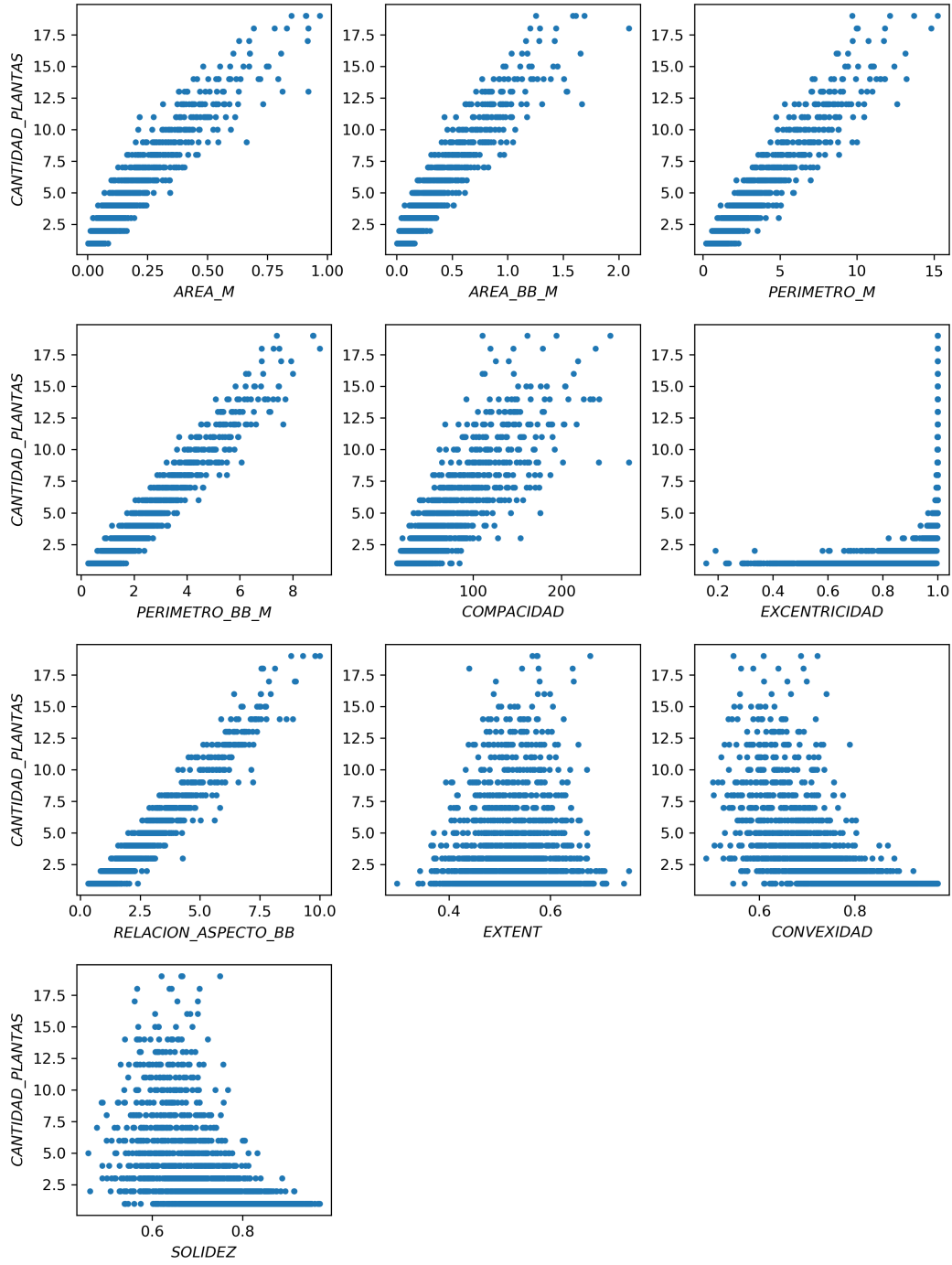


Figura 4.22: Relaciones entre la variable de respuesta y las restantes

La validación cruzada proporciona una forma sencilla de abordar éste problema [34].

Elegimos un conjunto de valores de λ y calculamos el error de la validación cruzada k-fold para cada valor de λ . Luego seleccionamos el valor del parámetro de ajuste para el cual el error de validación cruzada es más pequeño y re-entrenamos el modelo utilizando



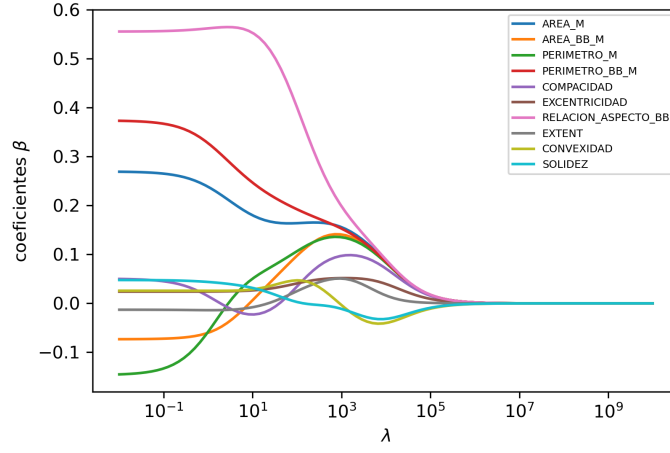
Figura 4.23: Coeficientes de correlación de Pearson (debajo de la diagonal) y sus valores absolutos (arriba de la diagonal)

dicho valor.

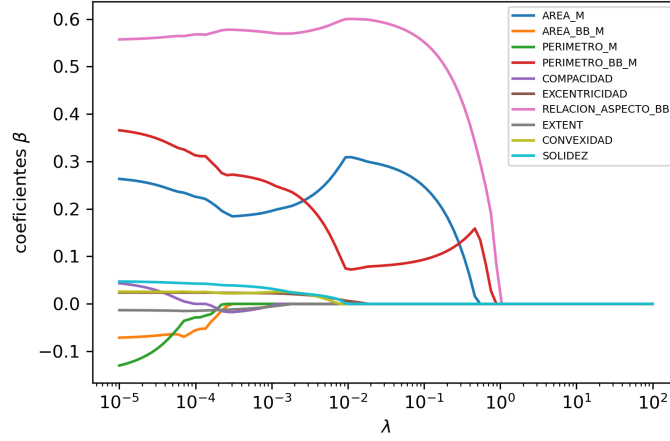
Para Ridge se seleccionaron 100 valores espaciados uniformemente en escala logarítmica entre 10^{-2} y 10^{10} , y en el caso de Lasso entre 10^{-5} y 10^2 . Estos valores se seleccionaron a partir de las muestras de entrenamiento para cubrir una amplia gama de escenarios: desde el aproximadamente modelo de mínimos cuadrados, es decir $\lambda = 0$, hasta el modelo nulo, es decir, el modelo que contiene únicamente la intercepción. En ambos casos se utilizó $k = 10$ para la validación cruzada k-fold interna del método.

En las figuras 4.24a y 4.24b se muestran los valores de los coeficientes de la regresión Ridge y Lasso respectivamente en función del valor de λ . Se puede observar que a medida que λ es mayor, la contracción es mayor y los valores de los coeficientes se reducen. En el caso de Lasso algunos coeficientes se hacen cero de forma abrupta con muy poca contracción.

Para la implementación de la SVR se seleccionaron valores de $\lambda = [0.01, 0.1, 1, 10, 100]$. Cabe aclarar que a menudo el parámetro utilizado para entrenar modelos SVR al igual que



(a) Ridge



(b) Lasso

Figura 4.24: Valores de los coeficientes β en función de λ para Ridge y Lasso

la SVM recibe el nombre de C , y que resulta $C = 1/\lambda$.

4.7.7. Validación cruzada repetida

La validación cruzada k-fold consiste en dividir aleatoriamente un conjunto de muestras en k particiones de igual tamaño, para luego entrenar un modelo utilizando $k-1$ particiones y testear con la partición restante. Este procedimiento se realiza k veces, dejando en cada iteración una partición distinta para testeo.

La validación cruzada repetida k-fold consiste en repetir el procedimiento de validación cruzada k-fold, n veces, donde en cada repetición dividimos aleatoriamente en k particiones distintas el conjunto de muestras. De ésta manera se obtienen $k \times n$ modelos distintos.

Para evaluar los modelos de regresión utilizaremos la validación cruzada k-fold repetida sobre el conjunto de muestras entrenamiento con $k = 10$ y $n = 10$.

Capítulo 5

Resultados

En éste capítulo se presentarán los resultados obtenidos a lo largo de éste trabajo.

Primero se analizarán los resultados del método de segmentación propuesto en la sección 4.1. A continuación se mostrarán los resultados del algoritmo de detección de la orientación de las hileras presentado en la sección 4.2.

Luego se evaluará el algoritmo de detección de las hileras desarrollado en la sección 4.4. Después se expondrán los resultados de los modelos generados utilizando los métodos de regresión detallados en la sección 4.7.

Finalmente se evalúa la metodología en su conjunto, a partir de los modelos generados.

5.1. Segmentación

El método de segmentación propuesto resultó efectivo para separar los objetos de plantas de girasol del fondo, tanto en días soleados como nublados, con y sin riego.

A continuación se muestran los resultados para 4 imágenes que cubren distintos escenarios: soleado y nublado, con y sin riego, plantas pequeñas y grandes, poca y mucha maleza.



Figura 5.1: Máscara de la imagen camet1cr.jpg

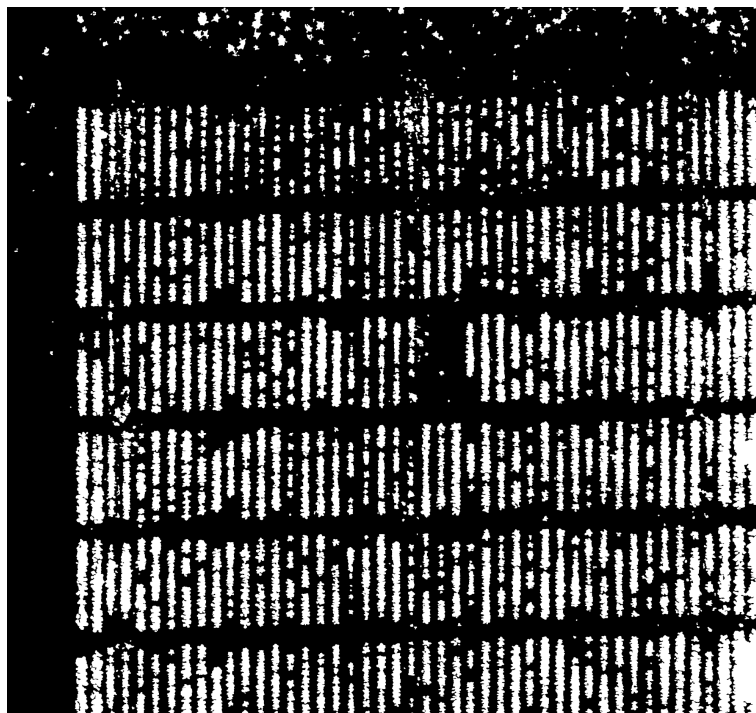


Figura 5.2: Máscara de la imagen camet2sr.jpg

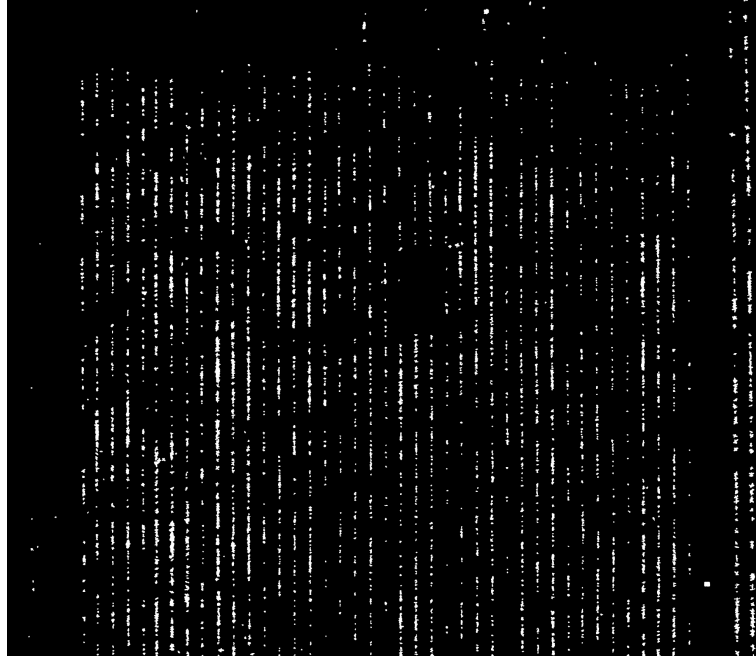


Figura 5.3: Máscara de la imagen quemulcr.jpg

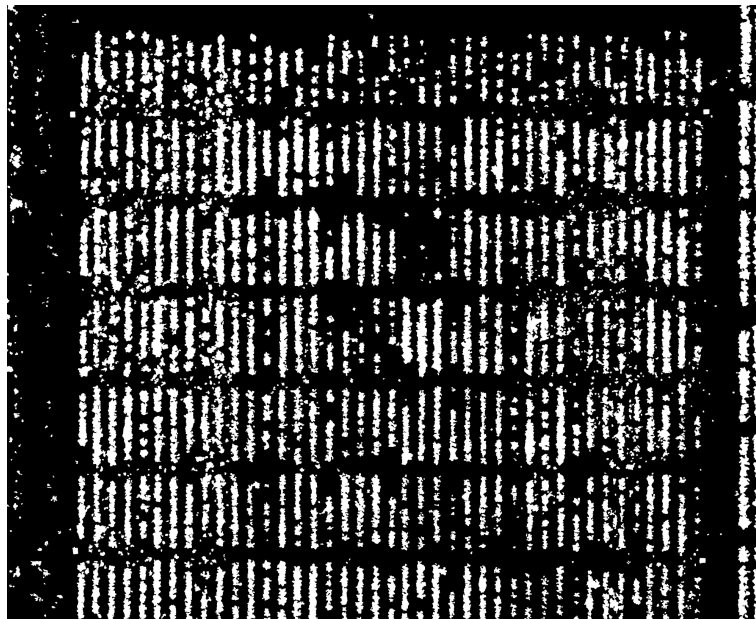


Figura 5.4: Máscara de la imagen quemu2sr.jpg

La imagen quemu2sr.jpg es la más desafiante de todas por presentar mucha maleza, parte de ella se clasificó como objetos de plantas de girasol. Una desventaja del método, es que al basarse en el color verde, no discrimina entre distintos tipos de elementos que tienen una tonalidad similar. Cabe destacar, por otra parte, que las ventajas del método son que tiene un bajo costo computacional y que no requiere entrenamiento.

5.2. Detección de la orientación de las hileras

El algoritmo de detección de la orientación de las hileras detectó correctamente la orientación en todas las imágenes. En la tabla 5.1 se detallan la orientación obtenida en cada caso.

Imagen	Grados
camet1.tif	42.07
camet1cr.jpg	-90.05
camet1sr.jpg	-89.93
quemu1cr.jpg	-89.89
quemu1sr.jpg	-90.52
camet2cr.jpg	-89.89
camet2sr.jpg	-90.13
quemu2cr.jpg	-89.48
quemu2sr.jpg	-90.14

Tabla 5.1: Orientación de las imágenes

5.3. Detección de hileras y etiquetado de objetos

La evaluación del desempeño del algoritmo de detección de hileras y etiquetado de objetos se realizó sobre la imagen *camet1x.tif*, calculando los valores *precision* y *recall* utilizando como referencia el conjunto de coordenadas de plantas *camet1x-completo*.

Los valores *precision* y *recall* se calcularon mediante las formulas

$$Precision = \frac{TP}{TP + FP} \quad (5.1)$$

$$Recall = \frac{TP}{TP + FN} \quad (5.2)$$

donde *TP* es la cantidad de objetos detectados que contienen plantas, *FN* es la cantidad de objetos no detectados que contienen plantas, y *FP* es la cantidad de objetos detectados que no contienen plantas.

También se calcularon dichos valores directamente sobre los objetos segmentados, sin detección de hileras, para comparar si efectivamente había una mejora.

Los resultados se muestran en la tabla 5.2 y se observa que el algoritmo mejora la precisión de la detección, ya que filtra objetos que en realidad no son plantas.

	Precision	Recall
Sin detección de hileras	0.90	0.94
Con detección de hileras	0.95	0.94

Tabla 5.2: Precision y recall

5.4. Regresión

5.4.1. Entrenamiento

A partir de la validación cruzada repetida k-fold se obtuvieron 100 modelos para cada método de regresión propuesto. Para cada método se calculó el R^2 promedio y el desvío estándar. Dichos valores se muestran en la tabla 5.3 y se grafican en la figura 5.5.

Podemos apreciar que con los distintos métodos obtenemos resultados prácticamente iguales. Los desvíos estándar son muy bajos y no se produjeron valores *outliers* en ninguno de los casos.

Método de regresión	R^2	Desvío estándar
Mínimos cuadrados	0.96	0.01
Ridge	0.96	0.01
Lasso	0.96	0.01
SVR	0.96	0.01

Tabla 5.3: R^2 promedio y desvío estándar

Por cada método de regresión se seleccionó el modelo con mejor R^2 , los coeficientes de los mismos se muestran en la tabla 5.4. Puede observarse que los modelos tienen poco sesgo y que en el caso de Lasso se hicieron cero 3 de las 10 variables, dando lugar a un modelo más interpretable.

5.4.2. Testeo

Se aplicaron los modelos de la tabla 5.4 sobre el conjunto de muestras de testeo y se calcularon los R^2 contra dicho conjunto. También se calcularon los R^2 contra el conteo del INTA. Dichos resultados se muestran en la tabla 5.5.

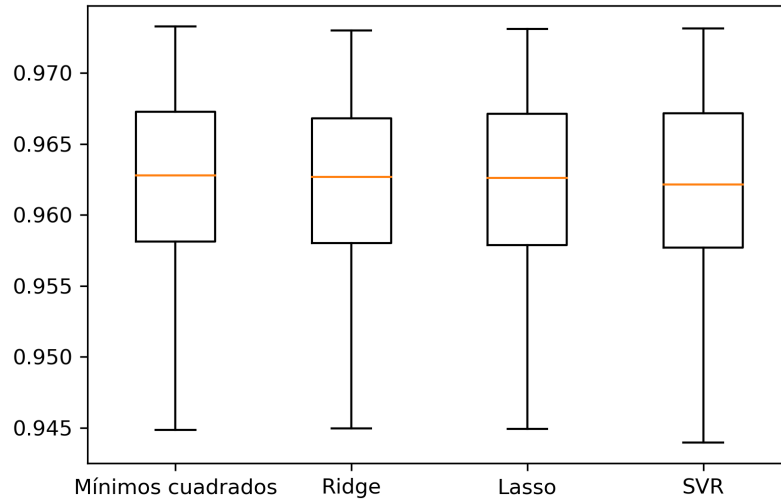


Figura 5.5: Boxplots para cada modelo

Variable	Mínimos cuadrados	Ridge	Lasso	SVR
intercept	0.002	0.002	0.002	-0.005
<i>AREA_M</i>	0.243	0.165	0.154	0.383
<i>AREA_BB_M</i>	-0.096	-0.017	0.000	-0.131
<i>PERIMETRO_M</i>	-0.114	0.053	0.000	-0.300
<i>PERIMETRO_BB_M</i>	0.399	0.277	0.299	0.396
<i>COMPACIDAD</i>	0.059	-0.009	0.000	0.147
<i>EXCENRICIDAD</i>	0.025	0.026	0.023	0.027
<i>RELACION_ASPECTO_BB</i>	0.542	0.550	0.565	0.546
<i>EXTENT</i>	-0.013	-0.010	-0.007	-0.013
<i>CONVEXIDAD</i>	0.032	0.032	0.026	0.029
<i>SOLIDEZ</i>	0.051	0.038	0.039	0.057

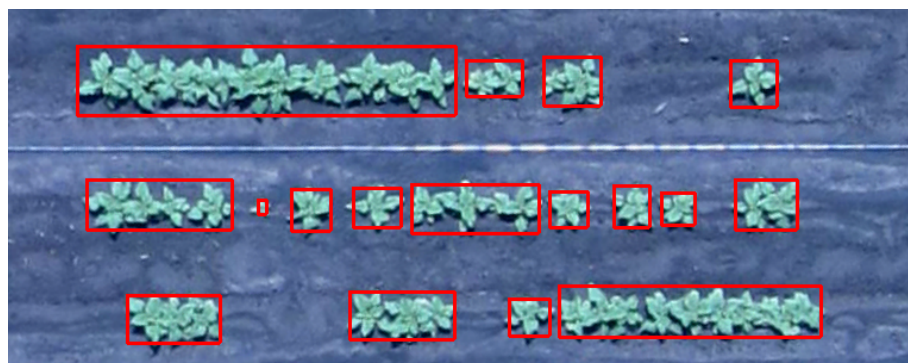
Tabla 5.4: Coeficientes de modelos con mejor R^2

Método de regresión	Conjunto de muestras de testeo	Conteo INTA
Mínimos cuadrados	0.96	0.86
Ridge	0.96	0.86
Lasso	0.96	0.86
SVR	0.96	0.85

Tabla 5.5: R^2 de los modelos

A continuación se muestran las imágenes correspondientes a las parcelas de testeo recortadas automáticamente a partir del conjunto de coordenadas de parcelas camet1x-

parcelas-testeo, junto con los valores de las predicciones de cada modelo, la cantidad de plantas marcadas y la cantidad del conteo del INTA.



Mín. cuad.: 47.59
Ridge: 47.54
Lasso: 47.67
SVR: 47.01
Marcadas: 49
INTA: 50

Figura 5.6: Conteo en parcela 8 de camet1x.tif



Mín. cuad.: 50.65
Ridge: 50.61
Lasso: 50.69
SVR: 50.33
Marcadas: 50
INTA: 57

Figura 5.7: Conteo en parcela 34 de camet1x.tif



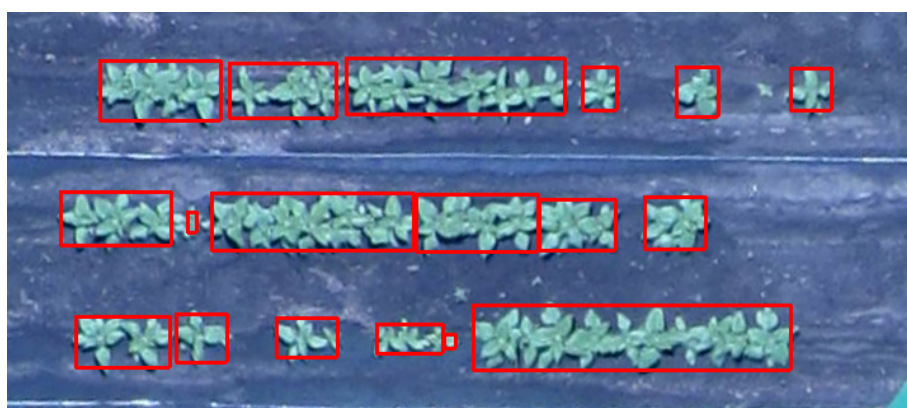
Mín. cuad.: 54.37
Ridge: 54.42
Lasso: 54.40
SVR: 53.74
Marcadas: 51
INTA: 56

Figura 5.8: Conteo en parcela 41 de camet1x.tif



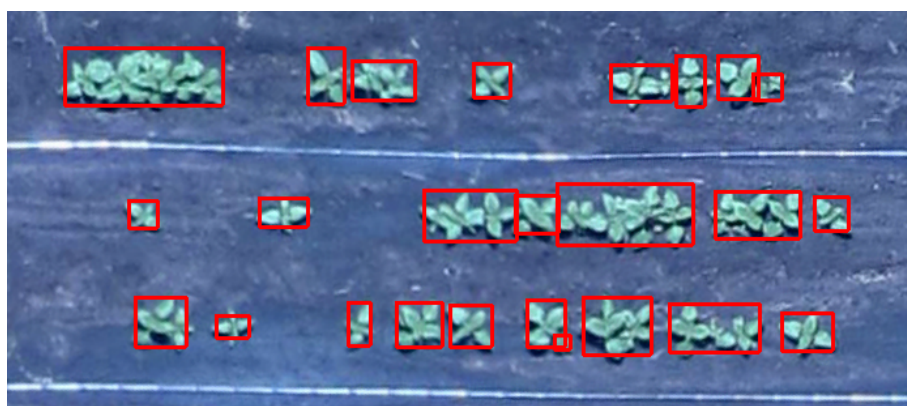
Mín. cuad.: 46.15
 Ridge: 46.12
 Lasso: 46.09
 SVR: 45.69
 Marcadas: 48
 INTA: 51

Figura 5.9: Conteo en parcela 52 de camet1x.tif



Mín. cuad.: 47.96
 Ridge: 47.97
 Lasso: 48.01
 SVR: 47.62
 Marcadas: 52
 INTA: 56

Figura 5.10: Conteo en parcela 55 de camet1x.tif



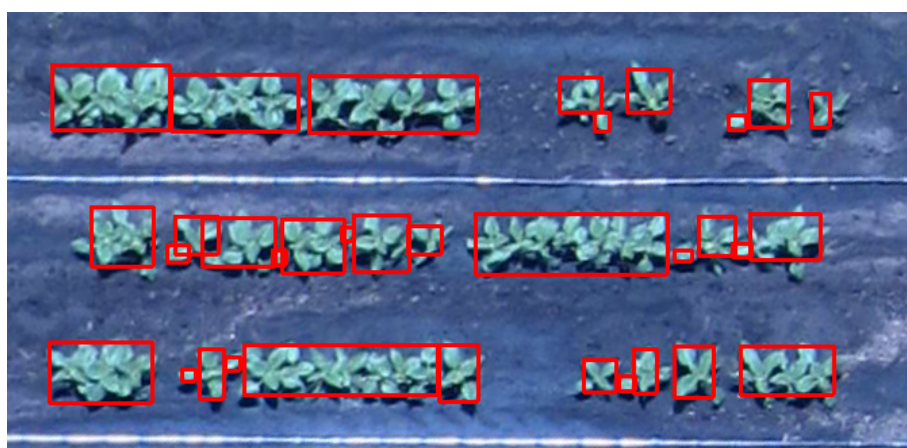
Mín. cuad.: 39.83
 Ridge: 39.94
 Lasso: 39.74
 SVR: 39.56
 Marcadas: 39
 INTA: 41

Figura 5.11: Conteo en parcela 63 de camet1x.tif



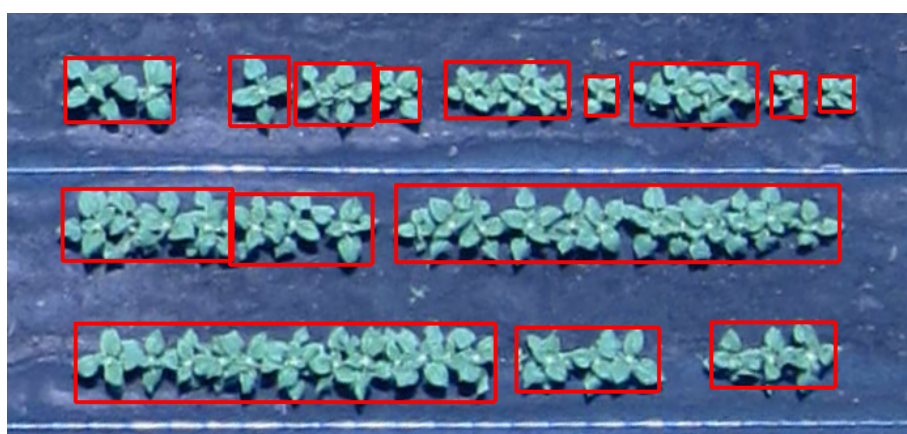
Mín. cuad.: 52.28
 Ridge: 52.20
 Lasso: 52.30
 SVR: 51.65
 Marcadas: 49
 INTA: 54

Figura 5.12: Conteo en parcela 64 de camet1x.tif



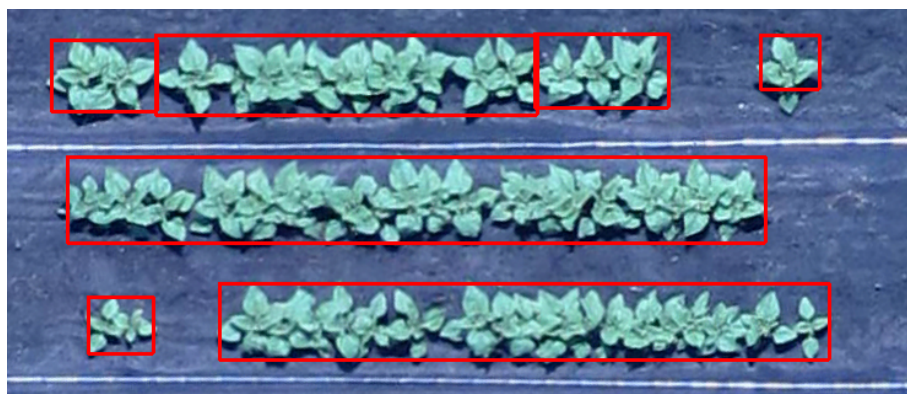
Mín. cuad.: 42.93
 Ridge: 42.63
 Lasso: 42.59
 SVR: 43.45
 Marcadas: 41
 INTA: 50

Figura 5.13: Conteo en parcela 72 de camet1x.tif



Mín. cuad.: 50.61
 Ridge: 50.95
 Lasso: 50.82
 SVR: 50.09
 Marcadas: 51
 INTA: 55

Figura 5.14: Conteo en parcela 74 de camet1x.tif



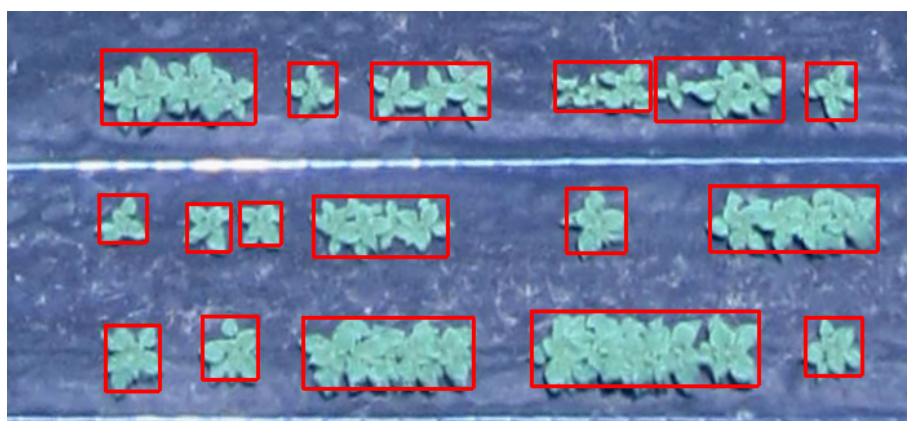
Mín. cuad.: 54.16
 Ridge: 54.54
 Lasso: 54.23
 SVR: 54.06
 Marcadas: 48
 INTA: 57

Figura 5.15: Conteo en parcela 77 de camet1x.tif



Mín. cuad.: 44.78
 Ridge: 44.62
 Lasso: 44.58
 SVR: 44.18
 Marcadas: 41
 INTA: 42

Figura 5.16: Conteo en parcela 87 de camet1x.tif



Mín. cuad.: 43.95
 Ridge: 43.54
 Lasso: 43.59
 SVR: 43.43
 Marcadas: 41
 INTA: 40

Figura 5.17: Conteo en parcela 101 de camet1x.tif

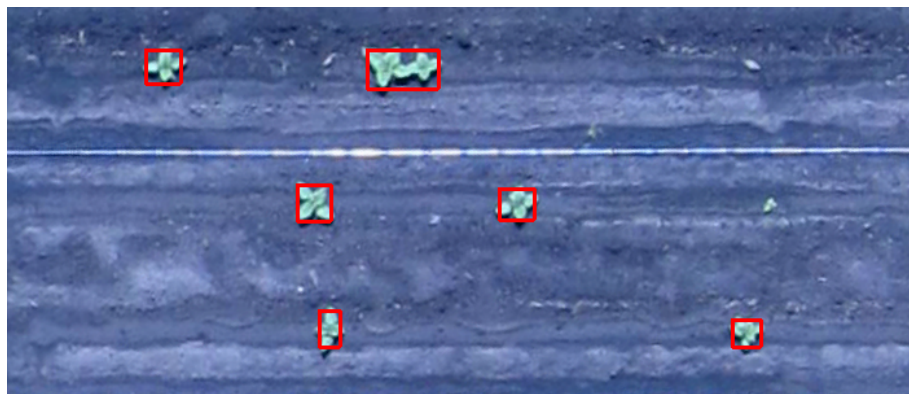


Figura 5.18: Conteo en parcela 108 de camet1x.tif

Mín. cuad.: 7.18
 Ridge: 7.22
 Lasso: 7.17
 SVR: 7.15
 Marcadas: 7
 INTA: 8



Figura 5.19: Conteo en parcela 115 de camet1x.tif

Mín. cuad.: 49.48
 Ridge: 49.52
 Lasso: 49.31
 SVR: 49.16
 Marcadas: 53
 INTA: 57

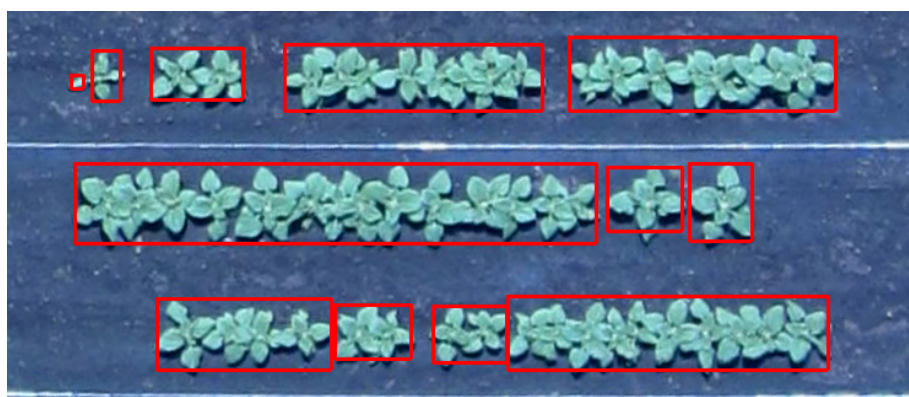


Figura 5.20: Conteo en parcela 135 de camet1x.tif

Mín. cuad.: 53.05
 Ridge: 52.99
 Lasso: 52.87
 SVR: 52.81
 Marcadas: 51
 INTA: 59

5.5. Metodología completa

Se aplicó la metodología completa utilizando los modelos de la tabla 5.4 a la imagen camet2crx.tif, donde las plantas se encuentran en un estadio de crecimiento más avanzado,

y se calcularon los R^2 de cada modelo frente al conteo del INTA.

Los valores obtenidos de R^2 fueron de 0.43 para Mínimos cuadrados, 0.41 para Ridge, 0.40 para Lasso, y 0.47 para SVR. Sin embargo, se observó que la parcela 28 se había segmentado incorrectamente ya que 2 objetos de hileras diferentes estaban conectados entre sí como se muestra en la figura 5.21. Ésto ocasionó que el algoritmo de la sección 4.4 no etiquetara tales objetos dentro de ninguna hilera. Se repitió el experimento excluyendo la parcela 28 y se obtuvo un R^2 de 0.78 para Mínimos cuadrados, 0.76 para Ridge, 0.75 para Lasso, y 0.80 para SVR.



Figura 5.21: Dos objetos de diferentes hileras conectados entre sí (color azul)

A continuación se muestran las imágenes correspondientes a las parcelas de la 1 a la 10, junto con los valores de las predicciones de cada modelo y del conteo del INTA, extraídas de la imagen camet2crx.tif a partir del conjunto de coordenadas de parcelas camet2crx-parcelas.



Mín. cuad.: 55.66
Ridge: 54.34
Lasso: 54.14
SVR: 58.75
INTA: 59

Figura 5.22: Conteo en parcela 1 de camet2crx.tif



Mín. cuad.: 47.24
 Ridge: 46.76
 Lasso: 46.79
 SVR: 47.52
 INTA: 54

Figura 5.23: Conteo en parcela 2 de camet2crx.tif



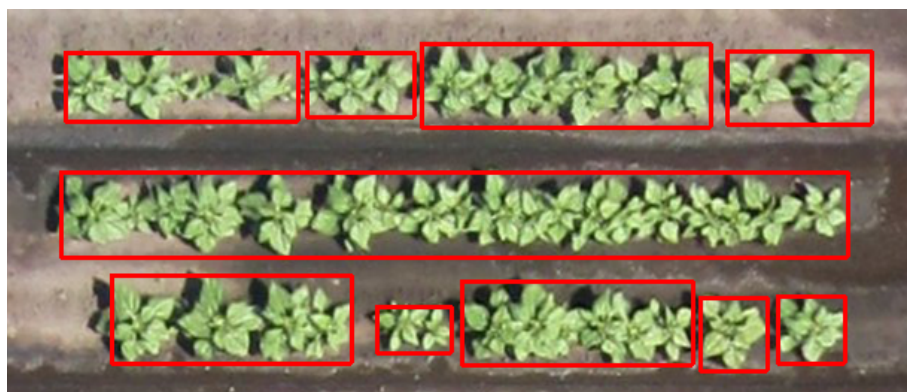
Mín. cuad.: 45.89
 Ridge: 46.19
 Lasso: 46.00
 SVR: 45.37
 INTA: 49

Figura 5.24: Conteo en parcela 3 de camet2crx.tif



Mín. cuad.: 50.33
 Ridge: 49.56
 Lasso: 49.71
 SVR: 50.11
 INTA: 53

Figura 5.25: Conteo en parcela 4 de camet2crx.tif



Mín. cuad.: 55.06
 Ridge: 54.93
 Lasso: 54.93
 SVR: 54.85
 INTA: 58

Figura 5.26: Conteo en parcela 5 de camet2crx.tif



Mín. cuad.: 51.40
 Ridge: 51.67
 Lasso: 51.46
 SVR: 51.57
 INTA: 53

Figura 5.27: Conteo en parcela 6 de camet2crx.tif



Mín. cuad.: 48.83
 Ridge: 48.37
 Lasso: 48.21
 SVR: 50.22
 INTA: 50

Figura 5.28: Conteo en parcela 7 de camet2crx.tif



Mín. cuad.: 51.83
 Ridge: 51.35
 Lasso: 51.18
 SVR: 52.80
 INTA: 50

Figura 5.29: Conteo en parcela 8 de camet2crx.tif



Mín. cuad.: 39.86
 Ridge: 39.70
 Lasso: 39.64
 SVR: 39.97
 INTA: 45

Figura 5.30: Conteo en parcela 9 de camet2crx.tif



Mín. cuad.: 53.36
 Ridge: 52.26
 Lasso: 52.20
 SVR: 55.61
 INTA: 53

Figura 5.31: Conteo en parcela 10 de camet2crx.tif

Capítulo 6

Conclusiones

Actualmente el INTA realiza el conteo de plantas de girasol de forma manual, contando de forma individual cada planta a partir de imágenes UAV. En la actualidad no se ha publicado un método específico para realizar ésta tarea de forma automática.

En éste trabajo se propuso un conjunto de algoritmos para el conteo automático de las plantas de girasol en cultivos en hileras a partir de imágenes UAV, que abarcan, la detección de los objetos de plantas de girasol, la corrección de la orientación de la imagen para que las hileras de los cultivos queden en forma horizontal, la detección de los objetos que componen las hileras, y por último, el conteo de plantas de girasol mediante modelos de regresión lineal previamente entrenados a partir de descriptores de forma y tamaño que son extraídos de los objetos detectados que contienen algún punto en el conjunto de coordenadas de plantas creado durante éste trabajo.

Para la detección de objetos de plantas de girasol se obtuvo una *precision* y *recall* de 0.95 y 0.94, respectivamente.

Durante el entrenamiento se obtuvo un valor de R^2 promedio de 0.96 para cada uno de los cuatro métodos de regresión propuestos: Modelo lineal multivariado con mínimos cuadrados, regresión Ridge, Lasso, y SVR lineal.

Los valores de R^2 obtenidos en el testeó fueron de 0.96 para todos los métodos de regresión. Los valores de R^2 comparando frente al conteo del INTA fueron de 0.86 para los métodos modelo lineal multivariado con mínimos cuadrados, regresión Ridge y Lasso, y de 0.85 para SVR lineal.

También se testeó en el estadio de crecimiento 27 DDE, cuando las plantas son más grandes y complejas, y se obtuvo un valor de R^2 de 0.78 para el modelo lineal multivariado con mínimos cuadrados, 0.76 para Ridge, 0.75 para Lasso, y 0.80 para SVR.

Uno de los problemas que se presentó durante el desarrollo de éste trabajo, fue la falta

de un conjunto de coordenadas de plantas generado por un experto en el campo de la agricultura, motivo por el cual, se creó uno propio en su lugar, el cual tenía discrepancia con la cantidad de plantas contadas por parcelas por el INTA.

Otro de los problemas surgió durante la etapa de testeo en el estadio de crecimiento 27 DDE donde las plantas eran lo suficientemente grandes como para tocarse entre hileras. Esto dio como resultado que dos objetos de plantas de girasol se segmentaran como uno sólo y que el algoritmo de la sección 4.4 no etiquetara dichos objetos a ninguna hilera.

En trabajos futuros se podría trabajar sobre los problemas mencionados anteriormente. Sería conveniente contar con conjuntos de coordenadas de plantas, en distintos estadios de crecimiento, provistos por expertos del área. De ésta manera se podrían entrenar modelos de mejor calidad para estadios de crecimiento concretos. También sería conveniente trabajar sobre la separación de plantas de distintas hileras que se conectan entre sí por estar encimadas.

Otra opción interesante a explorar sería la segmentación de las plantas de girasol con un enfoque de aprendizaje automático supervisado utilizando como entradas las componentes correspondientes a los espacios de color RGB, HSV, y $L^*a^*b^*$ como propusieron en su trabajo Guo *et al* [8] para la detección de las panojas de sorgo.

Por último, en el caso de las plantas de girasol ya adultas, se podría aprovechar que la flor de girasol es visible y tiene una tonalidad amarilla distinguible a simple vista del verde de las hojas, para detectar las plantas. Pero para ésto se requiere que las imágenes posean la menor perspectiva posible. Por ejemplo, para minimizar la influencia de la distorsión de la lente de la cámara, Guo *et al* [8] utilizaron sólo la región central de cada imagen, abarcando tan sólo el 10% de los píxeles de cada una, aproximadamente.

Bibliografía

- [1] F. Andrade and V. Sadras, *Bases para el manejo del maíz, el girasol y la soja*. Instituto Nacional de Tecnología Agropecuaria, 2000.
- [2] J. Ribera, Y. Chen, C. Boomsma, and E. J. Delp, “Counting plants using deep learning,” in *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1344–1348, 2017.
- [3] PrecisionHawk, “Automated Plant Counts with Drones for Precision Agriculture.” <https://www.precisionhawk.com/media/topic/understanding-your-aerial-data-plant-counting>, 2019.
- [4] F. Gnädinger and U. Schmidhalter, “Digital Counts of Maize Plants by Unmanned Aerial Vehicles (UAVs),” *Remote Sensing*, vol. 9, no. 6, 2017.
- [5] T. Liu, W. Wu, W. Chen, C. Sun, X. Zhu, and W. Guo, “Automated image-processing for counting seedlings in a wheat field,” *Precision Agriculture*, vol. 17, 08 2016.
- [6] Y. Wang, X. Zhu, and B. Wu, “Automatic detection of individual oil palm trees from UAV images using HOG features and an SVM classifier,” *International Journal of Remote Sensing*, vol. 40, pp. 1–15, 09 2018.
- [7] B. Zhao, J. Zhang, C. Yang, G. Zhou, Y. Ding, Y. Shi, D. Zhang, J. Xie, and Q. Liao, “Rapeseed Seedling Stand Counting and Seeding Performance Evaluation at Two Early Growth Stages Based on Unmanned Aerial Vehicle Imagery,” *Frontiers in Plant Science*, vol. 9, p. 1362, 2018.
- [8] W. Guo, B. Zheng, A. B. Potgieter, J. Diot, K. Watanabe, K. Noshita, D. R. Jordan, X. Wang, J. Watson, S. Ninomiya, and S. C. Chapman, “Aerial Imagery Analysis – Quantifying Appearance and Number of Sorghum Heads for Applications in Breeding and Agronomy,” *Frontiers in Plant Science*, vol. 9, p. 1544, 2018.
- [9] Koh Joshua C. O., Hayden Matthew, Daetwyler Hans, and Kant Surya, “Estimation of crop plant density at early mixed growth stages using UAV imagery,” *Plant Methods*, vol. 15, no. 1, p. 64, 2019.

- [10] J. Wu, G. Yang, X. Yang, B. Xu, L. Han, and Y. Zhu, “Automatic Counting of in situ Rice Seedlings from UAV Images Based on a Deep Fully Convolutional Neural Network,” *Remote Sensing*, vol. 11, no. 6, 2019.
- [11] L. P. Osco, M. d. S. de Arruda, J. Marcato Junior, N. B. da Silva, A. P. M. Ramos, É. A. S. Moryia, N. N. Imai, D. R. Pereira, J. E. Creste, E. T. Matsubara, J. Li, and W. N. Gonçalves, “A convolutional neural network approach for counting and geolocating citrus-trees in UAV multispectral imagery,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 160, pp. 97–106, 2020.
- [12] Valente João, Sari Bilal, Kooistra Lammert, Kramer Henk, and Múcher Sander, “Automated crop plant counting from very high-resolution aerial imagery,” *Precision Agriculture*, vol. 21, no. 6, pp. 1366–1384, 2020.
- [13] Y. Pang, Y. Shi, S. Gao, F. Jiang, A.-N. Veeranampalayam-Sivakumar, L. Thompson, J. Luck, and C. Liu, “Improved crop row detection with deep neural network for early-season maize stand count in UAV imagery,” *Computers and Electronics in Agriculture*, vol. 178, p. 105766, 2020.
- [14] H. Fang and S. Liang, “Leaf Area Index Models,” in *Reference Module in Earth Systems and Environmental Sciences*, Elsevier, 2014.
- [15] D. M. Woebbecke, G. E. Meyer, K. V. Bargaen, and D. A. Mortensen, “Plant species identification, size, and enumeration using machine vision techniques on near-binary images,” vol. 1836, may 1993.
- [16] M. W. D., E. M. G., V. B. K., and A. M. D., “Color Indices for Weed Identification Under Various Soil, Residue, and Lighting Conditions,” *Transactions of the ASAE*, vol. 38, no. 1, pp. 259–269, 1995.
- [17] T. Kataoka, T. Kaneko, H. Okamoto, and S. Hata, “Crop growth estimation system using machine vision,” in *Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2003)*, vol. 2, pp. b1079–b1083 vol.2, 2003.
- [18] G. E. Meyer, T. W. Hindman, and K. Laksmi, “Machine vision detection parameters for plant species identification,” vol. 3543, jan 1999.
- [19] G. E. Meyer and J. C. Neto, “Verification of color vegetation indices for automated crop imaging applications,” *Computers and Electronics in Agriculture*, vol. 63, no. 2, pp. 282–293, 2008.
- [20] Hague T., Tillet N. D., and Wheeler H., “Automated Crop and Weed Monitoring in Widely Spaced Cereals,” *Precision Agriculture*, vol. 7, no. 1, pp. 21–32, 2006.

- [21] M. Louhaichi, M. M. Borman, and D. E. Johnson, “Spatially Located Platform and Aerial Photography for Documentation of Grazing Impacts on Wheat,” *Geocarto International*, vol. 16, no. 1, pp. 65–70, 2001.
- [22] M. Rico-Fernández, R. Rios-Cabrera, M. Castelán, H.-I. Guerrero-Reyes, and A. Juarez-Maldonado, “A contextualized approach for segmentation of foliage in different crop species,” *Computers and Electronics in Agriculture*, vol. 156, pp. 378–386, 2019.
- [23] Z. Diao, H. Wang, Y. Song, and Y. Wang, “Image segmentation method for cotton mite disease based on color features and area thresholding,” *Journal of Theoretical and Applied Information Technology*, vol. 48, pp. 527–533, 01 2013.
- [24] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [25] R. Gonzalez and R. Woods, *Digital Image Processing*. Pearson, 2018.
- [26] J. Russ, *The Image Processing Handbook*. CRC Press, 2016.
- [27] M. G. Larese, *Desarrollo de Algoritmos Automatizados Aplicados a Bioinformática*. PhD thesis, 2011.
- [28] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, vol. 1, pp. 886–893 vol. 1, 2005.
- [29] P. Lottes, R. Khanna, J. Pfeifer, R. Siegwart, and C. Stachniss, “UAV-based crop and weed classification for smart farming,” in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3024–3031, 2017.
- [30] R. Ji and L. Qi, “Crop-row detection algorithm based on Random Hough Transformation,” *Mathematical and Computer Modelling*, vol. 54, no. 3, pp. 1016–1020, 2011. Mathematical and Computer Modeling in agriculture (CCTA 2010).
- [31] M. Yang, K. Kpalma, and J. Ronsin, “A Survey of Shape Feature Extraction Techniques,” vol. 15, 11 2007.
- [32] M. Kutner, C. Nachtsheim, J. Neter, and W. Li, *Applied Linear Statistical Models*. McGraw-Hill Irwin, 2005.
- [33] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.
- [34] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*. Springer, 2013.