

Horticultura periurbana y teledetección

Herramientas analíticas de valoración y
cuantificación de la producción hortícola
basada en sensores remotos



María Victoria Marinelli

Horticultura periurbana y teledetección

Herramientas analíticas de valoración y
cuantificación de la producción hortícola
basada en sensores remotos



María Victoria Marinelli

*Centro Regional
Córdoba*



Instituto Nacional de
Tecnología Agropecuaria

Secretaría de Agricultura,
Ganadería y Pesca



Ministerio de Economía
Argentina

Marinelli, Maria Victoria

Horticultura periurbana y teledetección : herramientas analíticas de valoración y cuantificación de la producción hortícola basada en sensores remotos / Maria Victoria Marinelli ; prólogo de Beatriz Giobellina ... [et al.]. - 1a ed. - Ciudad Autónoma de Buenos Aires : Ediciones INTA, 2022.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-987-679-356-8

1. Horticultura. 2. Ordenamiento Territorial. I. Giobellina, Beatriz, prolog. II. Título.

CDD 635.04

Este documento es resultado del financiamiento otorgado por el Estado Nacional, por lo tanto, queda sujeto al cumplimiento de la Ley N° 26.899.

Se agradece el aporte de:

- Ministerio de Ciencia y Tecnología de la Provincia de Córdoba (MinCyT) - Convocatoria Proyectos de Investigación Orientados a Demanda (PIODO 2018) - Proyecto: Relevamiento geo-referenciado y caracterización de las unidades productivas de alimentos de proximidad del Cinturón Verde metropolitano de Córdoba. Finalizada en 2020.
- ADEC - Agencia de Desarrollo Económico de Córdoba - Convocatoria Fondo de Competitividad 2018 - Proyecto “Herramientas de caracterización espacial multiescala de las unidades productivas de alimentos de proximidad del Cinturón Verde de Córdoba”. Entidad Solicitante APRODUCO (Asociación de Productores Hortícolas). Entidades de apoyo técnico: INTA - Instituto Nacional de Tecnología Agropecuaria. O AUPA (Observatorio de Agricultura Urbana, Periurbana y Agroecología). CONAE- Instituto de Altos Estudios Espaciales “Mario Gulich” Universidad Nacional de Córdoba Comisión Nacional de Actividades Espaciales. Finalizada en 2019.

Se enmarca dentro de los Proyectos Estructurales de INTA:

- “Alternativas socio-agro-ambientales: prospectiva, observatorios y ordenamiento territorial para la sustentabilidad agroalimentaria” (PE I205)
- “Prevención y evaluación de la emergencia y desastre agropecuario” (PE I064)
- “Plataforma de Periurbanos” (PT I048)

Ese libro
cuenta con licencia:



Dedicatoria

*{ A mi madre María Eugenia Vargas, por quedarse a nuestro
lado y ser un apoyo incondicional, en este hermoso proceso de
aprendizaje. }*

Resumen

En esta tesis se exploraron las herramientas de sensado remoto y su procesamiento, para abordar las preguntas *¿Dónde?* y *¿Qué?* hortalizas se producen en el territorio periurbano. El caso de estudio es el territorio periurbano de Córdoba Capital y localidades aledañas.

Para responder *¿Dónde* se producen hortalizas en el periurbano?, se siguieron y compararon dos aproximaciones. Para la primera, se construyó una metodología de clasificación de pixel jerárquica, basada en información espectral, textural y de contexto, de imágenes Sentinel-2. En la segunda, se modificó y ejecutó una cadena de procesamiento de análisis de imagen basado en objetos geográficos (GEOBIA), la misma se aplicó tanto a imágenes Sentinel-2 como a imágenes de alta resolución espacial (VHRI). A fin de abordar la pregunta de *¿Qué* hortalizas se producen?, se planificó y operó vuelos de un vehículo aéreo no tripulado (VANT), se procesaron las imágenes resultantes de una cámara Parrot Sequoia+ y se clasificaron en un esquema GEOBIA.

Se obtuvo como resultados una metodología de clasificación espectral-textural (automatizable) de horticultura periurbana, basada en tecnología de acceso libre, así como un identificador de zona de interfaz rural-urbana (i.e. periurbano). La clasificación espectral-textural del cinturón verde de Córdoba Capital, obtuvo una precisión general de $OA = 0,89$ y un índice $\kappa = 0,86$, y las métricas de precisión para la clase Hortícola alcanzaron un puntaje $FIScore = 0,9$. Fruto de la tesis se cuenta con una cadena de procesamiento GEOBIA, operativa que logró identificar lotes hortícolas, caracterizarlos espectral, textural y morfológicamente, para clasificarlos con una precisión general de $OA = 0,75$ y un $\kappa = 0,64$, sobre una relación muestra/población del 60%. Dicha cadena se aplicó también, al análisis de las imágenes de resolución espacial de 10 cm de pixel, para obtener una clasificación de cultivos hortícolas por especie. En dicha clasificación se logra una precisión sobre los cultivos hortícolas de 0,69 y un $\kappa_{horticola} = 0,58$, con buen resultado en Brócoli (en cosecha) con $FIScore = 0,65$.

A partir de los resultados y experiencia adquirida en las metodologías exploradas, se espera que la segmentación de líneas de cultivo y plantas individuales de especies hortícolas, junto con modelos de estimación de biomasa y detección de estrés de los cultivos, sea la base para la construcción de mapas de rendimiento hortícola y del servicio de abastecimiento de alimentos frescos.

Palabras clave: Horticultura, GEOBIA, VANT, Interfaz Rural-Urbana, Texturas.

Contenido

Dedicatoria	3
Resumen	4
Agradecimientos	8
Prólogo I	9
Prólogo II	11
Prólogo III	13
Prólogo IV.....	15
CAPÍTULO 1 - Geociencias y teledetección en horticultura	17
1.1. Introducción	17
1.2. Importancia de la gestión hortícola	19
1.2.1. Suministro de alimentos.....	19
1.2.2. Servicios ecosistémicos de regulación y moderación	22
1.3. Horticultura y teledetección.....	23
1.3.1. Experiencias de teledetección en horticultura	23
1.3.1.1. A escala de paisaje	24
1.3.1.2. A escala de lote.....	27
1.3.2. Desafíos y retos.....	29
1.4. Motivación y objetivo de tesis.....	33
CAPÍTULO 2 - Clasificación de zona hortícola basada en información espectral y textural a nivel de píxel	35
2.1. Introducción	35
2.2. Materiales y métodos.....	36
2.2.1. Área de Estudio	36
2.2.2. Imágenes Satelitales.....	38
2.2.3. Algoritmos de Clasificación	41
2.2.4. Definición de verdad de campo (VC) y clases de cobertura del suelo	42
2.2.5. Evaluación de la precisión de clasificación	42
2.2.6. Estrategias de análisis de contexto y textura	44
2.2.6.1. Máscara de cobertura "urbana"	44
2.2.6.2. Máscara de interfaz rural-urbana	46
2.2.7. Clasificación Jerárquica. Producto Final	47
2.3. Resultados y Discusión.....	48
2.3.1. Evaluación de la precisión de la clasificación.....	48
2.3.2. Evaluación de áreas por tipo de cobertura de suelo	50
2.3.3. Resultados de Clasificación Jerárquica. Producto Final.	55
2.4. Conclusión	59

CAPÍTULO 3 - Clasificación de zona hortícola basada en GEOBIA (Geographic Object-Based Image Analysis)

62

3.1. Introducción	62
3.1.1. Problema, antecedentes y objetivo	62
3.1.2. Requerimientos de usuario	63
3.1.3. Requerimientos de software.....	64
3.1.3.1. Entradas.....	64
3.1.3.2. Salidas	64
3.1.4. Segmentación.....	65
3.1.4.1. Algoritmos de segmentación.....	67
3.1.4.1.1. Basados en Bordes	67
3.1.4.1.2. Basado en Regiones	68
3.1.4.2. Parámetros de los algoritmos de segmentación	72
3.1.4.3. Evaluación de segmentación	73
3.1.4.3.1. Evaluación por Suma de Espindola	74
3.1.4.3.2. Evaluación por F de Johnson	75
3.2. Materiales y Métodos	76
3.2.1. Cadena de procesamiento	76
3.2.1.1. Módulo: Generación de mosaicos o tiling	77
3.2.1.2. Módulo: Segmentación	78
3.2.1.2.1. Fuente de datos para segmentación.	78
3.2.1.3. Módulo: Clasificación de segmentos	85
3.2.1.3.1. Fuente de datos para clasificación	85
3.2.1.3.2. Caracterización de segmentos	85
3.2.1.3.3. Verdad de campo y validación cruzada	85
3.2.1.3.4. Modelos de clasificación y esquema votación	86
3.3. Resultados y Discusiones.....	86
3.3.1. Resultado de segmentaciones.....	86
3.3.2. Resultados de clasificaciones.....	91
3.3.3. Discusión de resultados	94
3.4. Conclusión	99

CAPÍTULO 4 - Clasificación hortícola en base a imágenes de altaresolución espacial y uso de Vehículo Aéreo No Tripulado (VANT)

101

4.1. Introducción	101
4.1.1. VANT en Agricultura	101
4.2. Materiales y Métodos	103
4.2.1. Instrumento y sensor: eBee SQ y Parrot Sequoia+	104
4.2.2. Plan de vuelo y procesamiento	106
4.2.3. Imágenes Parrot Sequoia+.....	109
4.2.4. Segmentación imagen multiespectral Parrot Sequoia+.....	111
4.2.5. Clasificación imagen Parrot Sequoia+.....	114
4.2.6. Verdad de campo.....	114
4.2.7. Separabilidad de clases sobre imagen multiespectral de Parrot Sequoia+	114
4.2.8. Caracterización de los segmentos.....	115

4.2.9. Clasificación de los segmentos	116
4.3. Resultados.....	117
4.3.1. Segmentación imagen Parrot Sequoia+	117
4.3.2. Clasificación imagen Parrot Sequoia+	119
4.4. Discusión y Conclusiones	121

CAPÍTULO 5 - Conclusión **125**

BIBLIOGRAFÍA	129
ANEXO A	139
ANEXO B	141
ANEXO C	143
Acrónimos	147
Glosario	148

Agradecimientos

Esta Tesis ha sido posible gracias a la contribución de muchas personas que aportaron conocimientos, experiencia, curiosidad, generosidad, calidez humana y paciencia. A la ternura y el apoyo de mi familia, en especial de *Tati*.

Agradecimiento especial al equipo de investigadoras e investigadores, profesores y profesoras, estudiantes y ex-maestrandos del Instituto Gulich (IG), colaboradores activos y una fuente de incentivo continuo. Al equipo de ciencia de datos del IG y compañeras de la MAIE, que leyeron la tesis o escucharon las ideas todavía sin pulir y aportaron calma y sugerencias para mejorar el resultado. A los revisores de la tesis que con dedicación y profundo conocimiento sugirieron mejoras para esta tesis y mi formación profesional.

Al equipo de trabajo del Observatorio de Agricultura Urbana Periurbana y Agroecología (O-AUPA) del Instituto Nacional de Tecnología Agropecuaria (INTA), por motivar este estudio e integrar conocimientos.

A mis directores de doctorado, Beatriz y Marcelo, que mantienen la templanza, la mirada en el futuro, la visión crítica y la dulzura de los grandes guías. A mi co-director Nicolás, por las palabras, la guía y la invitación a conocer el IG y la MAIE. A la investigadora, profesora y tutora Valeria Rulloni por su constante paciencia y cariño.

A productoras y productores hortícolas que permitieron el ingreso a sus propiedades y brindaron su amabilidad y generosidad durante las jornadas de muestreo. Por el diálogo de saberes y el compañerismo.

La tesis se encuentra en el marco de trabajo, y en parte financiada, por el proyecto "Herramientas de caracterización espacial multiescala de las unidades productivas de alimentos de proximidad del cinturón verde de Córdoba (CVC)", de la Federación Agencia para el Desarrollo Económico de la Ciudad de Córdoba (ADEC), en articulación con APRODUCO (Asociación de Productores Hortícolas), INTA Agencia de Extensión Rural Córdoba, O-AUPA y Instituto de Altos Estudios Espaciales "Mario Gulich" (UNC-CONAE). Coordinado por la Dra Beatriz Giobellina y el Dr. Diego Pons. Agradezco su estímulo y apoyo.

Al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), y el Instituto Nacional de Tecnología Agropecuaria (INTA) por la beca interna doctoral que permitió este estudio.

Prólogo I

Es una felicidad presentar esta publicación en Ediciones INTA que recoge la tesis de **Maestría en Aplicaciones de Información Espacial** de María Victoria Marinelli. Viki para quienes trabajamos con ella y la apreciamos como persona y como profesional.

Viki se incorporó al O-AUPA (Observatorio de Agricultura Urbana, Periurbana y Agroecología) en 2018, aportando una ya valiosa experiencia en extensión agronómica y trabajo de base con productores en países latinoamericanos, como Haití y Brasil, así como en la provincia de Córdoba, principalmente en el Arco Noroeste. Entre sus herramientas profesionales también contaba con algunas habilidades y conocimiento de teledetección, que desarrollaba en equipo con Nicolás Mari y Diego Pons de INTA. Fue su perfil profesional diverso el que la hizo merecedora de la Beca INTA-CONICET que gestionamos para el O-AUPA. Esa beca doctoral es la que le permitió integrarse al equipo interdisciplinario que estábamos formando, abocado a estudiar los procesos territoriales y las dinámicas que afectan al Cinturón Verde de Córdoba, contando a partir de ese momento con la colaboración en su dirección del Dr. Marcelo Scavuzzo, del Instituto de Altos Estudios Espaciales Mario Gulich (IG).

El O-AUPA impulsa desde 2013, fecha de inicio de actividades, una visión integral, transdisciplinaria y multisectorial en un trabajo de Investigación-Acción Participativa que procura aportar a políticas públicas que desarrollen y protejan el área de regadío, principal abastecedora de alimentos frescos de proximidad del área Metropolitana de Córdoba. Para ello, trabaja en red con otras instituciones públicas, organizaciones, grupos de investigación del sistema público de Ciencia y Técnica, asociaciones de productores, entre otros.

Herramientas analíticas de valoración y cuantificación de la producción hortícola basada en sensores remotos es un desarrollo aplicado que estábamos necesitando y pidiendo en el O-AUPA, que desde hace años pregunta ¿Dónde y Qué se produce en el Cinturón Verde de Córdoba? La posibilidad de contar con esta tesis se dio gracias a esta excelente Maestría ofrecida por el IG y la Facultad de Matemática, Astronomía y Física de la Universidad Nacional de Córdoba, que permitió que Vicky la tomara no sólo como fin en sí misma, sino también en su proceso de formación hacia el doctorado. Esta tesis fue realizada en el marco de algunos proyectos que impulsa el O-AUPA en una colaboración entre INTA, el IG, Asociación de productores Hortí-

colas de Córdoba (APRODUCO), la Subsecretaría de Agricultura Familiar de la provincia de Córdoba (SSAF), equipos de investigación de la FCM y FCA de la UNC y de la UCC, entre otros. Proyectos que lograron financiación entre 2018 y 2019, tanto de Agencia de Desarrollo Económico de Córdoba (ADEC) como del Ministerio de Ciencia y Técnica de la provincia de Córdoba. La tesis aborda una forma de aproximación al mapeo, sistemático y operativo mediante algoritmos, del periurbano de Córdoba, donde se desarrollan múltiples actividades que representan diferentes usos de suelo. Debo resaltar que nos brinda la primera imagen integrada de un área que en el debate académico tiene muchas formas de mencionarse: periurbano, área rururbana, interfase rural-urbana, borde o frontera urbano-rural, periferia... Un territorio en estudio que empieza a ser concebido como una entidad que tiene reglas de funcionamiento y dinámicas propias, que es cambiante y diverso, que es una porción de territorio que no es ni urbano ni rural y que demanda nuevos marcos teóricos y herramientas para estudiarlo, comprenderlo, monitorearlo y definir acciones sobre él en función de generar ciudades y territorios más resilientes y sustentables. Un periurbano donde en la mayoría de las ciudades del mundo aún quedan relictos de los sistemas agro-alimentarios de proximidad, sistemas de regadío de valor incalculable que se están perdiendo aceleradamente, al ritmo del avance del agronegocio y del negocio inmobiliario, que disputa el agua y la tierra para otros usos diferentes al abastecimiento de alimentos a las poblaciones, que fueron su función histórica y principal.

Esta tesis no solo es un hito que certifica la formación de posgrado de una nueva profesional, quien ha adquirido las más altas capacidades en las tecnologías que están marcando el rumbo para los estudios de la Tierra y las funciones y actividades que en ellas se representan; también aporta una herramienta accesible que nos permitirá mapear los periurbanos productivos de lo que se conocía como los principales cinturones verdes de las ciudades argentinas, y poder hacer estudios comparados al tiempo que construir una mirada provincial y nacional integradoras, superadoras de la invisibilización de estos territorios estratégicos para el país. Es un desarrollo transferible a municipios y otras entidades públicas, que aporta nuevas herramientas para la planificación urbana, para el Ordenamiento Ambiental Territorial, para la promoción y protección de los sistemas de abastecimiento local y nacional de alimentos frescos de proximidad.

Córdoba 2020

Dra. Arq. Beatriz Giobellina

Co-Directora de Beca INTA-CONICET

Coordinadora del O-AUPA INTA / FAUD-UNC

Prólogo II

Para mi, tanto en lo personal como en lo profesional, es un gran honor y orgullo ser parte de la presentación de esta publicación de Ediciones INTA incluyendo la tesis de María Victoria Marinelli para optar por el título de la Universidad Nacional de Córdoba de Maestría en Aplicaciones de Información Espacial.

Esta tesis representa de manera acabada un conjunto de atributos que intentamos impulsar desde el Instituto de Altos Estudios Espaciales Mario Gulich (IG) como lo son, el impacto social, la excelencia, la innovación, la cooperación interinstitucional y la internacionalización. Todos y cada uno, abordados por Victoria en una notable conjunción de saberes, actitudes y compromisos.

Es así como durante su desarrollo el trabajo ha consolidado la relación entre el IG y el O-AUPA (Observatorio de Agricultura Urbana, Periurbana y Agroecología). La tesis ha generado valorables productos de teledetección enfocados en una problemática de alto impacto socio-ambiental como lo son los procesos territoriales y las dinámicas que afectan al Cinturón Verde de Córdoba, como área piloto para extenderse a otras metrópolis del país y región. Además consolida una colaboración con académicos italianos que han dado origen a publicaciones internacionales. Finalmente desde lo estrictamente metodológico la tesis aborda ampliamente cuestiones novedosas en la teledetección en horticultura como lo son el uso de las texturas de las imágenes para su clasificación, el uso de herramientas de clasificaciones basadas en objetos y adicionalmente la utilización de imágenes sub-métricas de drones en el abordaje del problema.

Es importante señalar también que la culminación de esta Maestría única en el país generada en forma conjunta entre el IG (CONAE-UNC) y la Facultad de Matemática, Astronomía, Física y computación de la UNC, ha permitido que Victoria sumara a su formación de base en agronomía con un perfil extensionista, una innumerable cantidad de herramientas matemáticas, físicas y de procesamiento de imágenes y datos que le abren un enorme perspectiva profesional hacia el futuro, que en lo inmediato se vinculan a su doctorado en Geomatica y Sistemas Espaciales.

Así y sobre la base de lo dicho anteriormente y las no menores dificultades que han significado para completar un trabajo de tesis en este año 2020 tan particular a nivel global, nuevamente, como director del IG y como director de tesis de Victoria, es

una satisfacción enorme compartirla este producto con la comunidad toda, tanto por su contenido técnico, científico y metodológico, como por lo que su desarrollo deja como ejemplos a seguir.

Córdoba 2020

Dr. Carlos Marcelo Scavuzzo

Director del IG

Director de Tesis

Prólogo III

La soberanía alimentaria es un derecho de los países que busca garantizar el abastecimiento local de alimentos, desarrollar las economías regionales y generar mecanismos de comercio justo que preserve a los productores agropecuarios a nivel local. Sin embargo, y como consecuencia de la enorme demanda internacional de commodities y de alimentos a nivel mundial, países productores como la Argentina, han destinado enormes superficies del territorio nacional para proveer y abastecer dicha demanda. De la mano del libre comercio y de políticas liberales, como también de la adopción de tecnologías que favorecieron la concentración de grandes aglomerados económicos, la agricultura industrial junto con la ganadería extensiva se propagó a lo largo y a lo ancho del país, generando múltiples impactos colaterales a nivel ecológico y a nivel socio-económico, atentando en última instancia contra la soberanía alimentaria del país. Los principales impactos ecológicos y ambientales que se evidencian, son los relacionados a la pérdida de biodiversidad y todos aquellos asociados a la pérdida de servicios eco-sistémicos, muchos de los cuales están directa o indirectamente relacionados a los cambios en el uso del suelo, principalmente debido al reemplazo de bosques por otras coberturas culturales, como también el crecimiento urbano sobre áreas naturales y entre otros disturbios de origen antrópico. Los impactos sociales y económicos se relacionan en general al cambio estructural que se produce como consecuencia de la concentración del capital, que expulsa a las pequeñas economías regionales, extinguiéndolas y generando procesos de migración hacia grandes centros urbanos, con el consecuente abandono del territorio rural, también, impactando en la capacidad productiva de generar alimentos diversos para el país.

Es notable como dichas problemáticas han recibido mayor atención a nivel de grandes ecosistemas, habiendo pocos estudios enfocados al diagnóstico de la agricultura peri urbana y su caracterización desde el punto de vista de la producción de alimentos sanos y cercanos a los centros de consumo, como son las grandes ciudades del interior del país. Es en este sentido que el trabajo de Victoria asume una gran relevancia actual, por tratarse de un estudio que busca caracterizar un fenómeno menos visibilizado, como es la pérdida de espacios productivos periurbanos.

Este trabajo aporta un importante acervo de información útil para la caracterización de ambientes periurbanos, con novedosas técnicas de relevamiento de información y diagnóstico, de suma importancia en el ámbito de la planificación y el ordenamiento

territorial. Desde el punto de vista metodológico el trabajo se destaca por la habilidad lograda para el manejo de distintas tecnologías de observación de la tierra, que permitieron lograr facilidades para la adquisición, procesamiento y obtención de información de calidad, precisa, consistente y de bajo costo operativo.

Sin dudas, el trabajo logrado por Victoria es un esfuerzo que sienta importantes antecedentes para una línea de investigación de altísima relevancia, que se espera, impacte positivamente en el ámbito de la producción de alimentos, y la planificación socio ambiental del territorio.

Córdoba 2020

Mgr. Nicolás Alejandro Mari

Investigador INTA

A.E.R Cruz del Eje, E.E.A Manfredi

Co-Director de Tesis

Prólogo IV

La presente tesis representa el punto de partida para la generación de cartografía detallada en áreas de interfase urbano rural, considerando usos de suelo productivos, así como uso urbano, infraestructura y vegetación nativa, mediante una metodología automatizable, útil para la implementación de algoritmos que generen periódicamente dicha información. La generación automática de este tipo de productos espaciales, derivados de imágenes satelitales de mediana resolución y acceso gratuito con software libre, abre la puerta a aplicaciones concretas en el monitoreo de planes de desarrollo urbano y rural, conservación de recursos naturales, evaluación de riesgos y provisión de servicios eco sistémicos.

Desde el punto de vista de la gestión integral de riesgos y la adaptación al cambio climático, contar con información detallada de estas zonas de interfase, permite identificar vulnerabilidades en cuanto a infraestructura, sistemas productivos y población en general. Asimismo, es un insumo indispensable para la cuantificación de eventos climáticos extremos de impacto local, como tormentas severas, granizo y heladas, información requerida para la aplicación de planes de financiación y equipamiento, y seguros agrícolas, destinados a la adaptación ante eventos extremos. Por otra parte, se explora la viabilidad de la aplicación de imágenes de alta resolución en la identificación de cultivos hortícolas, mediante Vehículos Aéreos No Tripulados (VANTs o drones), mostrando la potencialidad de este tipo de productos para el monitoreo de las superficies de cultivos intensivos. Contar con dicha información permitiría a los mercados que abastecen de alimentos a las ciudades, prever y programar los volúmenes de venta, precios y logística en forma más eficiente.

Desde el INTA y en particular desde el proyecto PE I064 “Prevención y Evaluación de la Emergencia y Desastre Agropecuario”, que me toca dirigir, estimo su aporte como trascendental en la gestión territorial del riesgo agroclimático, y la construcción de estrategias y medidas de adaptación a la variabilidad creciente de los eventos extremos, debido al cambio climático, tales como sistemas de alerta temprana e identificación de vulnerabilidades.

La tesis no solo presenta información detallada de la Ciudad de Córdoba y su área periurbana, sino que también presenta una metodología replicable a todo el territorio nacional, proponiendo una solución concreta al vacío de información referido a las

áreas periurbanas y en particular, a la caracterización de sistemas de producción de alimentos de proximidad, de trascendental importancia para la lucha contra el hambre, la soberanía alimentaria y el ordenamiento territorial.

Invito al lector a recorrer estas páginas, de lectura amena y precisa, e interiorizarse en esta problemática trascendental, más en tiempos de pandemia, analizando los avances metodológicos desarrollados en la presente tesis de maestría, no solo como un ejercicio académico, sino como fuente de solución a realidades similares en otras regiones alimentarias.

Córdoba 2020

Dr. Diego Hernán Pons

Coordinador PE I064

Prevención y evaluación de la emergencia y desastre agropecuario

Cambio Climático

Grupo Recursos Naturales y Manejo de cultivos, Área Agronomía

INTA EEA Manfredi - Córdoba

Capítulo 1

Geociencias y teledetección en horticultura

1.1. Introducción

La horticultura es la cultura del trabajo, la ciencia, la tecnología, las personas y las empresas que se dedican a la producción de hortalizas para el consumo. La horticultura es generalmente considerada como la técnica de cultivo de plantas que se desarrollan en los huertos. El término viene etimológicamente de las palabras latinas *hortus* ('jardín', 'huerto', 'planta') y *cultura* ('cultivo'), es decir, "cultivo en huertos". Según la Sociedad Internacional de Ciencias Hortícolas (ISHS) [3], la horticultura incluye la floricultura (producción de flores cortadas y plantas ornamentales), la olericultura (plantas oleráceas, es decir, hortalizas y verduras en general), la fruticultura y el cultivo de especies aromáticas o medicinales.

Los cultivos hortícolas desempeñan un papel importante en la mejora de la productividad de la tierra, la generación de empleo, el aumento de las exportaciones, la mejora de las condiciones económicas de los agricultores y empresarios, y la seguridad alimentaria y nutricional de la población [4]. Especialmente, la horticultura alrededor de las ciudades se llama *cinturón verde* (CV) o *área de producción de alimentos de proximidad* y hoy en día se identifican como una herramienta moderna de planificación urbana que contribuye a la seguridad alimentaria local, protege la integridad ecológica, conserva la biodiversidad, cuida la cantidad y calidad del agua local, proporciona áreas recreativas y constituye un freno a la expansión urbana [5; 6]. En particular, para una mejor gestión de los cultivos existentes y para aumentar la superficie dedicada a cultivos hortícolas, se necesita una base de datos actualizada y precisa para la planificación y la toma de decisiones sistemáticas.

La geociencias y la teledetección (GRS) son herramientas avanzadas que ayudan a recopilar y actualizar información para desarrollar ciencia para los planes de gestión. Muchos tipos de sensores, como los radiómetros de microondas, los medidores láser, los sensores magnéticos y las cámaras multiespectrales, recogen datos electromagnéticos para obtener información precisa y a gran escala sobre la superficie y la atmósfera de la Tierra. Recientemente, la mejora en términos de resoluciones espaciales¹, temporales² y radiométricas³; y la accesibilidad de los sensores aerotransportados o de los vehículos aéreos no tripulados (VANT) ha cambiado el desempeño y el enfoque de la teledetección hortícola. Debido a que estos datos e imágenes son digitales, pueden ser fácilmente cuantificados y manipulados usando computadoras, desarrollando procesos computacionales avanzados y software libres. Este último punto es clave en el avance actual de la ciencia de datos, mayor capacidad de cómputo y algoritmos de procesamiento y aprendizaje profundo (*deep learning*) [7; 8; 9; 10; 11].

La GRS se puede utilizar en los esfuerzos para reducir el riesgo (mediante estudios de vulnerabilidad climática y económica) y así minimizar los daños en la horticultura. Los mismos datos pueden ser analizados de diferentes maneras para diferentes aplicaciones. Varios estudios tienen como objetivo la identificación de cultivos, la estimación de áreas, la identificación de enfermedades y plagas, entre otros, utilizando datos satelitales. Sorprendentemente, una búsqueda en la base de datos “Scopus” incluyendo título, resumen y palabras clave como: *sensado AND remoto AND (horticultura OR hortícola OR vegetales)*⁴ arrojó sólo 521 resultados hasta diciembre de 2019, esto representa sólo el 2,7% del total de trabajos en teledetección agrícola (19.495)⁵. En los últimos 3 años se encontró sólo 50 trabajos publicados por año, y para señalar el hecho de que es un campo de aplicación muy nuevo en 2001 y 2002 podemos encontrar sólo 6 publicaciones en cada año en el mundo. China, EE.UU., India, Malasya e Italia son los cinco primeros países en términos de producción de publicaciones científicas en este campo [12].

1 Resolución espacial es una medida del objeto más pequeño que el sensor puede resolver, o el área del terreno fotografiada para el campo de visión instantáneo (IFOV) del sensor, o la dimensión lineal en el suelo representada por cada pixel.

2 Resolución temporal es una medida de la frecuencia con la que un satélite es capaz de obtener imágenes de una determinada área. También se denomina intervalo de revisita.

3 Resolución radiométrica hace referencia al número de niveles digitales utilizados para expresar los datos recogidos por el sensor.

4 La búsqueda se realizó en inglés con el siguiente formato: *remote AND sensing AND (horticultural OR horticulture OR vegetables)*

5 La búsqueda general se realizó en inglés: *remote AND sensing AND (agriculture OR agricultural)*

En este marco, este capítulo revisa brevemente el uso potencial de las técnicas de GRS en horticultura para explorar las herramientas disponibles para una gestión eficiente de los cultivos con el fin de mejorar la planificación y gestión territorial. Además, se examina la importancia de la horticultura desde varios puntos de vista.

1.2. Importancia de la gestión hortícola

1.2.1. Suministro de alimentos

La producción hortícola mundial alcanzó 298 millones de toneladas en el año 2018, presentando una tasa de crecimiento entre el año 1998 al 2018 de 66,7%. Esta mayor producción, se acompaña del aumento de la superficie cosechada con una tasa de 64,2%, en base a datos de FAOSTAT [13]. Según la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), en Argentina la producción hortícola es de 960 mil toneladas, es decir un 0,3% de lo producido a nivel mundial. La misma se realiza sobre 62.910 ha, alcanzando un rendimiento estimado de 1,53 kg/m². Estos valores dan cuenta de la oferta y no de la producción, dado que son obtenidos por Mercado Central de Buenos Aires (MCBA) y reportados a FAO. Se comprende que estos valores no contemplen las pérdidas intermedias (cosecha, transporte, almacenamiento y comercialización), subestimando así el suministro de alimentos fruto de la horticultura nacional. Más del 90% de la oferta nacional tiene como destino el mercado interno, tanto para consumo fresco como para materia prima para industria, y alrededor del 7,5% se destina a la exportación [14;15].

La Estrategia Mundial sobre Régimen Alimentario, Actividad Física y Salud de la OMS (2004) [16] y el Plan de Acción Mundial para la Prevención y el Control de las Enfermedades Crónicas No Transmisibles de la OMS (2013-2020) [17] piden a los gobiernos que adopten medidas para aumentar la disponibilidad, asequibilidad y consumo de frutas y hortalizas en la población. El objetivo es reducir el impacto de las enfermedades no transmisibles y proteger la salud. Una revisión internacional de alto nivel sobre el consumo de frutas y verduras y el riesgo de cáncer coordinada por la Agencia Internacional para la Investigación del Cáncer (IARC)⁶ concluyó que el consumo de frutas y verduras puede reducir el riesgo de cáncer, y en particular el cáncer

⁶ <http://www.iarc.fr/>

gastrointestinal. La IARC estima que la fracción prevenible de los cánceres debido a la ingesta insuficiente de frutas y verduras en todo el mundo oscila entre el 5 y el 12%, y entre el 20 y el 30% para los cánceres gastrointestinales. Además, la Asociación Americana del Corazón⁷, la Organización Mundial de la Salud (OMS) y otras organizaciones internacionales recomiendan que nuestra dieta incluya 5 o más porciones de frutas y verduras al día [16; 17]. En el caso de Argentina sólo el 6% de la población alcanza ese consumo mínimo. En la Tabla 1.1, se comparan consumo mínimo de frutas o verduras informados en la 4^a Encuesta Nacional de Factores de Riesgo [18], si bien estos datos muestran un leve aumento en los últimos 10 años; al ampliar el rango de años analizados a los últimos 20 años, el consumo de frutas disminuyó un 41% y el de hortalizas un 21% [18].

Tabla 1.1: Resultados de la variable “consumo de frutas o verduras al día” de la 4^a Encuesta Nacional de Factores de Riesgo comparación 2009-2018. (68)

	2009 (IC 95)	2013 (IC 95)	2018 (IC 95)
Consumo de al menos 5 porciones de frutas o verduras al día	4,8% (4,4 - 5,3)	4,9% (4,3 - 5,5)	6,0% (5,5 - 6,6)

Los Cinturones Verdes de la ciudad, son una zona de interfaz crítica entre lo rural y lo urbano, constituyen un territorio patrimonial [19], que alberga sistemas de producción de alimentos de proximidad, con baja huella ambiental asociada al transporte. Los CVs generan servicios ecosistémicos de alcance local, actuando en muchos casos como verdaderos pulmones verdes en áreas densamente pobladas, contribuyen a la captación de agua y a la regulación del agua del territorio urbano, albergan biodiversidad, carbono, recursos genéticos locales, y generan empleo para un número significativo de personas [20].

En los últimos 40 años, el estudio de la agricultura periurbana ha ganado importancia. En 1979, el Comité de Agricultura de la OCDE organizó en París un debate entre sus Estados miembros sobre “La agricultura en la planificación y gestión de las zonas periurbanas”. A partir de ahí se recupera la planificación estratégica, urbana y regional, pensando en estos espacios en conflicto con la amenaza de la desaparición de las tierras agrícolas y los espacios abiertos [21].

⁷ <https://www.international.heart.org/es>

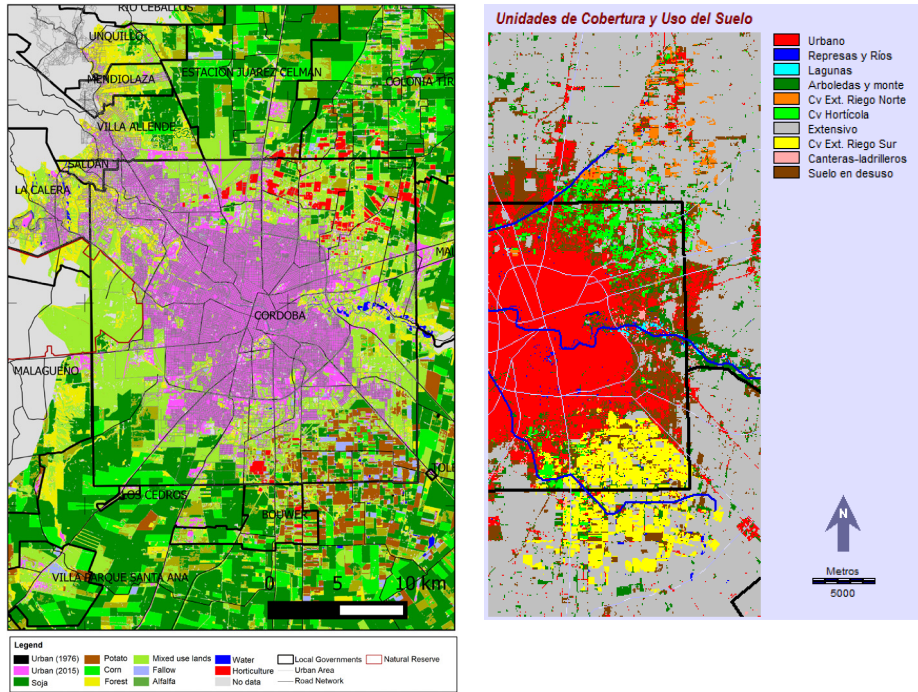


Figura 1.1: Mapas de cobertura y usos de suelo de la ciudad de Córdoba y zonas aledañas. Antecedentes sobre imágenes Spot de 2014 y 2015. Extraídos de [23] y [22].

Como ejemplo cercano, en la ciudad de Córdoba los últimos reportes de superficie hortícola cultivada basada en sensores remotos difieren entre 3.167 ha en 2014 (Barchuk *et al.*) [22] y 1.780 ha en 2015 (Mari *et al.*) [23], generando mapas de cobertura y uso (Figura 1.1)⁸. Se advierte una drástica disminución del área hortícola al comparar con los datos de Lanfranconi [24] para 1987 que indican se explotaban 28.238 hectáreas en el CV (norte, sur y Chacra de la Merced) donde más de la mitad (14.771 hectáreas) eran cultivos de hortalizas. En 2012, sólo 5.500 ha son reportadas por Fernández Lozano [25]. Los trabajos de Barchuk *et al.* y Mari *et al.* coinciden en que la disminución de la superficie hortícola se debe a un avance desmesurado de la frontera urbana [19; 22; 23], que crece sobre sistemas naturales y agrícolas estratégicos para la sostenibilidad de las ciudades. Según los datos de Mari *et al.* [23], se registra un aumento de superficie urbana de 6.700 ha entre 1974 y 2015 para toda la periferia de la Ciudad, significando un crecimiento del 46,5% en 40 años.

⁸ Las diferencias en las áreas reportadas responden a las diferentes metodologías usadas y serán analizadas en el capítulo 2 de esta tesis.

No existen suficientes datos en cuanto a la oferta de alimentos que el área productiva periurbana aporta a las ciudades. Esta falta de información propicia una planificación territorial deficiente, que no tiene en cuenta la seguridad alimentaria local.

El crecimiento demográfico, la gestión inadecuada de la tierra, el cambio climático, los fenómenos estocásticos extremos y los cambios en el uso de la tierra aumentan las tasas de erosión y degradación del suelo y reducen la cantidad y la calidad del agua disponible [26; 27]. Estos cambios generan impactos negativos en la salud humana y los ecosistemas naturales, disminuyen la producción agrícola e industrial y aumentan la aparición de conflictos sociales [28; 29]. Los ecosistemas productivos (dícese agroecosistemas o agrosistemas) sufren una severa degradación de sus recursos naturales y biodiversidad, principalmente debido al alto grado de fragilidad y difícil reversibilidad de algunos procesos biológicos y socioeconómicos afectados, que ponen en riesgo el suministro de alimentos y la seguridad alimentaria. Poco se ha hecho para mejorar la adaptabilidad de la agricultura industrial a los eventos climáticos extremos y cambiantes [30].

En general, los agrosistemas periurbanos se encuentran en una encrucijada donde los beneficios económicos a corto plazo están frecuentemente vinculados al deterioro de los recursos productivos a mediano y largo plazo [31]. En gran número de ciudades de países en desarrollo, los CV todavía conservan, la estrategia de producción fruti-hortícola al aire libre, que permite el uso de la teledetección para sensar su producción y analizar así sus relaciones ecosistémicas a nivel de paisaje.

1.2.2. Servicios ecosistémicos de regulación y moderación

Tanto el ‘suministro de alimentos’, como la ‘regulación biológica de plagas’, la ‘regulación de los caudales de agua’, la ‘moderación de fenómenos extremos’, entre otros, son servicios ecosistémicos que surgen de las propiedades emergentes de los agroecosistemas [1]. Los Servicios Ecosistémicos (SE) son los aspectos de los ecosistemas utilizados (activa o pasivamente) para producir bienestar humano [32]. Estos servicios interactúan entre sí; y con diferentes sistemas de producción.

En general, en las áreas hortícolas, es mayor la superficie de producción al

aire libre que bajo cubierta⁹. Entonces las herramientas de sensado remoto resultan significativas para generar información sobre los servicios de aprovisionamiento y regulación generados por los CVs, a efectos de su ordenación y planificación óptima. Aún así, geolocalizar y cuantificar las unidades productivas bajo cubierta en el territorio es una posibilidad desde el sensado remoto que permite analizar su dinámica espacial y la interacción con el medio [33], dado que estas unidades de producción utilizan mayor variedad de agroquímicos en su ciclo, especialmente vinculados al control de plagas animales (insectos, ácaros y/o nematodos) [34] afectando así el SE de ‘regulación biológica de plagas’. La regulación biológica de plagas es una forma de reducir el uso de plaguicidas y mejorar la biodiversidad [35; 36; 37], garantizar la producción de cultivos y el suministro de alimentos.

Es importante analizar y evaluar la regulación de los caudales de agua, dado que la agricultura es un gran consumidor de agua y tiene una influencia importante en ella [38; 39; 40]. La gestión de los agrosistemas puede contribuir a disminuir los efectos de las inundaciones y ser considerada un mecanismo de alcance regional para controlar los caudales de agua [41]. En los últimos años, la creciente variabilidad del clima ha provocado episodios de sequía cada vez más graves y frecuentes, lo que influye en los sistemas agrícolas en todas las fases de crecimiento y, en consecuencia, en el rendimiento de los cultivos [42]. Por tal motivo es de vital importancia desarrollar estrategias para mitigar los efectos de los eventos extremos [43; 44].

1.3. Horticultura y teledetección

1.3.1. Experiencias de teledetección en horticultura

Desde la década del 80’, una tecnología prometedora como la teledetección, junto a los SIG y las tecnologías de la información, constituyen un conjunto de herramientas convenientes para proporcionar grandes bases de datos sobre áreas de interés. La forma en que se presentan los avances en el área de teledetección hortícola, se basa en sus diferencias en cuanto a su escala de análisis (paisaje y lote) y sus objetivos. Así mismo en la Tabla 1.2 se com-

⁹ Superficie de suelo en producción ‘cubierta o semicubierta’ por plástico, vidrio, media sombra, malla antigranizo. Las especies hortícolas bajo esta tecnología requieren un uso y dependencia intensiva de los recursos y por ende mayor inversión inicial.

plementa c con la reseña de los avances en GRS hortícola según métodos y herramientas desarrolladas.

1.3.1.1. A escala de paisaje

Trabajar a escala de paisaje permite analizar y entender la variabilidad y dimensión espacial de un fenómeno ecológico; así como determinar el reparto de variación de los fenómenos a lo largo las escalas espaciales.

Detección y mapeo de horticultura:

El Departamento de Agricultura de los Estados Unidos (USDA), en 2011, produjo un mapa de cobertura de la tierra con formato raster, georeferenciado y específico para cada cultivo, llamado Cropland Data Layer (CDL) [45]. Sus entradas o fuentes de datos incluyen imágenes satelitales de resolución espacial media (AWiFS, Landsat TM y ETM, datos satelitales MODIS), y fuentes de datos auxiliares. Se utilizó un método de clasificación supervisado por árboles de decisión. En ella se reporta una clase única de verduras/frutas/nueces.

En 2014, Crnojevic *et al.* [46], proponen una clasificación de las tierras de cultivo basadas en píxeles de los pequeños campos agrícolas mediante la combinación de imágenes Landsat-8 y RapidEye, para el norte de Serbia. Se basa en una serie temporal de imágenes multiespectrales adquiridas a diferentes resoluciones. A pesar de que los cultivos mapeados son no hortícolas, el tamaño de los campos de Vojvodina es similar al de la horticultura, por lo que la metodología que se utilizó y la mejora atañen a este capítulo. A pesar de no tener especies hortícolas, este trabajo se destaca por la escala de lotes de cultivo identificados.

Desde 2017 a la actualidad, Gomez *et al.*, en el marco del proyecto europeo SENSAGRI¹⁰ producen un mapeo anual de cobertura con 122 clases en Castilla y León¹¹, basado en la fusión de imágenes de Sentinel-2, Landsat 8 y Deimos-1 y otros datos auxiliares. Utilizando el algoritmo C5.0, que genera árboles de decisión a partir de un conjunto de datos de referencia [47].

Inventarios de cosechas y las previsiones de producción:

¹⁰ <http://sensagri.eu/>

¹¹ http://mcsncyl.itacyl.es/visor_datos

Los estudios para inventariar, prever, estimar cosecha en horticultura se basan principalmente en sensores remotos de alta resolución o hiperespectrales y en el análisis de series temporales. En particular, los sensores hiperespectrales nos permiten medir la energía reflejada por el sol en numerosos canales (normalmente más de 200). De esta manera es posible construir con gran precisión la firma espectral de los blancos de interés y luego caracterizarlos radiométricamente. Como ejemplo de ello se muestra en la Figura 1.2 firmas espectrales de especies hortícolas a campo obtenidas con radiómetro, donde se puede observar claramente la posibilidad de diferenciarlas en determinados canales espectrales. Esta línea no ha sido aun profundizada para la detección y mapeo de la diversidad de especies hortícolas. El registro obtenido, sin embargo, abre la posibilidad de trabajo a futuro con algunos avances observables en [48]. Desafortunadamente, actualmente no existen plataformas satelitales operativas con sensores hiperespectrales, por lo que estas herramientas aún no están disponibles para su uso en la planificación y gestión de cultivos hortícolas.

En cuanto a sensores de alta resolución y análisis de series temporales, en 2004, Rao *et al.* [49] usó el Satélite de Teledetección de la India (IRS)-ID y sensor de auto-escaneado de imágenes lineales (LISS)-III de resolución espacial media (23,5-70,5 m), para estimar la producción de algunos cultivos de frutas y verduras cultivadas. Aplicaron un clasificador de máxima verosimilitud (MLC, por sus siglas en inglés), e interpretación visual de la textura del sensor PAN. Construyeron la respuesta anual del NDVI para cada especie estudiada a los fines de definir qué fecha era mejor para diferenciar cada cultivo pre-cosecha. Finalmente obtuvieron resultados en hectáreas que compararon con los provistos por la Subdirección de Agricultura de India, arrojando una sobre estimación de 0,3% del área ‘total de vegetales (papa, tomate y otros)’.

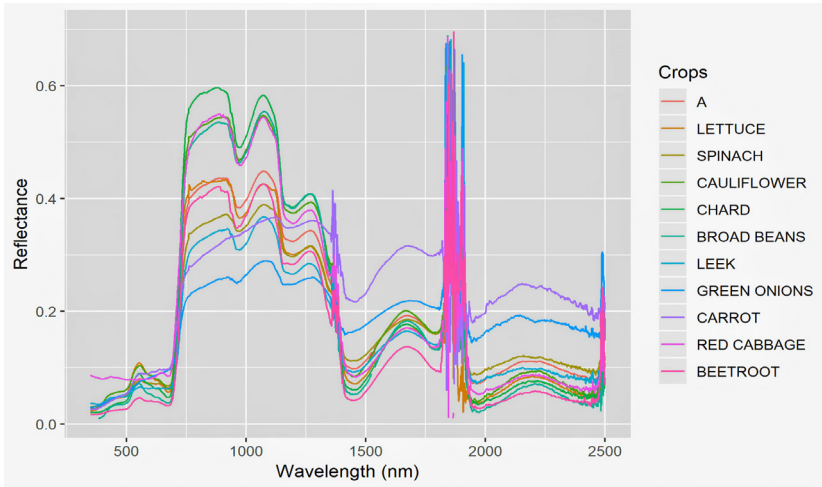


Figura 1.2: Firma espectral de especies hortícolas. Se muestra la media de mediciones por especie. Se captura 100 veces cada especie para poder construir su 'respuesta espectral' en el rango de 220 a 2500 nm con resolución de 1 nm. Adquisición propia CONAE [48]. 'A' es el promedio de respuesta espectral de las 10 hortícolas presentadas.

Resulta evidente que para estudiar la presencia y anticipar la cosecha de los cultivos hortícolas se precisa resoluciones espaciales y espectrales mayores a las de los sensores de libre acceso como MODIS, Landsat OLI 8, o Sentinel-2. Los sensores multiespectrales de alta resolución espacial o aquellos hiperespectrales montados en vehículos aéreos no tripulados (VANT o DRON) son una línea poco estudiada aún en agricultura a escala paisaje (sí a escala lote). Un motivo para esto, es que el procesamiento de mayor densidad de datos a escala paisaje requiere abordajes computacionales costosos o complejos como aquellos orientados a objetos o superpíxeles.

Evaluación de daños causados por sequía e inundaciones:

Usha & Singh [50], mencionan como un uso potencial, dado que se ha empleado para describir las condiciones del cambio climático en grandes áreas y las áreas de provisión de alimentos hortícolas podrían ser consideradas en el futuro.

Seguimiento y gestión del suelo:

En coordinación con la planificación territorial de la provisión de alimentos, Barchuk *et al.* y Mari *et al.* [22; 23], estudian la dinámica territorial y la distribución de los recursos en el área de producción hortícola periurbana de Córdoba, con imágenes satelitales SPOT 5. Estos estudios son analizados y detallados en el capítulo 2 de esta tesis. Así mismo, Baldini *et al.* realizó una clasificación multitemporal de Landsat en el partido de La Plata, parte importante del área hortícola de Buenos Aires [51]. En Kenia [52], estudia la dinámica y los conflictos reales entre tamaños y escalas socio-económicas de las empresas de producción hortícola. Este área de seguimiento y gestión del suelo, detecta la horticultura en el paisaje, y la analiza en interacción con otras coberturas, por ello se conforma en una herramienta necesaria para la planificación territorial.

1.3.1.2. A escala de lote

Monitoreo de cultivos:

Los trabajos de Jonnson *et al.*; Min *et al.*; Okamoto *et al.* y Panda *et al.*, [53; 54; 55; 56] son algunos de los que ofrecen las ventajas del monitoreo de cultivos desde el sensado remoto. En particular el trabajo de Smart *et al.* [57] muestra el monitoreo de rendimiento, desde el manejo del dosel en viñedos. Dado que la GRS y los índices derivados son útiles para detectar el dosel, su estado y densidad, y esto es elemental para mejorar la producción y/o la calidad del vino, reducir la incidencia de enfermedades y facilitar la mecanización. Con imágenes RGB y estereoscópicas, segmentación y modelos de calidad vegetal, Parsons *et al.* [58] proponen y mejoran la evaluación en decisiones complejas en entornos de cultivos protegidos en invernaderos, a través de la programación y la gestión de la calidad de un cultivo vegetal en crecimiento.

Gestión del riego:

O'Connell *et al.* [59], trabajó con Landsat 5 TM, encontró relaciones consistentes de evapotranspiración (ET) - NDVI en cultivos de uvas, almendras y cítricos, demuestra el valor de usar GRS como una herramienta práctica para

mejorar los sistemas de riego y el manejo del agua de riego. Una revisión bibliográfica sobre el aporte a la gestión del riego desde las GRS es Alvino & Marino de 2017 [60].

Crecimiento y salud de las plantas (enfermedades, estrés y plagas):

Las técnicas de detección a corta distancia y de teledetección han demostrado un alto potencial en la detección de enfermedades y en el seguimiento de las parcelas de cultivo de las áreas con plantas infectadas [61]. Los tipos de sensores más prometedores son la termografía, la fluorescencia clorofílica y los sensores hiperespectrales [62; 63; 50].

Para la gestión de malezas y horticultura de precisión, un estudio de Okamoto *et al.* en 2008 [55], probó una cámara hiperespectral portátil, y un entorno de software con un modelo orientado a detección de objetos, logrando identificar cultivos, suelo y malezas; su caso de estudio fue la remolacha y cítricos verdes. Es un estudio a nivel de lote, basado en un sistema de captura de imágenes de bajo costo y el marco un software versátil orientado a objetos para imágenes hiperespectrales multipropósito. Un algoritmo de separación para poder trabajar a futuro con imágenes hiperespectrales de todo el lote. Lamentablemente no reporta las medidas de precisión que obtuvieron. En este área son vastos los estudios para agricultura extensiva (cereales y oleaginosas) [64; 65; 66; 67]. El manejo de cultivos sitio-específicos (SSCM, por sus siglas en inglés) es una faceta de la agricultura de precisión, ampliamente aplicada en la agricultura de cultivos en hilera, pero tiene muy poca aplicación en horticultura. Por ello Panda *et al.* [56] revisó aquellos que utilizaban satélites de alta resolución, hiperespectrales, LIDARs, VANT, etc.; y modelos espaciales sistemas de información geográfico (SIG) para huertos frutales.

Todos los trabajos mencionados se examinan brevemente en la Tabla 1.2, según herramientas usadas y línea de investigación que les permitió un avance en el aporte al estudio de la horticultura (olericultura) desde el GRS. Se puede consultar la revisión de Usha *et al.* [50], que contempla el estudio de GRS en cultivos de frutas y nueces.

Tabla 1.2: Reseña de los avances de la teledetección usados o potencial uso en horticultura. Producción propia basada en referencias citadas.

Avances	Herramientas Usadas	Referencias
Sensores	Alta Resolución Espacial	Spot (CNES)/ IKONOS/Pleiades/Planet/Rapid Eye/Deimos/National Agricultural Imagery Programme (NAIP)/ Quick Bird/VANT
	Libre Acceso	Sentinel (ESA) Landsat (NASA)
Índices	Índices	NDVI (Rouse et al., 1974) (128) NIR/Red ratio (Jordan, 1969) (74) Soil-adjusted VI (Huete, 1988) (65)
	Banda de Adquisición	NIR/Green Red Edge
Procesamiento Computacional	Fusión de Datos (en base a:)	SAR:Terra SAR X/ Sentinel 1/ COSMO Skymed/SAOCOM (CONAE)
		Hyperspectral
		LIDAR
	Machine Learning	Las técnicas avanzadas de clasificación de imágenes, como ISODATA, WARD, la técnica de mínima varianza, agrupamiento de k-medias, SOM, algoritmo de clasificación bayesiano, Fuzzy ARTMAP, análisis de agrupamiento multivariado difuso y redes neuronales de retro-propagación y técnicas de mapas auto-organizados también se utilizan para delinear cultivos hortícolas en imágenes satelitales.

1.3.2. Desafíos y retos

El sensado remoto de la actividad hortícola desde plataformas satelitales es un desafío para los investigadores debido a las características fenológicas, diversidad, tamaño de las unidades productivas, ubicación, y manejo agrícola de la horticultura. Algunas de sus características particulares son el **contexto** y la **fragmentación**. La horticultura generalmente se establece cerca de un río, en las márgenes de una ciudad o en los valles intermontanos, lo que dificulta la observación de la diferencia dentro y fuera de los huertos. Además de

eso, el **tamaño o escala** de los lotes (Figura 1.3) y la **diversidad de cultivos** (Figura 1.4) que componen los huertos, interfieren en la recepción espectral del sensor y hacen aún más importante que éste tenga una alta resolución espacial, espectral y radiométrica [50].



Figura 1.3: La escala y diversidad de cultivos, de los campos en el área de interfaz rural-urbano en el periurbano cordobés. Referencia 100m (fuente de la imagen: Fotogrametría, IGN, 2016)

Los retos a los que se enfrenta la tecnología de sensado remoto (GRS) en la horticultura se pueden resumir en dar respuesta a: **¿Dónde?**, **¿Qué?** y **¿Cuánto?** se produce. Al responder a **DÓNDE**, se trata de trazar un mapa del área de cultivo hortícola en una región determinada, por ejemplo, el cinturón verde de una gran ciudad. Dado que el cultivo hortícola es básicamente un uso de la tierra y no una cubierta vegetal, no tiene una respuesta radiométrica única, por lo que su clasificación no es una cuestión sencilla. La respuesta al **QUÉ**, es aumentar el nivel de complejidad que interesa conocer, es decir, diferenciar el tipo de cultivo hortícola. Básicamente significa saber si es lechuga, repollo, remolacha u otros. Estas diferentes especies de hortalizas pueden tener diferentes firmas espectrales, pero la escala espacial de cada cultivo es demasiado pequeña para la teledetección multiespectral o hiperespectral y la resolución espacial, métrica o sub-métrica, necesaria no es hasta ahora fácilmente accesible. Finalmente, y aún más complejo, es tratar de cuantificar la producción

de cada cultivo. Es decir, CUÁNTO o cuántos kilogramos por metro cuadrado (kg/m^2) de un determinado vegetal están disponibles para ser cosechados. Otro desafío importante a la hora de usar GRS en horticultura es la diferencia en las **fechas de cosecha**, derivada de la diversidad de cultivos y de las estrategias de mercado. Las etapas de crecimiento y la fenología de muchos cultivos hortícolas tienden a ser difíciles de generalizar como resultado de las grandes diferencias en las variedades, las densidades de plantación y las prácticas de cultivo [50]. A pesar de ello, la GRS es una herramienta útil para identificar las fechas óptimas de cosecha [49; 55]. Sin embargo, la dependencia de la energía solar por parte de estos sistemas ópticos, limita la adquisición de imágenes en condiciones **nubladas o nebulosas** (Figura 1.4). Las imágenes adquiridas durante estos períodos son normalmente de poca utilidad en la cartografía debido a la alta cobertura de nubes/granizo. Actualmente, hay mayor disponibilidad de los sistemas de radar de apertura sintética (SAR) que son casi independientes de las condiciones meteorológicas. A diferencia de los sensores ópticos, los sistemas de radar activos tienen su propia fuente de energía, transmitiendo ondas de radio y recibiendo los ecos reflejados de los objetos en la superficie de la Tierra [85]. Un creciente número de estudios ha explorado la fusión de datos SAR con datos ópticos de alta resolución y datos auxiliares para superar este problema [50; 69], lamentablemente estos trabajos no están aplicados a cultivos hortícolas. Por ejemplo, Forkuor *et al.* indican que la integración de los datos RapidEye y TerraSAR-X mejoró la precisión de clasificación en un 10% y 15% sobre el uso de RapidEye solamente, y precisan que si los datos ópticos y SAR están disponibles para toda la temporada de cultivo, se pueden lograr precisiones de clasificación de hasta el 75%.



Figura 1.4: Fotografía de campos de cultivos hortícolas diversos, en día nublado (izq.).
Radiómetro de campo: herramienta para mediciones espectrales, usado en horticultura (der.).

Otra cuestión es **la falta de datos de campo**, datos de base sobre la existencia, ubicación y producción de las unidades hortícolas, estos datos resultan esenciales para la planificación hortícola y el ordenamiento territorial. Los datos de relevamiento de campo y su consiguiente seguimiento temporal resultan costosos en tiempo, personas idóneas y fondos para adquirirlos. En este escenario, el carácter multi-fecha de las imágenes satelitales permite realizar un seguimiento de las características dinámicas del paisaje y, por lo tanto, proporciona un medio para detectar cambios importantes en la cubierta terrestre y cuantificar las tasas de cambio [50]. El conocimiento previo y la información auxiliar obtenida por la GRS y los sensores no destructivos de suelos y plantas pueden utilizarse para diseñar protocolos que proporcionen un equilibrio adecuado entre la precisión estadística y el costo [86]. Sin embargo, para entrenamiento y validación algunas herramientas y modelos a menudo requieren una gran cantidad de datos de campo, y de gestión, que son laboriosos, caros de obtener y/o requieren una calibración específica en el sitio [32].

Herramientas más avanzadas y accesibles, desde una perspectiva computacional, abren más oportunidades de utilizar los nuevos datos de teledetección para aplicaciones operativas en la horticultura. En particular, metodologías como el análisis de imágenes basado en objetos (OBIA), y el aprendizaje profundo y automático que se utiliza habitualmente en el campo de la visión por ordenador (*computer vision*, en inglés) están estrechamente relacionadas con

el acceso a imágenes de muy alta resolución espacial (VHRI, por sus siglas en inglés, *very high spatial resolution images*) con tamaños de píxel inferiores al metro, por ejemplo, de sensores montados en VANT. En la Figura 1.5 se presenta una imagen de una cámara RGB sobre un VANT sobre un huerto de brócoli. Esta imagen tiene una resolución espacial de 2,5 cm donde cada planta puede ser vista individualmente, por lo que herramientas tales como las redes neuronales convolucionales, utilizadas en *computer vision*, y fácilmente accesibles desde, por ejemplo, librerías de procesamiento de imágenes en Python, podrían ser aplicadas aquí para un conteo exhaustivo del número de plantas de este cultivo.

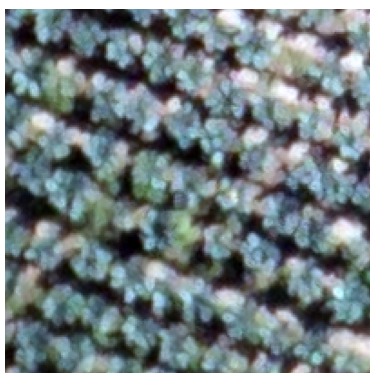


Figura 1.5: Imagen de VANT RGB en un huerto de brócoli. Adquisición propia. Escala 1:20.

1.4. Motivación y objetivo de tesis

Aunque los sensores remotos tienen muchas ventajas para determinar y estudiar los principales cultivos, esos estudios no han incluido la mayoría de las especies hortícolas [87]. Es sorprendente cómo a pesar de la importancia de la horticultura en la seguridad alimentaria, la salud, los aspectos sociales y laborales; las herramientas geoespaciales se utilizan sólo marginalmente en la gestión operativa de los huertos.

En consonancia con estos dos aspectos mencionados anteriormente, la producción científica en el ámbito de las aplicaciones de la GRS en la horticultura, como se muestra aquí, sigue siendo muy limitada e incipiente. Sin embargo, la GRS en horticultura tiene un futuro prometedor en cuanto a sensores y plataformas con mayor resolución, imágenes de colección de libre acceso (*ej.*

Sentinel-ESA y Landsat-NASA), sensores montados en aviones, VANT, la potencia del procesamiento computacional, fusión de datos, mayores accesos a grandes bases de datos digitales, y la información histórica de rendimiento. Estos avances y nuevas metodologías sumado al contexto de ‘seguridad humana’ desde el enfoque de las necesidades alimentarias y la gestión territorial, aportan una nueva perspectiva para analizar y estudiar la horticultura.

En este marco es que se formula el **objetivo general** de esta tesis de maestría. El objetivo es **explorar las herramientas de sensado remoto y su procesamiento para caracterizar la producción hortícola en los territorios periurbanos**, tomando Córdoba Capital como caso de estudio.

Siguiendo con esta línea directriz se plantean como **objetivos específicos**:

1. Identificar la horticultura periurbana a través de la generación de procedimientos de clasificación basada en píxeles incorporando características de contexto.
2. Implementar una cadena de procesamiento completo para generar clasificaciones basadas en objetos del área hortícola.
3. Explorar el uso de imágenes de muy alta resolución tomadas por un vehículo aéreo no tripulado (VANT) sobre cultivos hortícolas, incluyendo la programación de vuelos, su ejecución, el pre-procesamiento y la posterior clasificación de las imágenes obtenidas.

Capítulo 2

Clasificación de zona hortícola basada en información espectral y textural a nivel de píxel

2.1. Introducción

La cartografía continua de los cultivos hortícolas en zonas periurbanas, es muy importante para la gestión de la actividad hortícola por sus funciones en el suministro de alimentos, así como por otros servicios ecosistémicos que proporciona. El área periurbana destinada a cultivo de hortalizas disminuye drásticamente [88; 25; 89; 22], debido a un avance extralimitado de la frontera urbana, que crece sobre sistemas naturales y agrícolas estratégicos para la sostenibilidad de las ciudades [23].

En Córdoba, la historia de la identificación remota de la horticultura multiespecífica, en la zona de interfaz rural-urbano¹, muestra resultados que difieren entre 3.167 ha para 2014 [22], y 1.780 ha para 2015 [23]. Ambas investigaciones se basaron en imágenes Spot 5 con una resolución espacial de 10 m/píxel. Los algoritmos de clasificación utilizados son diferentes, así como el método de colecta de la verdad de campo. Los dos trabajos reportan estadísticas globales de precisión similares con κ de 0,75 y 0,76, respectivamente. Los métodos aplicados por Barchuk *et al.* [22] son clasificaciones *por píxel* y la asignación de la clase mayoritaria a los objetos obtenidos mediante un algoritmo de segmentación en el programa TerrSet™ (bajo licencia comercial). En el caso de Mari *et al.* [23] es una clasificación por píxel utilizando el algoritmo de aprendizaje automático Máquina de Soporte Vectorial o *Support Vector Machine* en inglés (SVM). En particular, para la horticultura (multiespecífica),

¹ El término “interfaz”, deriva del ámbito del desarrollo territorial y la planificación. El concepto surge de la traducción de interface (inglés) y en la bibliografía en español se traduce tanto como *interfaz* o *interfase*. En idioma castellano, según la Real Academia Española, la palabra *interfase* “se refiere a la superficie de separación entre dos fases. En esta tesis, se consideró que la raducción correcta de la palabra inglesa interface sería el término “interfaz”, que se refiere a una conexión física y funcional entre dos aparatos o sistemas independientes [90; 91]

[23] hizo una identificación visual y digitalización de las *parcelas hortícolas*, en Google Earth™, que luego se utiliza como máscara en la clasificación. Si bien ambas clasificaciones son recientes y contemporáneas, indican una diferencia de 1.387 ha, además de que los estudios se basaron en imágenes comerciales, o programas bajo licencia y los métodos usados fueron costosos dado que requirieron de numerosas salidas a campo y de muchas horas en gabinete para la identificación manual de lotes o corrección de la clasificación. Por ello se consideran difíciles de replicar y falta un método más adaptable.

En este contexto, el objetivo es presentar una clasificación basada en píxeles para la identificación y cuantificación de áreas de producción hortícola en la interfaz rural-urbano que rodea el área urbana de Córdoba. Se usan imágenes de resolución 10m/px, Sentinel-2, y algoritmos de aprendizaje automático *Support Vector Machines* (SVM) y *Random Forest* (RF). Con el fin de mejorar las experiencias anteriores, se propone el uso de información espectral y espacial, es decir, se añaden variables de textura y contexto a las bandas espectrales para describir los cambios en la vecindad de cada píxel. Todos los procesos tienen como objetivo acercarse a la escala de la horticultura y a su mapeo operativo y sistemático, utilizando imágenes de libre acceso y procesándolas en software abierto para generar un análisis automatizable y replicable.

2.2. Materiales y métodos

2.2.1. Área de Estudio

El estudio se llevó a cabo en un área de 180.000 ha, que incluye parte de la región metropolitana de Córdoba y se enfocó en comprender la zona histórica de riego y uso hortícola, circundante a la ciudad de Córdoba. El área delimitada al *oeste* coincide con el límite oeste del departamento capital y al *este* limita con el departamento Río Primero. Al *norte*, limita con la localidad de Estación General Paz, y al *sur* con las localidades de Lozada y Rafael García, a la altura de la ruta nacional C-45 (Figura 2.1).

El Cinturón Verde histórico de Córdoba está conformado por tres zonas definidas: sector *norte* regado por el Canal Maestro Norte que comprende el noreste del departamento Capital y parte del departamento Colón (Villa Retiro, Villa Esquiú, El Quebrachal, El Gateado, Guiñazú y parte de Colonia Tirolesa); el

sector *este* es la zona de Chacras de la Merced a las orillas del Río Suquía, que comprende parte de tres departamentos Capital, Colón y Santa María; y el sector *sur* ocupando parte del sureste del departamento Capital y parte del departamento de Santa María (camino a San Carlos, camino a 60 Cuadras, camino a San Antonio, Ferreyra y La Carbonada), históricamente definido por el sistema de riego Canal Maestro Sur, que derivaba del dique San Roque. Actualmente, la zona sur se abastece del dique Los Molinos.

La zona estudiada presenta una temperatura media anual de 17 °C, con amplitud térmica de 14 °C, manifestándose un período libre de heladas que alcanza los 270 días entre septiembre y mayo. Las precipitaciones oscilan entre 750 mm al oeste y 800 mm al este, con una distribución estacional de tipo monzónico presentándose un déficit hídrico de 180 mm al este y 240 mm al oeste de la ciudad de Córdoba [92]. Los vientos predominantes tienen dirección NE, con una velocidad media de 7 km/h [93].

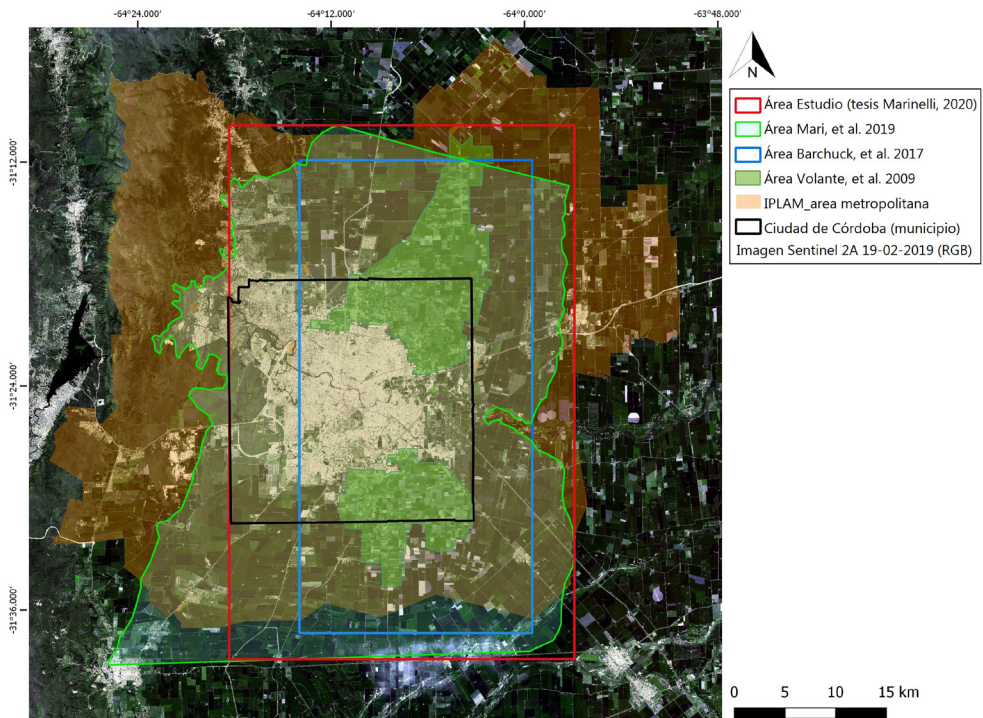


Figura 2.1: Área de estudio en el Cinturón Verde de Córdoba, comparación con áreas de estudio de trabajos antecedentes [94; 95; 22; 23].

La producción hortícola diversificada o multiespecífica se concentra en la zona periférica norte. Se caracteriza por tener establecimientos de tipo familiar con menos de 20 hectáreas cultivadas. La zona sur está dominada por productores de carácter emprendedor, poco diversificados y con uso intensivo de mano de obra mecanizada, en lotes superiores a 20 hectáreas, dedicados principalmente al cultivo de la papa, asociada a la zanahoria y maíz dulce [25].

Los principales cultivos anuales son la soja y el maíz (con la variante de invierno: trigo en rotación con soja) y se distribuyen en gran medida fuera del departamento de la capital, pero con una participación significativa en la interfaz rural urbana [89; 22; 23].

Dentro del área de estudio, el foco de esta tesis está puesto en las áreas hortícolas. Se sigue la definición de [96; 21; 25] que sostienen que las regiones alimentarias de proximidad (llamadas Cinturones Verdes) son el territorio ocupado por huertos familiares y comerciales que rodean a las ciudades, y donde se producen hortalizas para abastecer a la población urbana; que pertenecen al mismo sistema de riego común y cuya distancia al centro urbano no supera los 30 km.

2.2.2. Imágenes Satelitales

El área de estudio fue representada por un mosaico de dos imágenes (JML y JLL) Sentinel-2A (ESA) del 19/02/2019 preprocesadas a reflectancia de superficie. Se utilizaron las bandas azul (B2 ~ 493 nm), verde (B3 ~ 560 nm), rojo (B4 ~ 665 nm), infrarrojo cercano (B8 ~ 833 nm) con 10 m de resolución espacial, y 4 bandas estrechas en el dominio espectral de borde rojo de vegetación (B5 ~ 704 nm, B6 ~ 740 nm, B7 ~ 783 nm, B8a ~ 865 nm) y 2 bandas infrarrojas de onda corta, más anchas (B11 ~ 1610 nm y B12 ~ 2190 nm) remuestreadas bilinealmente a un píxel de 10 m.

Para incorporar información del contexto del píxel, se construyeron cinco características texturales-contextuales que se agregan a las 10 bandas espectrales de Sentinel-2A mencionadas para la clasificación. Estas características texturales-contextuales se basan en el *filtro de varianza del NDVI*. La va-

rianza calculada en una ventana de análisis o kernel de 3×3 fue binarizada para realzar la presencia de la varianza característica de los campos hortícolas (líneas de cultivo). Dicha binarización se realizó mediante un umbral que permitió diferenciar la varianza de zona hortícola, de la de zona urbana o rural (a partir de puntos de verdad de campo). Sobre la clasificación binaria se procesaron cinco filtros de media, para describir la proporción de esta varianza en contextos de 0,25 ha; 0,81 ha; 2,25 ha; 4,41 ha y 9,61 ha² (kernels de 5×5 ; 9×9 ; 15×15 ; 21×21 y 31×31 respectivamente). Como se observa en la Figura 2.2, estas características resaltan el patrón espacial que caracteriza a los lotes de producción hortícola.

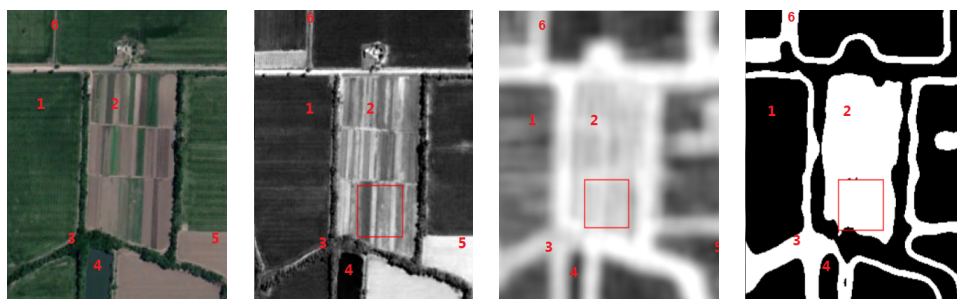


Figura 2.2: Características auxiliares. Gráfica ilustrativa del efecto buscado con la construcción de características auxiliares para identificar horticultura. (a) Imagen RGB, (b) NDVI, (c) Filtro de media sobre varianza NDVI, (d) Binarización sobre filtro de media. Se observa 1: cultivo extensivo (soja), 2: cultivos hortícolas, 3: arboleda, 4: represa de agua, 5: suelo arado, 6: borde de lote.

Se seleccionó una fecha de la segunda quincena de febrero (mes estival) para acentuar la diferencia entre los cultivos. Así los sembrados extensivos (maíz, soja, alfalfa) están en plena etapa vegetativa, con los lotes cubiertos en más de un 90%. En el caso de las plantaciones de papa, febrero es una de las dos fechas anuales de siembra, entonces todo suelo arado de superficie media (más de 10 ha) está sembrado con papa, y se visualizó claramente en la recolección de datos de campo. Los cultivos hortícolas se encuentran constantemente en una dinámica de siembra y cosecha, y en el periodo estival están en plena actividad vegetativa, de cosecha y recambio.

A los fines de mejorar el procesamiento y evaluar la exactitud, se trabajó con la definición de tres áreas: i) total (180 mil ha), ii) área con máscara de cobertura urbana y iii) interfaz ruralurbana. Cada área fue caracterizada con las 15

2 Superficie media de parcela o lote hortícola es 8,7 ha (144).

capas descriptas (10 bandas espectrales y 5 características textural-contextuales), para ser clasificadas (entrenadas y validadas) bajo los mismos modelos. Las áreas fueron generadas con distintos métodos (descritos en el apartado 2.2.6) a los fines de abordar las deficiencias de resultados preliminares de clasificación de la horticultura periurbana (Tabla 2.1).

Tabla 2.1: Métodos e información satelital seleccionada de Sentinel-2A del 19/02/2019, utilizados para caracterizar el Cinturón Verde de la ciudad de Córdoba (Argentina).

	Método	Bandas Sentinel-2A	Región Espectral	Long. de Onda Central (μm)	Resolución Espacial (m)
1) Máscara de cobertura urbana	Matemática de Bandas /	2	Azul	492,4	
	Filtro de valores y filtro de media /	3	Verde	559,8	10
	Binarización	4	Verde	664,6	
2) Máscara de Interfaz Rural-Urbana	Remuestreo / NDVI /	4	Rojo	664,6	10
	Filtro de varianza y filtro de media /	8a	NIR	864,7	20
	Selección de Rangos de Valores				
Todas las anteriores más:					
3) Cobertura y uso actual del suelo agrícola		5	Red Edge	704,1	
		6	Red Edge	740,5	20
	Clasificación Supervisada Random Forest (RF) y Support Vector Machine (SVM) /	7	Red Edge	782,8	
	Aplicación de máscaras y reclasificación	8	NIR	832,8	10
		11	SWIR	1613,7	
		12	SWIR	2202,4	20
más cinco características texturales-espectrales (NDVI)					10

Red Edge: Dominio espectral "Borde Rojo"vegetación (VNIR)

NIR: Infrarojo Cercano

SWIR: Infrarojo de onda corta

2.2.3. Algoritmos de Clasificación

Basado en los antecedentes de teledetección hortícola en Córdoba [22; 23], se usaron los modelos de clasificación supervisada *Random Forest* (RF) y *Support Vector Machine* (SVM). Para las clasificaciones se utilizaron algoritmos implementados en el software **R** [97], que usa el paquete *caret*³. Este paquete cuenta con un conjunto de funciones que tratan de simplificar el proceso de creación de modelos predictivos. Contiene herramientas para: división de datos, pre-procesamiento, selección de características, ajuste de modelos por remuestreo y otros.

En la tarea de clasificación, los SVM pertenecen a la categoría de clasificadores lineales, ya que inducen separadores lineales o hiperplanos de alta dimensión, ya sea en el espacio original de los datos de entrada, si éstos son separables o cuasi-separables (ruido), o en un espacio transformado (espacio de características), si los datos no son separables linealmente en el espacio original [98]. A diferencia de otros métodos de aprendizaje, el SVM minimiza el riesgo estructural. Separa las clases seleccionando un hiperplano de separación que es equidistante de los datos más cercanos de cada clase para alcanzar un margen máximo en cada lado del hiperplano. Para definir el hiperplano, sólo se consideran los datos de entrenamiento de cada clase que se encuentran en el límite de esos márgenes máximos, a los que se les llama vectores de apoyo [98]. El modelo *Support Vector Machine* se utilizó con un núcleo de funciones radial, el paquete *caret* utiliza la precisión para seleccionar el modelo óptimo para generar la predicción.

El otro modelo de clasificación supervisada utilizado es *Random Forest* (RF), un algoritmo de aprendizaje automático que ajusta iterativamente los modelos de árbol de regresión a subconjuntos aleatorios de los datos de entrada y utiliza el resultado combinado para la predicción [99]. Es decir un modelo RF está formado por un conjunto de árboles de decisión individuales, cada uno entrenado con una muestra ligeramente distinta de los datos de entrenamiento generada mediante (validación cruzada, k-iteraciones). La predicción de una nueva observación se obtiene agregando las predicciones de todos los árboles individuales que forman el modelo. El paquete *randomForest* se usa en **R** [100], e implementa el algoritmo clásico de Breiman [101]. Se entrenó el

³ *caret* por sus siglas en inglés Classification And REgression Training (Clasificación y Entrenamiento de Regresión)

modelo RF con el 70% de las muestras, 15 predictores, y las 10 clases; se remuestreó con validación cruzada de 10 carpetas y se utilizó la precisión para seleccionar el modelo óptimo utilizando el valor más grande.

2.2.4. Definición de verdad de campo (VC) y clases de cobertura del suelo

Se realizó la georeferenciación de 1.251 puntos de cobertura y usos del suelo como verdad de campo, los días 15 y 16 de febrero de 2019, con un GPS Garmin etrex 10. Los puntos se digitalizaron y procesaron en QGIS 2.18.25 [102]. El recorrido no fue preestablecido, se tomaron los puntos a los costados de caminos rurales, rutas y calles dentro de la zona histórica del sistema de riego. Se registraron todas las coberturas observadas con una distancia mínima de 15 metros.

Basado en el material bibliográfico [25; 103; 89; 22; 23], los puntos de VC registrados fueron agrupados en clases de cobertura de suelo, y se calcularon índices de separabilidad espectraltextural (Jeffries-Matusita y Divergencia transformada) entre pares de clases [104]. Estos valores van de 0 a 2, lo que indica cuán bien separadas estadísticamente están las clases seleccionadas. Para valores de separabilidad bajos (menores a 1), las clases se combinaron en una sola [105]. Todas las clases seleccionadas presentaron una separabilidad superior a 1,6.

Finalmente se proponen diez clases: Agua (represas y lagunas); Alfalfa; Arboleda (árboles en los bordes de los canales de riego, manchas de vegetación arbórea espontánea o implantada); Construcción (galpones y casas); Bajo Cubierta (cultivos bajo cubierta: antigranizo, mediasombra e invernaderos); Hortícola (cultivos vegetales diversificados); Maíz; Soja; Arados (arados, recién sembrados, en zona sur siembra de papa -verificados a campo-); Tierras de uso mixto (lotes para viviendas, campos sin uso, suelo desnudo -sin cobertura vegetal o artificial- y campos de fútbol).

2.2.5. Evaluación de la precisión de clasificación

El conjunto de los 1.251 píxeles de VC (Figura 2.3) se dividió por clase en 70% como datos de entrenamiento y un 30% de validación, con el fin de mantener los datos para la posterior prueba independiente. En este caso, la

biblioteca *caret* en R se asegura de que la distribución de las clases sea la misma en ambos conjuntos de datos.

Para determinar la exactitud de los mapas resultantes, se utilizaron las matrices de confusión, se calculó la Precisión Global (*OA* por sus siglas en inglés Overall Accuracy) y el coeficiente kappa (κ) [106], así como otras métricas de precisión por clase, de las que se destacan *Precision*, *Recall* y *F1 Score*⁴. La precisión global *OA*, se calcula con los píxeles correctamente clasificados sobre el total de datos de validación (píxeles). El κ calcula el acuerdo entre la imagen clasificada y la verdad del terreno, debido únicamente a la exactitud de la clasificación, suprimiendo el acuerdo que cabría esperar simplemente por casualidad [107]. En cuanto a las métricas de exactitud por clase, *Precision* es el número de predicciones positivas dividido por el número total de valores de clase positivos pronosticados. *Precision* puede considerarse como una medida de la exactitud de los clasificadores, es decir “precisión de usuario”. En cambio *Recall* es el número de verdaderos positivos dividido por el número total de elementos que realmente pertenecen a la clase positiva, es decir, la suma de verdaderos positivos y falsos negativos, que son elementos que no fueron etiquetados como pertenecientes a la clase positiva pero deberían haber sido. *Recall* se analiza como una medida de sensibilidad a los falsos negativos, es una “precisión de productor”. *F1 Score*, es calculado como $[2 * ((precision + recall)/(precision+recall))]$, transmite el equilibrio entre *precision* y *recall* [108].

⁴ Se utilizan el nombre de las métricas en inglés a los fines de no perder el sentido en la traducción y dado su uso generalizado en dicho idioma.

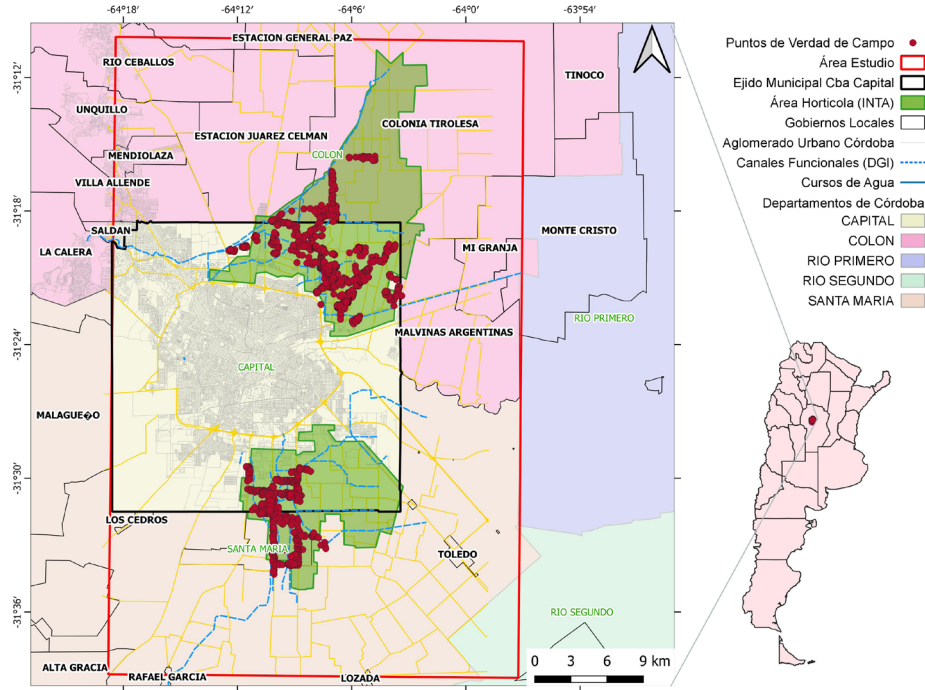


Figura 2.3: Mapa con geoposición de puntos de verdad de campo, canales funcionales del sistema de riego (Dirección General de Irrigación, DGI), referencia de área histórica de cultivos hortícolas (INTA) [94], y división política de Córdoba, en el área de estudio de la tesis.

2.2.6. Estrategias de análisis de contexto y textura

2.2.6.1. Máscara de cobertura “urbana”

Las construcciones en el contexto urbano de media densidad y en el periurbano de baja densidad representan un problema para los clasificadores a los fines de clasificar la horticultura periurbana. Esto se debe principalmente a que tales coberturas muestran “ruido” debido a píxeles de cobertura vegetal (jardines o arbolado público) en su interior, lo que disminuye la precisión de la clasificación. Este va depender de la resolución espacial de los datos usados, en este caso Sentinel-2 o cualquier otro de mediana resolución. En el caso de trabajar sobre imágenes de alta resolución espacial (métricas o sub-métricas), este “ruido” dentro de lo urbano se recomienda clasificar.

De acuerdo con el comportamiento de las construcciones, que muestran una alta reflectividad en el espectro visible, se calculó el *Brillo* mediante álgebra

de bandas con la fórmula $Brillo = [BandaAzul(B2) + BandaVerde(B3) + BandaRoja(B4)]/3$. En la imagen resultante se estableció un umbral de valores, basado en la selección de polígonos de muestra y puntos de verdad de campo, tomados en la ciudad y en construcciones periurbanas. Con este umbral se realizó una clasificación binaria de valor cero para la clase “construcciones”.

A fin de generar una máscara conectada o sin ruido, sobre la máscara de píxeles “construida”, se aplicó un filtro de media, con una ventana móvil de 11×11 píxeles, la dimensión de la vecindad es similar a 1 manzana en el centro urbano ($10.000 m^2$). Esto asignó a cada píxel un valor de *contexto*, donde el tamaño de la ventana establecía el vecindario de interés. Sobre estos elementos se aplicó una clasificación binaria, y se construyó una máscara para todos aquellos píxeles que en su vecindario (11×11 px) tenían al menos 4 vecinos *construidos*.

Con esta máscara urbana, se depuró la base de datos de la verdad de campo, dejando 1.015 píxeles, y se ejecutaron los dos clasificadores SVM y RF. La prueba se realizó con el 30% de los datos de la VC y además se calcularon las matrices de confusión, la precisión general *OA* y por clase, y el coeficiente κ . Se puede observar el proceso realizado en el diagrama de flujo: Figura 2.4.

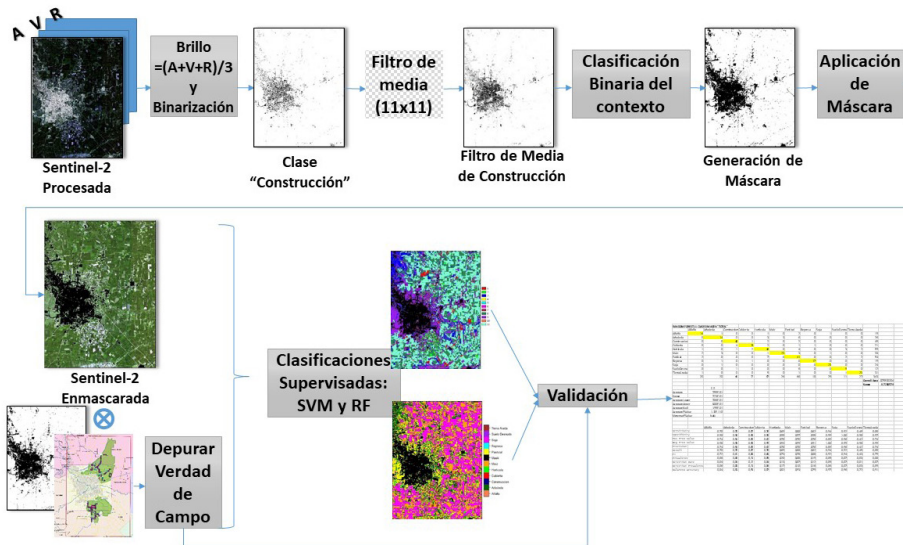


Figura 2.4: Diagrama de flujo de trabajo para generar máscara de cobertura urbana y clasificaciones supervisadas.

2.2.6.2. Máscara de interfaz rural-urbana

Se trabajó el análisis del patrón de ocupación del suelo. Se diferenciaron áreas: agrícola extensiva, hortícola y urbana, las mismas se distinguen por su organización sobre el terreno. La utilidad y el uso de la textura depende en gran medida de la escala de la intervención humana que creó la textura y de la resolución de las imágenes satelitales usadas.

El NDVI (Normalized Difference Vegetation Index) [70], ha demostrado ser un índice descriptivo eficaz para la detección de cultivos de regadío [87; 109; 40]. Se identificó el área hortícola periurbana, mediante el estudio de la variabilidad del NDVI en diferentes contextos o vecindades. Dicha variabilidad resalta el patrón de uso de suelo. Bajo la premisa de que las actividades productivas se aglomeran espacialmente debido a la distribución de un recurso o servicio (como agua de riego para horticultura, o líneas eléctricas y pavimento para urbanizaciones), esta aglomeración permite la extracción de esa porción para su estudio detallado. Por ello se extrajo el área de interfaz rural-urbano, donde la horticultura es predominante, aplicando un filtro de varianza del NDVI en una ventana móvil de 3×3 y el estudio del contexto del pixel en distintas vecindades.

Se observa que los lotes de horticultura presentan un *patrón* de alta varianza espacial en el NDVI, superior a la de los cultivos extensivos pero inferior al área urbana. En menos de 9 ha, tamaño medio de los lotes hortícolas [93], éste puede presentar varias líneas de cultivo de diferentes especies. Con el fin de interpretar la varianza del NDVI y con ello, definir *zonas*, se aplica un filtro de media lineal a la imagen *filtro de varianza del NDVI*, que asigna a cada píxel el valor de varianza en una vecindad de aproximadamente 4,5 ha (ventana móvil de 21×21 píxeles). Este procesamiento permitió definir rangos o umbrales de valores de varianza del NDVI, que discriminan entre patrones de uso: zonas urbanas, interfaz ruralurbano y agricultura extensiva. Se asignó valor cero a todos aquellos píxeles fuera del rango de valores que identifican la *interfaz rural-urbano*. Para la clasificación solo de la interfaz rural-urbana, se depuró la base de datos de verdad de campo, a 1.171 píxeles, y se ejecutaron los dos clasificadores SVM y RF. La validación se realizó con el 30% de los datos de la VC, por medio de las matrices de confusión calculadas, la precisión global OA y por clase, también el coeficiente κ , ver Figura 2.5.

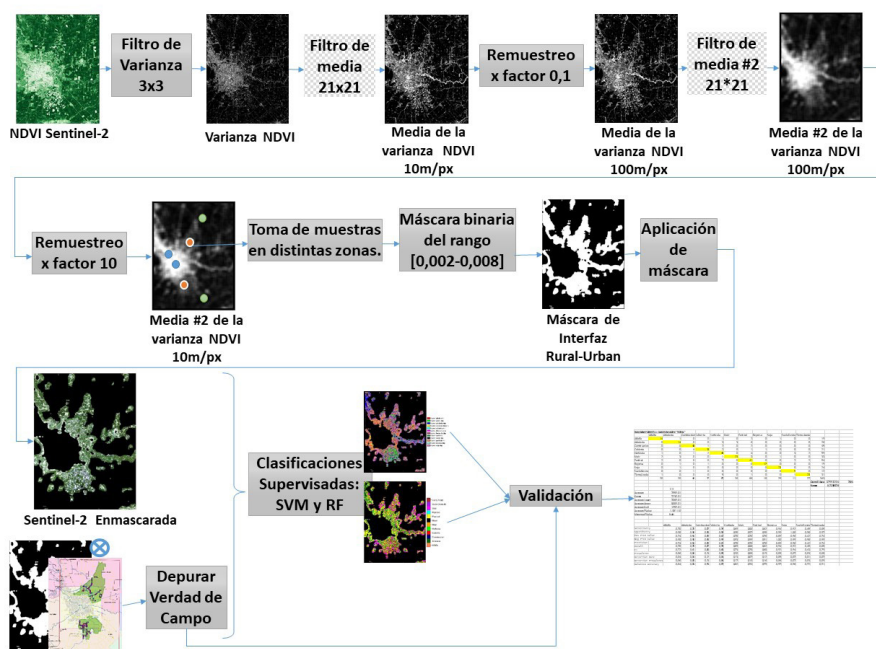


Figura 2.5: Diagrama de flujo de trabajo para generar máscara de interfaz rural-urbana y clasificaciones supervisadas.

2.2.7. Clasificación Jerárquica. Producto Final

Con el objetivo de mejorar la precisión de la detección de la clase Horticultura, se analizaron los resultados parciales de las tres áreas estudiadas comparativamente. A partir de los mejores resultados, se realizó una metodología de clasificación jerárquica que contempla el uso de las máscaras, interfaz rural-urbana y de cobertura urbana. En la clasificación RF de la interfaz rural-urbana, se extrajo la clase Horticultura y se seleccionaron 760 puntos. De los cuales 380 fueron *Hortícola Puro* (validado en la imagen fotogramática IGN 2016 y en los puntos de control de campo proporcionados por O-AUPA INTA) y los 380 puntos restantes fueron designados como *Bordes*, seleccionando aquellos que rodean los lotes agrícolas extensivos y los bordes de las avenidas. Se subclasificó (con un modelo RF) en las dos clases mencionadas, separando los puntos de entrenamiento (532 puntos) y validación (228 puntos). Los puntos de “Bordes” son reunificados con la clase Tierra de uso mixto. Así mismo se subclasificó la clase *Superficie cubierta*, con el mismo

método de selección de puntos y modelo. La clase residual se observó que pertenecían a áreas de lagunas y canteras del sureste del Río Suquía que habían sido asignadas erróneamente dicha clase, se reunificaron con la clase Agua.

Para el producto final se aplicó la máscara de uso urbano a la clasificación jerárquica obtenida en la interfaz rural-urbana, generando una caracterización continua de ese uso para las localidades o aglomerados urbanos periféricos al centro de Córdoba Capital.

La verdad de campo de 1.886 píxeles, se compuso de 715 píxeles identificados visualmente y digitalizados fruto de la sub-clasificación Horticultura, más los 1.171 píxeles del recorrido de VC. La clasificación se validó con el 30% por clase de los 1.886 píxeles totales. Se exhibe un diagrama de flujo de trabajo y método para la obtención del producto de clasificación jerárquica textural-espectral (Figura 2.6).

2.3. Resultados y Discusión

2.3.1. Evaluación de la precisión de la clasificación

En términos de las métricas de precisión (OA en % y κ como coeficiente de rango 0 a 1) de los clasificadores, *Random Forest* mostró mayor precisión en todas las pruebas: imagen total del área de estudio, imagen con máscara urbana e interfaz rural-urbano; la comparación se muestra en la Tabla 2.2.

Tabla 2.2: Mediciones de precisión global OA y coeficiente κ de clasificadores SVM y RF para las 3 pruebas realizadas: imagen total, imagen con máscara urbana, interfaz rural-urbana.

Métricas de Precisión	RF Total	SVM Total	RF Máscara Urbana	SVM Máscara Urbana	RF Interfaz	SVM Interfaz
OA (%)	75	71	77	70	77	64
kappa	0,71	0,66	0,74	0,65	0,73	0,58

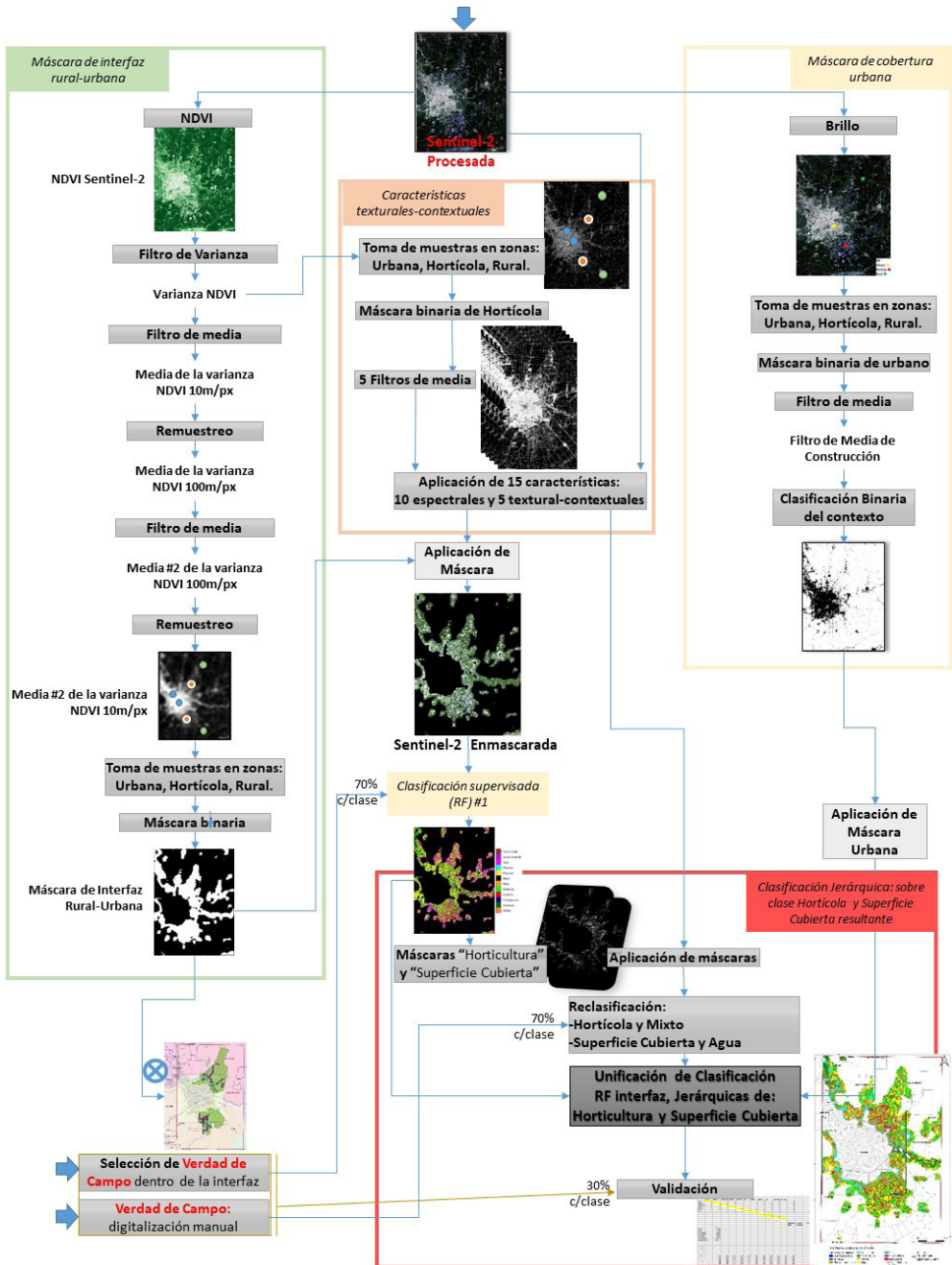


Figura 2.6: Diagrama de flujo de trabajo y método para la obtención del producto de clasificación jerárquica textural-espectral de la interfaz rural-urbana.

Siguiendo la escala de interpretación del coeficiente kappa, propuesta por Monserud *et al.* (96), los resultados de las clasificaciones RF fueron “muy buenos” (Tabla 2.3).

Tabla 2.3: Interpretación del Coeficiente kappa K

Valores κ	Clasificación
<a 0,40	Pobres
0,40 - 0,55	Suficientes
0,55 - 0,70	Buenos
0,70 - 0,85	Muy Buenos
>0,85	Excelentes

En particular, con la clasificación de RF en el área de la interfaz rural-urbano, se obtiene un OA de 77% y un índice κ de 0,73. Esta clasificación fue la elegida, siendo la de mayor precisión sobre la clase Horticultura con *precisión* de 0,82, *recall* de 0,72 y *F1 Score* de 0,77. En la Tabla 2.4 se comparan las métricas de la clase Horticultura de las tres áreas clasificadas.

Tabla 2.4: Métricas de Precisión, Recall y F1 Score de la clase Hortícola, obtenidas en las tres áreas clasificadas: imagen total, imagen con máscara urbana e interfaz rural-urbana.

	Clase Horticultura		
	RF Total	RF sin Urbano	RF interfaz
Precisión	0,74	0,82	0,82
Recall	0,69	0,71	0,72
F1 Score	0,71	0,76	0,77

2.3.2. Evaluación de áreas por tipo de cobertura de suelo

Se describen los resultados de las áreas de cada clase de cobertura y uso de suelo, para las clasificaciones de RF total, RF de máscara urbana y la de interfaz de RF rural-urbano, con el fin de mostrar las superficies clasificadas.

Las dos primeras clasificaciones, imagen total (180.109 ha) e imagen con máscara urbana (141.082 ha), indicaron que la clase Soja es la que ocupa más superficie con aproximadamente 50.000 ha que representan el 28% del área estudiada; una magnitud coincidente con las estimaciones de 2014 y

2015 [22; 23]. Es seguida por Maíz con 37.019 ha, 21% del total analizado, y Tierra de uso mixto y Arboleda con ocupaciones del 14% y 10% del área estudiada, respectivamente. Dentro de la clase Tierra de uso mixto se encuentran: lotes de viviendas, campos no utilizados, terrenos baldíos o sin cultivar y campos de fútbol. Esta clase conforma un conjunto heterogéneo que representa los espacios del territorio en tensión, tanto por la presión inmobiliaria para urbanizar, como por el avance de la agricultura industrial (soja, maíz, trigo y papa) [89; 22; 23]. Dadas las características de esta clase vinculadas al suelo de uso urbano, al aplicar la máscara se redujo la Tierra de uso mixto a 9% de ocupación (Tabla 2.5).

Tabla 2.5: Superficie en hectáreas y porcentajes de ocupación, obtenidos en las tres áreas clasificadas, RF total: clasificación del área de estudio centro de Córdoba; RF sin Urbano: clasificación imagen total con máscara urbana; RF interfaz: clasificación del área interfaz rural-urbana.

	RF Total	%	RF sin Urbano	%	RF interfaz	%
Agua	572	0%	472	0%	345	1%
Arboleda	17.572	10%	19.444	11%	6.445	11%
T. uso mixto	24.762	14%	16.499	9%	12.738	22%
Construcción	16.075	9%	513	0%	2.349	4%
T. Arada	6.944	4%	4.347	2%	3.888	7%
Alfalfa	7.060	4%	6.771	4%	2.491	4%
Maíz	37.019	21%	35.039	19%	12.075	20%
Soja	50.775	28%	49.771	28%	9.500	16%
Sup. Cubierta	1.859	1%	1.706	1%	1.477	3%
Horticultura	17.470	10%	7.242	4%	7.604	13%
Máscara	0	0%	38.307	21%	0	0%
TOTAL	180.109	100%	180.109	100%	58.912	100%

La superficie destinada al *cultivo de papa* se situó en torno a las 6.200 ha. Esta superficie se identificó como Tierra Arada, debido a que las fechas de recolección de VC coinciden con la época de plantación de papa, como se observa en (a) y (b) de la Figura 2.7 tomadas en el recorrido de VC. El resultado de superficie de papa fue del orden de lo reportado en los estudios de 2014 y 2015 [23; 22]⁵. Sin embargo, esta clase mostró confusión con la horticultura, omitiendo píxeles de papa, asignándolos a horticultura. Estos errores se

⁵ Además coincidió con la superficie reportada para el centro de Córdoba, del cultivo de papa “de segunda” (aquella que se siembra en febrero) según la Cámara Papera de Córdoba (comunicación personal con el presidente del Consorcio de Riego Zona Sur).

observan en las matrices de confusión de todas las áreas, aún en la de RF interfaz rural-urbana de la Tabla 2.6.

La clase urbana resultó en un 9% del área demarcada como total área de estudio mientras que la máscara de uso urbano representa el 21% de dicha área. La importancia del uso de la máscara se explica dada la resolución espacial de Sentinel-2A, con la que se obtuvo una descripción de la cobertura detallada y heterogénea espectralmente (arboleda, construcción, ruta, baldíos, jardines) dentro del uso “urbano”. El clasificador de imagen RF total discriminó y asignó píxeles de Arboleda en parques y plazas; la arboleda urbana irrigada se confundió con la clase Horticultura; los lotes baldíos y los jardines de las áreas residenciales se caracterizaron como Tierras de uso mixto y Horticultura. Siendo estos y otros, espacios típicos de las ciudades, fue necesario evitar esta caracterización, dado que este estudio se enfoca en el área no urbana. Por tal motivo, se enmascaro la zona de uso urbano de modo permanente. Así se eliminó la posibilidad de que otras clases caractericen esta zona (Figura 2.8). Esta máscara, y su método de construcción, puede ser usada de forma inversa para estudios de ocupación de suelo urbano.



(a) Campo Sembrado con Papa (Zona Sur).



(b) Campo Sembrado con Papa (Zona Norte).



(c) Campo Hortícola Sembrado. Influencia Suelo.



(d) Campo Hortícola (multiespecies).

Figura 2.7: Registro fotográfico del recorrido de Verdad de Campo, sobre campos sembrados con papa y hortícolas multiespecies ubicados en el Cinturón Verde de Córdoba. Febrero 2019.

En la clasificación RF de interfaz rural-urbana, las clases de cultivos extensivos (Soja, Maíz, Alfalfa), y papa (Arada) disminuyeron su proporción y dominancia. Resultando las Tierras de uso mixto un 22%, Maíz un 20% y Soja un 16% del área de interfaz. Las clase Arboleda% representa un 11% manteniendo esta proporción en todas las áreas (total, sin urbano e interfaz). Esta dominancia de Tierra de uso mixto, caracterizó al espacio periurbano de Córdoba. La identificación de esta cobertura, y el estudio y caracterización de la misma es un área de gran importancia para el ordenamiento del territorio. Dado que estas tierras pueden ser manejadas como reservorio de biodiversidad y aprovisionamiento de servicios ecosistémicos (SE) de regulación, y poder así, hacer una transición exitosa hacia producciones más sustentables. La distribución de estos ambientes, su conectividad y la interacción de estos ambientes y los cultivos al aire libre, es sumamente importante para garantizar la provisión de SE [51].

Las clases de Agua y Bajo Cubierta mostraron resultados prometedores en la detección de estas áreas de interés para la gestión de los recursos hídricos y la producción hortícola. En el caso de la clase Bajo Cubierta, se deben considerar subdivisiones de los resultados o se deben agregar al modelo original etiquetas tales como *canteras* y *ladrilleros*, como lo hace el trabajo de [22], dado que píxeles de lagunas y canteras del sureste del Río Suquía son asignados erróneamente a la clase Bajo Cubierta. Por tal motivo, se realizó una subdivisión de la clase Bajo Cubierta para el producto final clasificado.

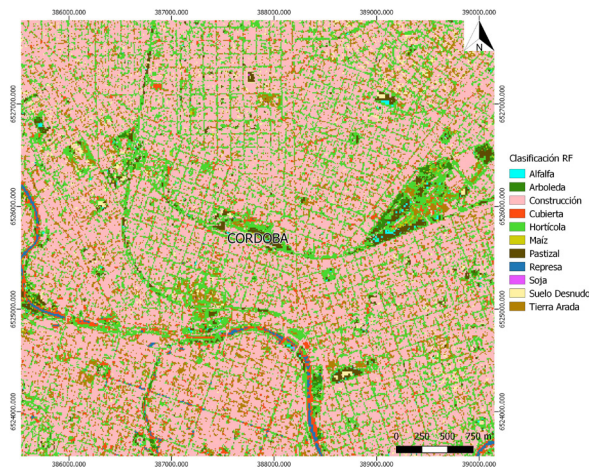


Figura 2.8: Zoom sobre el resultado de la clasificación Random Forest en contexto urbano. “Ruido” en contexto urbano.

Tabla 2.6: Matriz de Confusión de clasificación RF interfaz rural-urbano (valores en pixel).

Píxeles de clases asignadas en la imagen RF interfaz rural-urbano

		Agua	Arboleda	T. uso mixto	Construcción	T. Arada	Alfalfa	Maíz	Soja	Sup. Cubierta	Horticultura	TOTAL
P	Agua	16	0	0	0	0	0	1	0	1	1	19
i	Arboleda	0	20	2	0	0	0	1	0	1	1	25
x	T. uso mixto	0	3	45	1	0	2	4	0	0	8	63
e	Construcción	0	0	1	31	0	0	0	0	2	3	37
l	T. Arada	0	0	0	0	21	0	0	0	0	9	30
e	Alfalfa	0	2	5	0	0	11	0	0	0	1	19
s	Maíz	0	2	6	0	0	0	18	2	0	1	29
	Soja	0	1	0	0	0	0	0	20	0	0	21
V	Sup. Cubierta	0	0	2	2	0	0	0	0	18	1	23
C	Horticultura	0	1	4	0	5	0	1	0	3	65	79
	TOTAL	16	29	65	34	26	13	25	22	25	90	345

La clase Construcción, presentó una muy buena precisión en su detección (*Precisión* 0,84), el resultado se vio afectado por falsos positivos (errores de comisión) con píxeles asignados desde las clases Bajo Cubierta y T. usos Mixtos. En menor medida presentó errores por omisión con la clase Horticultura, Bajo Cubierta y T. usos Mixtos (Tabla 2.6). Construcción y Bajo Cubierta son dos coberturas que a nivel espectral y en el análisis de contexto pueden resultar similares, aún así los resultados para ambas fueron muy buenos (*F1 Score* 0,87 y 0,75, respectivamente) (Tabla 2.7).

La clase Horticultura fue sobre-estimada tanto en la clasificación de todo el área de estudio (17.385 ha), así como en el área con máscara de uso urbano (7.242 ha). A pesar de que, la clasificación RF interfaz rural-urbana generó una mayor precisión en la detección de la clase (*F1 Score* de 0,77), aún sobre-estimó la cantidad de hectáreas (7.604 ha). Estos resultados muestran que tanto las zonas densamente urbanizadas, las zonas suburbanas construidas y las zonas de cultivos extensivos afectaron negativamente a las clasificaciones de cultivos hortícolas de la interfaz rural-urbana.

Tabla 2.7: Precisión por clases, producto de clasificación Random Forest interfaz rural-urbana de Córdoba.

	Agua	Arboleda	Tierra de uso mixto	Construcción	Tierra arada	Alfalfa	Maíz	Soja	Sup. Cubierta	Horticultura
Sensitivity	1,00	0,69	0,69	0,91	0,81	0,85	0,72	0,91	0,72	0,72
Specificity	0,99	0,98	0,94	0,98	0,97	0,98	0,97	1,00	0,99	0,95
Pos Pred Value	0,84	0,80	0,71	0,84	0,70	0,58	0,62	0,95	0,78	0,82
Neg Pred Value	1,00	0,97	0,93	0,99	0,98	0,99	0,98	0,99	0,98	0,91
Precision	0,84	0,80	0,71	0,84	0,70	0,58	0,62	0,95	0,78	0,82
Recall	1,00	0,69	0,69	0,91	0,81	0,85	0,72	0,91	0,72	0,72
F1	0,91	0,74	0,70	0,87	0,75	0,69	0,67	0,93	0,75	0,77
Prevalence	0,04	0,08	0,18	0,09	0,07	0,04	0,07	0,06	0,07	0,25
Detection Rate	0,04	0,06	0,13	0,09	0,06	0,03	0,05	0,06	0,05	0,18
Detection Prevalence	0,05	0,07	0,18	0,10	0,08	0,05	0,08	0,06	0,06	0,22
Balanced Accuracy	1,00	0,84	0,82	0,95	0,89	0,91	0,84	0,95	0,85	0,84

Se observó exhaustivamente las áreas identificadas como *Horticultura* en la clasificación más precisa (RF de interfaz rural-urbano). Allí se encontraron errores de asignación, dado que en las 7.604 ha de clase Horticultura, el clasificador incluyó píxeles de límites de los campos y bordes de rutas, debido a la presencia de herbáceas en estos “bordes”. Esta asignación de “bordes” herbáceos a la clase Horticultura, fue causada por la inexistencia previa de esa clase en el modelo, pero especialmente de que las 5 características texturales-contextuales adicionales del modelo, singularizan las áreas hortícolas por la presencia de continuos “saltos” o “bordes” en la varianza del NDVI. Esto puede verse en la Figura 2.2, en la identificación 6 “borde”.

2.3.3. Resultados de Clasificación Jerárquica. Producto Final.

Las métricas de precisión global del producto de clasificación jerárquica en el área de interfaz rural-urbana fueron *OA* del 89% y un coeficiente κ de 0,86. Las precisiones por clase mejoraron de forma que superan al 0,82 y en particular las clases de interés hortícola presentaron medidas de F1 Score de 0,90 para Horticultura y 0,89 para Superficie Cubierta (Tabla 2.8). La clase Horticultura presentó errores de comisión (5/167) y omisión (22/189) respecto de la clase Tierra de uso mixto (Tabla 2.9).

En esta distribución final de las clases, la **Horticultura ocupa 1.649 ha**. Esta clase se caracteriza por la producción de hortalizas multiespecíficas en la interfaz rural-urbano o periurbana. Así mismo la clase Superficie cubierta reportó 839 ha. El resultado final de la clasificación jerárquica se puede ver en la Figura 2.9 y la distribución final de la superficie ocupada por clase en la Tabla 2.10.

A partir de resultados recientes del Censo Hortícola de la Ciudad de Córdoba⁶ (aun no publicados), se obtiene un conteo de hectáreas en producción hortícola de 1.062,22 ha solo del departamento Capital. Se procedió a medir la clasificación de la clase Horticultura en ese área lo que resultó en 1.025,26 ha. Dicho resultado conforma una validación empírica de la clasificación presentada.

Tabla 2.8: Precisión por clases, producto de clasificación Random Forest jerárquica de interfaz rural-urbana de Córdoba.

	Agua	Arboleda	T. Uso Mixto	Construcción	Tierra Arada	Alfalfa	Maíz	Soja	Sup. Cubierta	Horticultura
Sensitivity	1,00	0,89	0,85	0,68	0,96	0,93	0,97	0,89	0,88	0,97
Specificity	1,00	1,00	0,93	1,00	1,00	0,99	0,99	1,00	1,00	0,93
Pos Pred Value	0,95	0,89	0,84	1,00	0,93	0,82	0,86	0,89	0,91	0,84
Neg Pred Value	1,00	1,00	0,93	0,97	1,00	1,00	1,00	1,00	0,99	0,99
Precision	0,95	0,89	0,84	1,00	0,93	0,82	0,86	0,89	0,91	0,84
Recall	1,00	0,89	0,85	0,68	0,96	0,93	0,97	0,89	0,88	0,97
F1 Score	0,98	0,89	0,85	0,81	0,94	0,88	0,91	0,89	0,89	0,90
Prevalence	0,03	0,03	0,31	0,10	0,04	0,03	0,06	0,03	0,04	0,28
Detection Rate	0,03	0,03	0,26	0,07	0,04	0,02	0,05	0,03	0,04	0,27
Detection Prevalence	0,04	0,03	0,31	0,07	0,05	0,03	0,06	0,03	0,04	0,32
Balanced Accuracy	1,00	0,94	0,89	0,84	0,98	0,96	0,98	0,95	0,94	0,95

⁶ El Censo Hortícola es realizado por el proyecto “Herramientas de caracterización espacial multiescala de las unidades productivas de alimentos de proximidad del cinturón verde de Córdoba (CVC)”, financiado por la Federación Agencia para el Desarrollo Económico de la Ciudad de Córdoba (ADEC), en articulación con APRODUCO (Asociación de Productores Hortícolas), INTA (Instituto Nacional de Tecnología Agropecuaria) Agencia de Extensión Rural Córdoba - O AUPA (Observatorio de Agricultura Urbana, Periurbana y Agroecología) y CONAE (Comisión Nacional de Actividades Espaciales)- Instituto de Altos Estudios Espaciales “Mario Gulich” (UNC-CONAE).

Tabla 2.9: Matriz de Confusión de clasificación Random Forest interfaz rural-urbano jerárquica (enmascarado uso urbano), valores en píxeles.

Píxeles de clases asignadas en la imagen RF interfaz rural-urbano jerárquica (enmascarado uso urbano)

		Agua	Arboleda	T. uso mixto	Construcción	T. Arada	Alfalfa	Maíz	Soja	Sup. Cubierta	Horticultura	TOTAL
P	Agua	20	0	0	0	0	0	0	0	1	0	21
i	Arboleda	0	16	0	0	0	0	0	0	1	0	17
x	T. uso mixto	0	1	156	13	0	1	1	0	0	5	177
e	Construcción	0	0	0	41	0	0	0	0	0	0	41
l	T. Arada	0	0	0	0	25	0	0	0	0	0	25
e	Alfalfa	0	1	2	0	0	14	0	0	0	0	17
s	Maíz	0	0	2	1	0	0	32	2	0	0	37
	Soja	0	0	0	2	0	0	0	17	0	0	19
V	Sup. Cubierta	0	0	2	0	0	0	0	0	21	0	23
C	Horticultura	0	0	22	3	1	0	0	0	1	162	189
	TOTAL	20	18	184	60	26	15	33	19	24	167	566
	Overall Accuracy	89%										
	kappa	0,86										

Tabla 2.10: Superficie en hectáreas y porcentajes de ocupación, obtenidos en la clasificación espectral-textural basada en píxel con un modelo Random Forest jerárquico de interfaz rural-urbana final.

	RF jerárquica interfaz (ha)	%
Agua	342	1%
Arboleda	6.844	12%
T. uso mixto	14.467	25%
Construcción	5.866	10%
T. Arada	3.644	6%
Alfalfa	2.490	4%
Maíz	12.061	21%
Soja	9.499	16%
Sup. Cubierta	839	1%
Horticultura	1.649	3%
TOTAL	57.701	100%

El mapa de coberturas y usos de suelo generado para la interfaz rural-urbana de Córdoba del año 2019, brindó una línea de base concreta para analizar la situación actual y poder generar nueva información derivada de esta, que contribuya en un futuro proceso de planificación orientado al desarrollo de un territorio sustentable. Con una precisión global OA de 0,89 y un κ de 0,86. En este capítulo se logró cuantificar el área bajo producción hortícola de manera automatizable y mediante datos y programas de libre acceso, con una precisión de 0,90. Así como, se generó un método automatizable de identificación de interfaz rural-urbano y máscara de cobertura urbana.

La precisión alcanzada ($F1\ Score = 0,90$) en la identificación de horticultura diversa al aire libre mediante sensado remoto, en comparación con otros trabajos, es superior. Por ejemplo, Mari *et al.* [23] logró un $F1\ Score$ de 0,81 y Barchuk *et al.* [22] un $F1\ Score$ de 0,59 sobre el mismo área, pero con distintas metodologías y datos de base. También es superior a la clasificación multitemporal sobre un paisaje agrícola altamente fragmentado y heterogéneo realizado por Asgarian *et al.* [110]. En este último la clase Vegetales alcanzó un $F1\ Score$ de 0,80. Sin embargo, la precisión alcanzada en este capítulo, resulta inferior a otros recientes trabajos de identificación como el de Baldini [33] que obtuvo un $F1\ Score$ de 0,97 o Pulighe & Lupia [111] o Daga *et al.* [112], en los que las “áreas hortícolas a campo” se superpusieron como vectores al finalizar la clasificación. No obstante, los aportes de este capítulo resultaron más favorables que estos últimos, en cuanto a que aporta un método automatizable de la clasificación y es extrapolables a otras áreas. Así mismo, la verdad de campo y el método de recolección usado, requirió menor inversión en recorridos y/o tiempo de gabinete que los trabajos mencionados y es adaptable a métodos de recolección de datos de campo desde distintas fuentes, así como lo hace IDECOR [113] u otras plataformas de mapeos regionales [47]. Se abre la discusión en torno a la metodología de recolección colectiva y abierta de datos de verdad de campo.

La discusión en torno a la utilidad de un análisis multi-temporal para la identificación hortícola, no se abordó en este capítulo. Dado que la clase deseada es una clase de uso Hortícola y no un tipo de cultivo [114; 115; 116; 117; 118]. La horticultura de interfaz rural-urbana en Córdoba se compone de más de 30 especies y cada productor realiza en promedio 13 especies anualmente⁷. Por

⁷ Datos aún no publicados del Censo Hortícola realizado por el proyecto “Herramientas de caracterización espacial multiescala de las unidades productivas de alimentos de proximidad del cinturón verde de Córdoba (CVC)”.

ende la información que brinda un análisis multi-temporal en torno a la fenología de cultivo no fue necesaria, sino que se hizo incapié en el patrón espacial que caracteriza el uso hortícola periurbano. Incluso en pruebas preliminares, el uso de una serie temporal, entorpecía el resultado y fue más costoso computacionalmente. Se plantean como trabajos futuros para la detección multi-temporal de las coberturas, la clasificación sobre plataformas que operan con servidores en línea. Así como adjuntar características (bandas) provenientes de procesamientos sobre la serie temporal del NDVI (análisis de armónicos).

Las limitantes de la clasificación por píxeles, en base a sensores de media y alta resolución radica en encontrar una alta variabilidad de respuestas espectrales dentro de la clase que se desea identificar. Los lotes hortícolas cuentan con franjas de diferentes cultivos irrigados, en distintos estadios fenológicos, intercalados con tierra arada o bordes de plantas espontáneas. Esto implica encontrar en su interior distintas respuestas espectrales que pueden ser captadas por el sensado remoto, es decir, los lotes hortícolas están caracterizados por una alta variabilidad en su interior [23; 48]. Los errores de clasificación de píxeles de la clase hortícola, se deben principalmente, a píxeles no etiquetados como tales en el interior de lotes hortícolas. Esto debido a la presencia de tierra arada o a parcelas en descanso cubiertas de plantas de crecimiento espontáneo (i. e. malezas). Asimismo los errores de sobre-estimación de la clase, se ven en espacios verdes urbanos, bordes de lotes agrícola extensivos, arboledas de riberas (canales de irrigación) y banquetas⁸. Si bien, la aplicación de una máscara de uso urbano, suprime la asignación a la clase hortícola de píxeles correspondientes a arboleda urbana y espacios verdes (plazas, parques, otros), fuera de dicha máscara, la mayoría de los píxeles de bordes (de lotes y banquetas) son designados como clase hortícola. El uso de filtros de moda, mejora el resultado final, pero perjudica el nivel de detalle y la definición de límites de las coberturas.

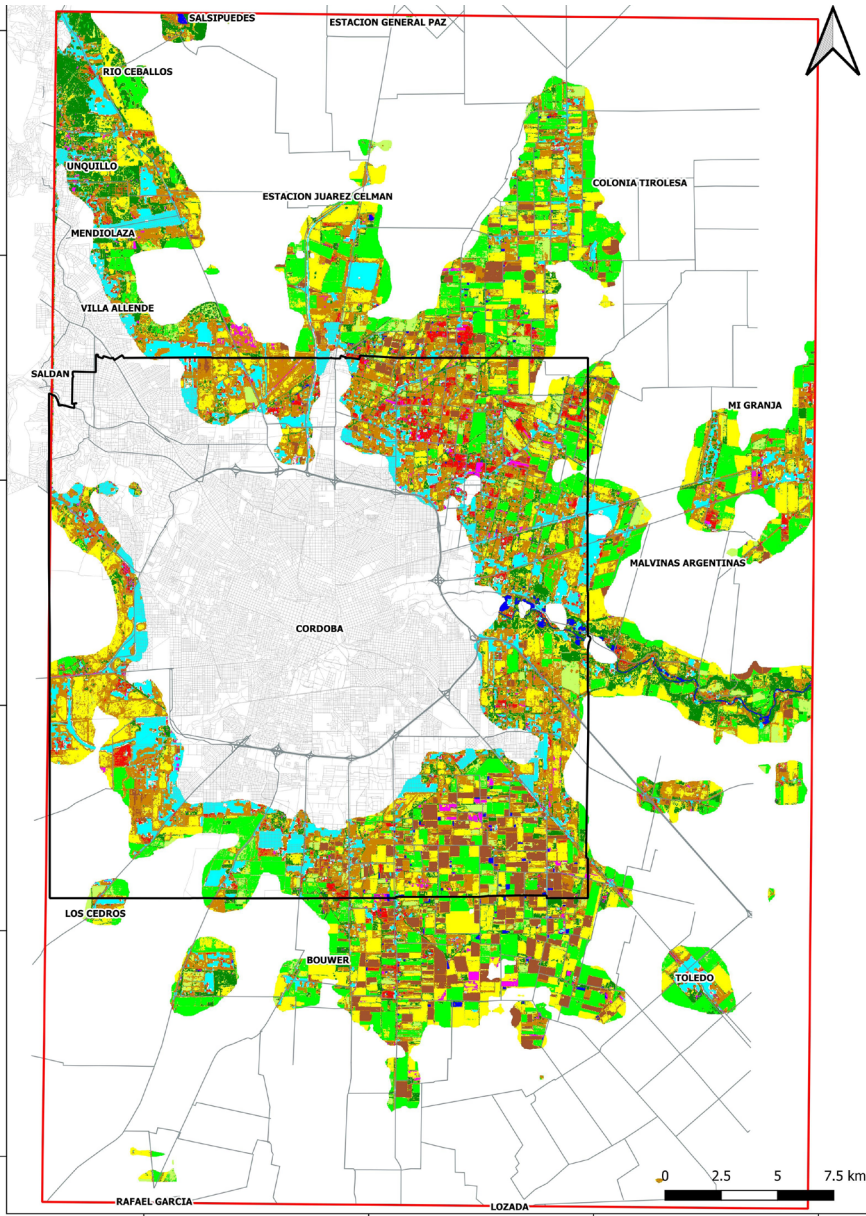
2.4. Conclusión

Este trabajo, aporta una respuesta a *¿Dónde?* se producen las hortícolas de proximidad a la ciudad, con una metodológica automatizable que permite el seguimiento a futuro de la identificación. A diferencia de los estudios antecedentes de clasificación hortícola del Cinturón Verde de Córdoba [22; 23],

⁸ Franja lateral nivelada entre el camino o ruta y el campo, donde los vehículos pueden detenerse.

se logra identificar y caracterizar la clase Horticultura desde un enfoque espectral integrado a la información de contexto de cada píxel. Se utiliza un abordaje jerárquico para resolver la sobre-estimación de la superficie total con cultivos hortícolas. Esta sobre-estimación se debe básicamente a que las características espectrales y texturales de la horticultura se mezclan con los bordes de otras coberturas con respuesta espectral similar. Además se destaca que si bien las métricas analíticas resultan convincentes, la validación con otras fuentes e incluso la mirada experta sobre el mapa generado, es de carácter esencial para definir si una clasificación es útil o no.

Como sub-producto de interés de este capítulo, se deriva el proceso por el cual se identifica la zona de interfaz rural-urbano, que permite su reproducción para detectar y cuantificar zonas periurbanas. La *máscara de interfaz rural-urbano*, permite la identificación de los procesos que ocurren en la zona periurbana de una ciudad. Identificación clave para los procesos de ordenamiento territorial. El trabajo muestra una herramienta de clasificación jerárquica basada en píxeles que, en base a sus resultados; la metodología y las herramientas utilizadas están en condiciones técnicas de ser operativas. El siguiente paso en esta trayectoria de mapeo hortícola es el uso del análisis de imágenes basado en objetos geográficos (GEOBIA) para el reconocimiento de patrones en las escenas periurbanas.



Interfaz Rural-Urbana de Córdoba

- | | | | |
|-----------------------------|--------------|-----------------|---------------------|
| Clasificación RF Jerárquica | Construcción | Soja | T Gobiernos locales |
| Cuerpos de Agua | Tierra Arada | Sup. Cubierta | □ Córdoba Capital |
| Arboleda | Alfalfa | Horticultura | — Aglomerado Urbano |
| Tierra de uso mixto | Maíz | Área de Estudio | — Red Vial |

Figura 2.9: Clasificación Random Forest jerárquica de la máscara de interfaz rural-urbano.

Capítulo 3

Clasificación de zona hortícola basada en GEOBIA (Geographic Object-Based Image Analysis)

3.1. Introducción

3.1.1. Problema, antecedentes y objetivo

La clasificación de imágenes satelitales por píxeles presentada en el capítulo anterior, significa un avance respecto a las metodologías usadas en Argentina para identificar cobertura y uso hortícola [94; 22; 23; 33]. Se incrementó la resolución espacial, se usaron métodos de clasificación no paramétricos basados en IA ¹, se aumentó la verdad de campo, se empleó una máscara de cobertura urbana y se explotaron las características texturales-contextuales junto con a las espectrales. Si bien todo ello resultó en una mejora de los productos de clasificación hortícola, todavía éstas presentan errores que vale la pena mejorar.

El *análisis de imágenes basado en objeto geográfico* (GEOBIA) [119; 120; 121] resulta una técnica que puede ayudar a superar los problemas anteriormente encontrados, a partir de identificar objetos representativos de la clase hortícola. GEOBIA es un modelo no supervisado que detecta objetos mediante la agrupación de píxeles *similares* en un mismo segmento (i.e., objeto). GEOBIA utiliza segmentos en lugar de píxeles como unidad de procesamiento. En el paisaje agrícola, los campos de cultivo aparecen como segmentos contiguos, lo que hace que GEOBIA sea una estrategia adecuada para interpretar imágenes satelitales de áreas rurales [122]. La medida para diferenciar objetos es una *distancia* (inversa a la similitud) en el dominio de las características de los segmentos. Estas características usualmente son espectrales, morfológicas y de contexto o vecindad. En esta propuesta se agregan características relacionadas a la textura, a los fines de abordar el *objeto lote hortícola*.

¹ Aprendizaje automático, inteligencia artificial o machine learning: SVM y RF.

El uso actual de las técnicas de detección de objetos o *segmentación* puede dividirse en dos enfoques, uno que asigna la clase mayoritaria al segmento, y otro que caracteriza al objeto y lo clasifica. El primer enfoque realiza una segmentación y una clasificación de píxeles, y propone la asignación de la clase mayoritaria al segmento que los contiene. Esta aproximación permite corregir píxeles erróneamente clasificados y establecer la unidad mínima de mapeo deseada [113; 22]². El segundo enfoque es un abordaje estadístico diferente, parte de una segmentación y describe a los segmentos con distintas métricas de forma y estadísticos sobre su contenido espectral. Luego se etiquetan los segmentos con la verdad de campo (VC) y se entrena el método de clasificación con los segmentos etiquetados para predecir sobre los no etiquetados. Este segundo enfoque mejora la relación muestra/población en comparación con una clasificación por píxeles (número de píxeles de VC versus objetos como VC). A partir de este método, se espera que la caracterización de los segmentos, permita identificar de manera automática aquello que el ojo experto discrimina al analizar visualmente e interpretar la imagen. En esta tesis se optó por el segundo enfoque.

La implementación de una cadena de procesamiento GEOBIA constituye un desafío metodológico así como una herramienta prometedora para la caracterización de áreas hortícolas. El objetivo del presente capítulo es explorar esta herramienta y disponibilizar un producto operacional que responda con mayor precisión a *¿Dónde?* se producen las hortalizas.

3.1.2. Requerimientos de usuario

Como productos se esperan, la detección de lotes agrícolas y la clasificación según su uso. Se precisa una clasificación en una zona periurbana, con un grado de precisión de clase Horticultura superior al 70%. Que el modelo clasifique mediante el ingreso de información de cobertura de suelo, menor al 1% de la zona de estudio total. Se requiere un procesamiento que permita un seguimiento temporal, sin costos adicionales de acceso.

² Unidad Mínima de Mapeo: es el objeto más pequeño distinguible en el mapa.

3.1.3. Requerimientos de software

La definición de los requerimientos de software, implica la descripción del modelo según sus funcionalidades, sin entrar en detalles de cómo lo tiene que hacer. Debe omitir terminología y detalles de implementación [123].

El objetivo de la cadena de procesamientos propuesta es, a partir de una imagen satelital, identificar objetos, caracterizarlos, y finalmente clasificarlos mediante métodos supervisados no paramétricos. Específicamente se requiere una secuencia de procesos en el que ingresen imágenes *raster* (matriz de datos georreferenciada) o *raster stack* para segmentar y aquellas necesarias para caracterizar los segmentos. La sucesión automatizada genera índices espectrales y filtros texturales, útiles en la caracterización de los objetos.

La cadena de procesamiento se implementa y basa en programas (i.e., *software*) de código abierto existentes, aprovechando las posibilidades ya implementadas de optimización. La clasificación es programada para hacerse con métodos supervisados y de validación cruzada. Las salidas tanto de la segmentación, como de la caracterización y clasificación pueden leerse por *software* de libre acceso. Se realiza el cálculo de métricas de desempeño de los procesos, y de tiempos de ejecución.

3.1.3.1. Entradas

El inicio de la cadena de procesamiento es la imagen de la zona de estudio. Los tipos de archivos que la cadena de procesamientos ingesta para comenzar con los procesos son de tipo *raster*, de formato JPEG o GeoTIFF (ver glosario).

La arquitectura es capaz de importar archivos vectoriales, tipo punto, como verdad de campo. Dicho archivo debe contener el par de coordenadas (latitud y longitud) del punto y la clase. Puede ser de tipo de formato ESRI shapefile o GeoPackage (ver glosario).

3.1.3.2. Salidas

Las salidas son archivos tipo *raster*: uno de *segmentos* donde cada pixel está etiquetado con el número de segmento al que pertenece, y otro de *clasificación* donde cada pixel está etiquetado con la clase (cobertura y uso de suelo).

En cuanto a la segmentación, también se requiere una salida como vector, con las columnas de atributos de: ID, características (estadísticas espectrales, texturales y morfológicas) y la columna clase resultante. Para dicha salida se recomienda usar el tipo de archivo *GeoPackage*.

De esta secuencia se obtienen matrices de confusión derivadas de los modelos de clasificación implementados (SVM y RF) y métricas de exactitud globales, así como tablas con los detalles de parámetros de dichos modelos.

3.1.4. Segmentación

La segmentación de imágenes implica subdividir una imagen en regiones homogéneas [124] que idealmente representan objetos de interés del mundo real. Los segmentos son regiones que se generan por uno o más criterios de homogeneidad en una o más dimensiones (del espacio de una imagen). El espacio n-dimensional sobre el que opera la segmentación, está conformado por propiedades específicas tales como la información espectral (color, o *gray level*), texturas, forma, tamaño y topología. La segmentación debe buscar (a) la uniformidad de objetos individuales (homogeneidad intra-segmentos) y (b) una diferencia significativa entre objetos adyacentes (heterogeneidad inter-segmentos) y estas pueden ser medidas, para evaluar la segmentación [125].

El reconocimiento de objetos es el proceso de agrupar píxeles similares en segmentos únicos mediante algoritmos basados en detección de bordes y crecimiento de regiones, descritos en la literatura [126; 124; 127]; trabajos más recientes incorporan los métodos híbridos y semánticos (63). Cada objeto encontrado durante el proceso de segmentación recibe un identificador único y es una colección de píxeles contiguos que cumplen con criterios de homogeneidad en las características.

Los segmentos tienen información espectral adicional comparada con los píxeles, ya que en el procesamiento, se calculan estadísticos sobre las variables espectrales y texturales (valores mínimos, máximos, media, desvíos, varianza, etc.). Asimismo la dimensión espacial de los segmentos en la imagen incluye atributos geométricos y la información contextual, en la que se evalúan distancias, vecindad y topologías. Este dominio de características es la pieza fundamental para los métodos GEOBIA, al tomar el segmento como

la unidad espacial básica, en contraposición al píxel. El propósito de la segmentación ha cambiado desde simplemente ayudar al etiquetado del píxel hacia una identificación del objeto [128]. Sin embargo, vale decir que los resultados de la segmentación de imágenes pueden ser útiles por sí solos. La segmentación puede ser enfocada como un pre-procesamiento, sin intervención en la clasificación. Solo utilizando los objetos de la segmentación para etiquetarlos con la moda de la clasificación de los píxeles que contienen. Esta técnica contribuye a superar el llamado “efecto sal y pimienta” [129], entendido como píxeles mal clasificados o “ruido” o para establecer una unidad mínima de mapeo [113].

La segmentación se ha mostrado particularmente beneficiosa cuando los objetos de interés en una imagen son más grandes que los píxeles de la misma, como generalmente sucede con las imágenes de alta resolución espacial (sub-métricas o llamadas *very high resolution VHR*) [120], pero también útil en el procesamiento de aquellas de resolución media (Sentinel-2 [130], Spot). Los diferentes tipos de imágenes requieren diferentes escalas de análisis y, además, en muchos casos los objetos significativos aparecen a diferentes escalas de análisis de la misma imagen [120]. La calidad de la segmentación impacta fuertemente en la precisión de la clasificación que se pretende [128], al agrupar los píxeles individuales en objetos significativos, objetos geográficos de interés en la escena.

El enfoque utilizado para construir objetos se remonta a los años 70’ [126; 124; 129; 120] desde el lanzamiento del Landsat-1 [128]. Sin embargo durante los años 80’ y 90’ fueron los abordajes de estudio de píxel o sub-píxeles los que prevalecieron en el análisis de la información satelital [120]. GEOBIA tuvo un impulso con el incremento de la resolución espacial a niveles sub-metro (IKONOS 1999, QuickBird 2001, OrbView 2003, o información local de LIDAR o VANTs) y los nuevos campos de aplicación en ciencia ambiental, como los programas supranacionales de monitoreo, evaluación de riesgo y daños; aún con imágenes de resolución media: Spot, Landsat8, Sentinel-2 que se lanzaron unos años más tarde. El abordaje GEOBIA se vio impulsado tanto por la disponibilidad de sensores remotos de mayor resolución espacial, como de nuevas áreas temáticas que demandan este abordaje. Este fomento, se yuxtapone al lanzamiento del software comercial *eCognition*© en el año 2000, siendo hasta la actualidad el más usado en las publicaciones científicas

de GEOBIA, representando 50-55% de las mismas [120; 128]. Se suman luego una serie de otros programas bajo licencia (Feature Analyst, ENVI Feature Extraction, ERDAS Imagine, IDRISI, ArcGIS, entre otros). Las aplicaciones GEOBIA se han desarrollado aún más rápidamente desde 2004/2005, y están siendo acompañadas con el desarrollo de códigos abiertos e implementaciones en software libres (SAGA³, OTB⁴, GRASS GIS⁵, SciKit Image [131; 132], RSGISLib⁶). La última fase de la investigación de GEOBIA (desde 2005) se dirige más hacia la automatización del procesamiento de imágenes [120; 133; 125; 130; 128].

3.1.4.1. Algoritmos de segmentación

Los algoritmos de segmentación están descriptos en una vasta literatura y los lectores pueden referirse a [126; 124; 127; 120; 128]. En este apartado se sigue la categorización sugerida por [127; 128], y se mencionan algunos de los métodos más usados en GEOBIA para el reconocimiento de objetos agrupando píxeles similares en segmentos únicos. La segmentación de imágenes en GEOBIA, desde una perspectiva algorítmica, se divide generalmente en cuatro categorías: (a) Basados en Bordes, (b) Basados en Regiones (Creciente o de División), (c) Híbridos y (d) Semánticos. Los métodos Híbridos integran los resultados de los métodos basados en bordes y basados en regiones, obtienen mejores resultados pero los algoritmos son difíciles de implementar, intensivos computacionalmente y no están disponibles en programas para ejecutarlos [128].

Los métodos Semánticos son abordajes supervisados, basados en algoritmos de *machine learning* (ML) (MRF, Redes Bayesianas, Redes Neuronales, SVM, Redes Neuronales Convolucionales, etc.). Requieren una gran cantidad de datos para entrenarlos y un número significativo de parámetros para ajustar [128].

3.1.4.1.1. Basados en Bordes Los algoritmos basados en bordes, primero identifican bordes y luego los unen usando algoritmos de contorno. Asumen que entre los bordes, las propiedades de los píxeles cambian abruptamente. Esos

3 http://www.saga-gis.org/saga_tool_doc/2.2.2/imagery_segmentation.html

4 <https://www.orfeo-toolbox.org/>

5 <https://grass.osgeo.org/grass74/manuals/i.segment.html>

6 <https://www.rsgislib.org/>

bordes se consideran como límites entre objetos. Según [128] la detección de borde requiere tres pasos: filtrado (reduce el desenfoque y desplazamiento de los bordes), realce (acentúa los píxeles donde la intensidad cambia, así como lo hace un filtro de bordes) y detección (une los bordes, cierra los límites, elimina los bordes ruidosos).

Estos algoritmos son menos complejos que los basados en región [128], funcionan mejor en imágenes con alto contraste entre los objetos y el fondo y son computacionalmente eficientes. Por el contrario, tienen problemas con imágenes “ruidosas”, de bajo contraste o transiciones difusas entre objetos. No son muy aplicados en imágenes de alta resolución, y presentan dificultades en las imágenes multiespectrales debido a la inconsistente locación de los bordes para las distintas bandas [128]. Algunos ejemplos de algoritmos de segmentación por detección de bordes son *Watershed* [134], que puede verse implementado en [131], y *Canny* [135; 136], usado en el capítulo 4 en el módulo de generación de mosaicos o “zonas morfológicas”.

3.1.4.1.2. Basado en Regiones Los algoritmos basados en regiones, empiezan desde el interior de un objeto y se expanden hacia afuera buscando la “suficiente similitud” hasta encontrar los límites del objeto, es decir no-similitud (según un umbral) [126; 128]. Los métodos basados en región asumen que los píxeles vecinos en una misma región tienen valores similares. Estos algoritmos realizan dos operaciones básicas: *fusión* y *división*. Para ello obtienen una primera sobre-segmentación (o infra-segmentación) de la imagen y luego fusionan los segmentos adyacentes que son similares (o separan los segmentos que son disímiles), y siguen una y otra vez hasta que no quedan segmentos que deben ser unidos (o separados) [128]. Estos métodos son útiles para generar segmentos a múltiples escalas [120]. Los usuarios pueden elegir los puntos semillas y los criterios de fusión (homogeneidad). Son menos sensibles al “ruido” que los basados en bordes.

Hay dos tipos más comunes de algoritmos de segmentación basados en regiones: división de región y fusión (a), y región creciente y fusión (b). Aquellos de **división de región y fusión**, comienzan con una semilla que es toda la imagen y dividen la imagen en segmentos según un criterio de no-homogeneidad de los mismos [120]. Combinan un abordaje de abajo hacia arriba, con uno de arriba hacia abajo. El enfoque de abajo hacia arriba genera objetos combinando y fusionando píxeles, mientras el otro enfoque de arriba hacia

abajo, separa la imagen entera en segmentos según una norma de no-homogeneidad. La desventaja de estos algoritmos es que generan segmentaciones cuadradas, y la ventaja es que son menos costosos computacionalmente dado que parten de una sola semilla. Los de **región creciente y fusión**, en cambio, realizan dos acciones: selección de semillas de cada región y cálculo de similitud. Luego de seleccionar las semillas, la región crece sumando píxeles adyacentes que son similares, de acuerdo a criterios específicos (definidos por el usuario), incrementando el tamaño de la región gradualmente [124; 128]. Son los más usados, se puede ver en la Tabla 3.1, el acceso a herramientas y programas vigentes [128].

De los algoritmos de segmentación “basados en regiones”, se describen los de “crecimiento y fusión de regiones” aplicables en GEOBIA, *Mean Shift* [137] y *Region Growing* [138].

Tabla 3.1: Programas y herramientas disponibles para segmentación. Modificado de Hossain y Chen 2019 (63)

Programa/ Herramienta	Referencia/ Desarrollador	Sitio web	Algoritmo basado en	Licencia / Disponibilidad
Orfeo ToolBox	Grinzonnet <i>et al.</i> (2017)	https://www.orfeo-toolbox.org/otb/	Región	Software Libre
RSGISLib	Buntin <i>et al.</i> (2014)	https://www.rsgislib.org/	Región	Software Libre
SCRM	Castilla <i>et al.</i> (2009)	http://www.castlink.ca/scrm/scrm	Región y Borde	Software Libre
GRASS GIS	Neteler <i>et al.</i> (2008)	https://grass.osgeo.org/grass76/manuals/i.segment.html	Región y Borde	Software Libre
SAGA	Böhner <i>et al.</i> (2006)	http://www.saga-gis.org/en/index.html	Región y Borde	Software Libre
SPRING	Cámara <i>et al.</i> (1996)	http://www.dpi.inpe.br/spring/english/index.html	Región y Borde	Software Libre
EDISON	Comaniciu y Meer (2002)	http://rci.rutgers.edu/~meer/RIUL/research/code/EDISON/doc/overview.html	Región	Software Libre
Scikit-image (Python)	van der Walt <i>et al.</i> (2014)	https://scikit-image.org/docs/dev/auto_examples/segmentation/plot_egmentations.html	Bordes	Código Abierto
GeoSegment	Chen (2018)	http://130.15.95.215/lagisa/	Región	Herramienta en línea
InterSeg	Happ <i>et al.</i> (2016)	http://www.lvc.ele.puc-rio.br/wp/?cat=41	Región (en la nube)	Disponible a pedido
RHSeg	Tilton <i>et al.</i> (2012)	https://opensource.gsfc.nasa.gov/projects/HSEG/	Región	Copia de Evaluación
ArcGIS Spatial Analyst	ESRI	http://desktop.arcgis.com/es/arcmap/10.3/tools/spatial-analyst-toolbox/an-overview-of-thesegmentation-and-classification-tools.html	Región	Comercial
Object Analyst	PCI Geomatics	http://www.pcigeomatics.com/geomatica-help/concepts/focus_oa_ntro.html	Región	Comercial
IDRISI GIS Tool	Clark Labs	https://clarklabs.org/terreset/idrisi-gis/	Borde	Comercial
ENVI Feature Extraction	Harris Geospatial Solutions	https://www.harrisgeospatial.com/docs/routines164.html	Borde	Comercial
IMAGINE Spatial Modeller	Hexagon Geospatial	https://community.hexagongeospatial.com/t5/SpatialModeler/tkb-p/eTSpatialModeler	Borde	Comercial
BerkeyImSeg	Clinton <i>et al.</i> (2010)	https://www.imageseg.com/	Región	Comercial
Feature Analyst	Opitz y Bundell (2008)	https://www.textronsystems.com/products/feature-analyst	Semantic	Comercial
SEGEN	Goffman (2006)	http://www.research.ibm.com/haifa/projects/image/seggen/index.html	Región	Comercial
eCognition	Baatz and Schäpe (2000)	http://www.ecognition.com/suite/ecognition-developer	Región y Borde	Comercial

Algoritmo Mean Shift (MS) Es un algoritmo de agrupamiento con una estimación de densidad no-paramétrica [137]. Segmenta la imagen agrupando todos los píxeles que son cercanos, en un dominio espectral y espacial, para luego conectar los puntos de correspondencia [128]. Es un algoritmo sencillo, que arroja buenos resultados pero computacionalmente costoso, por ello no es tan utilizado.

El algoritmo Mean Shift trabaja en dos etapas: primero le calcula a cada píxel el punto de densidad que lo “atrae” (usando los píxeles dentro de un radio espacial-espectral r), es decir busca máximos locales de densidad (Figura 3.1). Estos máximos locales están ligados a un área de interés llamado *radio r* , definido por el usuario. Comienza por un píxel atractor, establece el centro, y luego lo cambia a la media que los píxeles dentro del radio r determinen otro centro de atracción. Itera hasta que el cambio entre media y media sea menor que un umbral (establecido por el usuario) o una cantidad máxima de iteraciones. Devuelve *grupos o clusters*: con los puntos “atractores” (media de grupo) y *pertenencia*: un vector enunciando a cual grupo fue “atraído” cada píxel. Todo píxel que nos dirija hacia el mismo máximo local de convergencia forma parte de la misma región. Hace el idéntico proceso en paralelo barriendo la imagen y encontrando cuantos máximos locales existen. A partir de esto, etiqueta a los píxeles de la imagen según la región (segmento) al que pertenezcan [139].

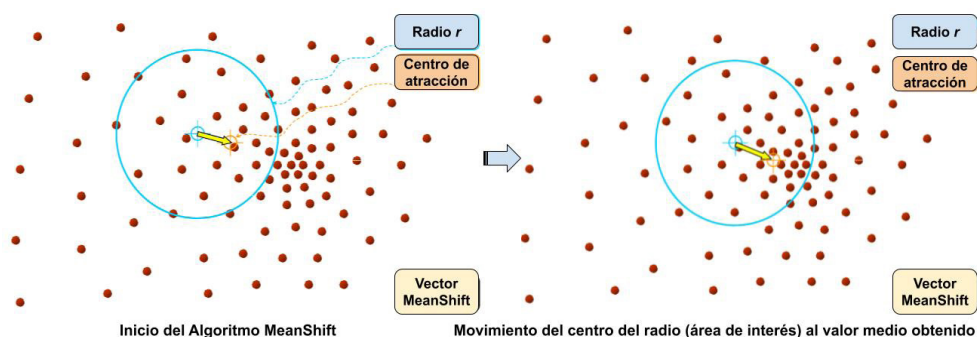


Figura 3.1: Esquema de algoritmo Mean Shift, un ciclo o iteración. Modificado de [139]

Algoritmo Region Growing (RG) Comienza con la generación de píxeles semilla, un proceso costoso computacionalmente, dado que en general, son todos los píxeles de la imagen. Algunas implementaciones permiten la adición de semillas definidas por el usuario.

El algoritmo RG de segmentación examina secuencialmente todos los segmentos actuales en el mapa ráster. La similitud entre el segmento actual y cada uno de sus vecinos se calcula de acuerdo con la fórmula de distancia dada. Los segmentos se fusionarán si cumplen una serie de criterios [140], que incluyen:

- El par es mutuamente más similar entre sí (la distancia de similitud será menor que la de cualquier otro vecino), y
- La similitud debe ser inferior al umbral de entrada. El proceso se repite hasta que no se realicen combinaciones durante un pase completo.

3.1.4.2. Parámetros de los algoritmos de segmentación

Los algoritmos de segmentación, precisan parámetros suministrados por el usuario. La calidad de la segmentación resultante depende de la elección de los mismos. Los parámetros descritos aquí son los pertinentes a los algoritmos Mean Shift (MS) y Region Growing (RG).

Los parámetros generales son: tamaño mínimo de objeto y umbral de similitud. El *tamaño mínimo* de objeto, es el número de píxeles mínimo de los segmentos, los algoritmos de región creciente iteran hasta alcanzarlo y siguen si otro de los parámetros aún lo determina. El *umbral* es un valor de similitud entre segmentos que se utiliza para determinar qué objetos (i.e., segmentos) se fusionan. El valor de *umbral* da cuenta de una distancia (Euclídea o Manhattan) en el espacio n-dimensional de características de los segmentos (ej. valores radiométricos). Los valores de distancia más pequeños indican una coincidencia más cercana, con una puntuación de similitud de cero para píxeles idénticos. Durante el procesamiento normal, las combinaciones sólo se permiten cuando la similitud entre dos segmentos es inferior al valor de umbral dado. Sin embargo, durante la última iteración, los objetos con un número de píxeles menor al tamaño mínimo, se fusionaron con su vecino más similar, incluso si son más disímiles que lo permitido por el umbral [141; 140]. El *umbral* debe ser mayor que 0 y menor que 1. Un umbral de 0 permitiría que solo se fusionaran píxeles de valor idéntico, mientras que un umbral de 1 permitiría que todo se fusionara [140].

En el caso del algoritmo *MeanShift*, además de umbral y tamaño mínimo de objeto, requiere los parámetros de *radio r* es decir la zona de interés, tanto *spectral* como *espacial o de distancia*.

3.1.4.3. Evaluación de segmentación

Los parámetros de segmentación se seleccionan a menudo después de un tedioso y largo proceso de prueba y error [120; 128]; la optimización de los mismos es de trascendental importancia para obtener buenos resultados de segmentación. Algunas medidas de precisión sobre las segmentaciones han sido descritas en [142] para evaluar la elección de parámetros. Se puede categorizar las propuestas de evaluación en dos: supervisadas (a) y no supervisadas (b). Para la evaluación **supervisada**, es necesario tener una segmentación previa como verdad de campo, con la que un experto de dominio compara visualmente, y con la ayuda de una serie de medidas (número de segmentos, tamaño de segmentos, discrepancia espectral del segmento, superposición de áreas, bordes, etc.) determina si la combinación de parámetros usados en el algoritmo es buena. Con la evaluación **no supervisada**, no es necesaria una previa segmentación como verdad de campo, y se trata de *distintas funciones de optimización* que combinan, de modos diferentes: la *varianza ponderada* (homogeneidad dentro del segmento) y la *autocorrelación espacial* (separabilidad entre segmentos). Los métodos no supervisados de evaluación, son analíticos y permiten mayor nivel de objetividad y automatización [143].

De los métodos de evaluación no supervisados se presentan brevemente la *F* de Jonhson [143] y la *F* o *Sum* de Espíndola[141]. Ambas funciones de optimización operan sobre la *varianza ponderada* como medida de homogeneidad interna del objeto (segmento) y la *autocorrelación espacial* como una medida de heterogeneidad entre los objetos (segmentos). La *varianza* es ponderada por el área de cada segmento (mayor para los segmentos más grandes, ver ecuación 3.1).

Homogeneidad dentro del segmento, Varianza Ponderada (VP):

Es calculada para cada banda espectral o característica textural sobre la que se segmenta, luego se promedia [143].

$$vp = \frac{\sum_{i=1}^n a_i v_i}{\sum_{i=1}^n a_i} \quad (3.1)$$

donde v_i es la *varianza* y a_i es el *área* del segmento i .

Heterogeneidad entre segmentos, índices de Autocorrelación Espacial (AE):

La autocorrelación espacial, mide el grado de asociación espacial, es decir rompe con el presupuesto de independencia de las observaciones [144]. Es una propiedad conocida de los datos espaciales, por la que los valores similares de una variable en ubicaciones cercanas, darán lugar a conglomerados espaciales [141]. La autocorrelación espacial es multidireccional y multidimensional, lo que la hace útil para encontrar patrones en conjuntos de datos complicados [141]. Puede ser medida mediante el índice I de Moran (1950) o C de Geary (1954).

El estadístico C mide la autocorrelación de una variable espacial. C no asume valores negativos. La media y la varianza del índice se obtienen bajo el supuesto de normalidad de los datos, o bajo el supuesto de aleatorización [144]. En ambos casos se prueba que $E(C) = 1$. Para su interpretación: $C < 1$ indica que en sitios conectados los valores de la variable son similares, $C > 1$ evidencia que la variable en los sitios conectados asume valores disímiles o diferentes. Para la segmentación se busca que los valores de C sean mayores a 1.

El índice I de Moran mide la autocorrelación lineal espacial entre valores de una variable (Moran, 1950 citado por (24)). Una variable espacial tiende a asumir valores similares en unidades geográficamente cercanas (Anselin, 2001 citado por [144]). La I de Moran aplicada a imágenes segmentadas captura como, en promedio, los valores medios de cada segmento difieren de los valores medios de los segmentos vecinos. A valores bajos de I , los segmentos vecinos son estadísticamente diferentes. Entonces un mínimo local de este índice señala mucha heterogeneidad entre los segmentos y esto se asocia a límites claros entre objetos. El índice de Moran oscila entre -1 y 1; los valores cercanos a 0 indican baja autocorrelación espacial, y hacia -1 y 1 indican autocorrelación espacial.

3.1.4.3.1. Evaluación por Suma de Espíndola La función objetivo de Espíndola *et al.*, pretende dar cuenta de una combinación óptima de parámetros, evaluando la segmentación [141]. Es una suma (i. e. *sum*), donde se busca la combinación de valores bajos de I (segmentos disímiles) y bajos de vp (segmentos internamente homogéneos), mediante la siguiente ecuación:

$$F(vp, I) = F(vp) + F(I) \quad (3.2)$$

donde las variables fueron previamente normalizadas:

$$F(x) = \frac{X_{max} - X}{X_{max} - X_{min}} \quad (3.3)$$

3.1.4.3.2. Evaluación por F de Johnson Es una medida de precisión global de la segmentación [143], que permite una evaluación no supervisada para optimizar los parámetros de la segmentación. Se calcula mediante la siguiente ecuación:

$$F = (1 + \alpha^2) \frac{F(I) \times F(vp)}{\alpha^2 \times F(I) + F(vp)} \quad (3.4)$$

donde α es un coeficiente de ponderación que controla los pesos relativos de $F(I)$ y $F(vp)$, y donde las variables fueron previamente normalizadas a valores de 0 a 1, con la ecuación 3.3. Si $F \approx 0$ la segmentación es de baja calidad y si $F \approx 1$ es de alta calidad.

Si $\alpha = 1$ indica dar igual peso a AE que a VP , si $\alpha = 0,5$ da la mitad del peso a VP (no tan importante que los segmentos sean homogéneos) y si $\alpha = 2$ duplica el peso de VP y la importancia de la homogeneidad intra-segmento. Entonces $\alpha > 1$ son más apropiados para obtener objetos más pequeños, propios de las segmentaciones más finas (los objetos son similares espectralmente a su alrededor). Pero si se busca objetos más grandes y heterogéneos, en una segmentación gruesa el α debe ser < 1 . Esto le permite adaptarse a los métodos multi-escala de GEOBIA (MS-GEOBIA) [143; 120].

La *Suma* de Espindola y F de Johnson son operadores de medición de optimización que pueden trabajar la autocorrelación espacial (AE) tanto con I de Moran como con la C de Geary [125]; aunque los trabajos originales hacen mención únicamente a la I de Moran [141; 143].

Estos conceptos serán empleados en la cadena de procesamiento GEOBIA. El **objetivo** del presente capítulo es explorar esta herramienta y disponibilizar un producto operacional que responda con mayor precisión a ¿*Dónde?* se producen las hortalizas. Se requiere una sucesión automatizada que a partir de

una imagen satelital, genere índices espectrales y filtros texturales. Luego que segmente lotes agrícolas en una zona periurbana, los caracterice y finalmente los clasifique.

3.2. Materiales y Métodos

3.2.1. Cadena de procesamiento

El sistema modular propuesto para analizar imágenes basadas en objeto geográfico (GEOBIA) consta de tres subsistemas o módulos, además de las entradas y salidas que cumplen con los requisitos de software de la sección anterior. La entrada inyecta los datos, externos al procesamiento, en formato raster y/o vector a cada módulo (Figura 3.2).

El primer módulo es de **Generación de Mosaicos o tiling** que divide la imagen a procesar en mosaicos. Tiene como función generar zonas donde se encontrarán distintas morfologías de objetos, por eso también se llama segmentación en zonas morfológicas. Se tiene en cuenta que, la división de la imagen en mosaicos rectilíneos presenta problemas para GEOBIA y puede dar lugar a defectos en el procesamiento sobre las líneas rectas en los límites de los mosaicos [133]. En esos límites se puede cortar objetos (semánticos) y perder información. Así mismo al superponer mosaicos rectilíneos se multiplica información y procesamiento. Por ello se trabaja con una macro-segmentación de la imagen basada en las morfologías de los objetos que se pretende diferenciar. La misma puede ser obtenida de modo no-supervisado (algoritmos de segmentación) o ingresando un vector o raster con dicha división morfológica de interés. De este primer bloque se obtienen tanto el vector como el raster de zonas morfológicas.

El segundo módulo es la **Segmentación** de lotes propiamente y la caracterización de los segmentos con estadísticos de la información espectral ingresada, la información textural generada y las descripción morfológica de cada segmento. La salida de este módulo es un raster de segmentos y un vector de segmentos, con las características extraídas.

El módulo final de la estructura es de **Clasificación**, en donde ingresa el vector de verdad de campo (con elementos tipo punto) con el que se etiquetan y

extraen los segmentos que contienen VC del vector generado en el módulo de segmentación. Los segmentos extraídos ingresan en un esquema de validación cruzada (de k-iteraciones) para entrenar modelos supervisados no paramétricos de clasificación. Las predicciones se realizan sobre todos los segmentos, y luego se validan, obteniendo métricas de exactitud general y por clase.

La salida de la estructura de procesamiento propuesta es un raster de clases, un vector de características y la predicción de clase; así como las métricas y detalle de la parametrización de cada modelo de clasificación implementado.

A continuación se describe la aplicación de la cadena de procesamiento al área de estudio. La misma se aplica a una porción de la imagen Sentinel-2 trabajada en el capítulo 2.

3.2.1.1. Módulo: Generación de mosaicos o tiling

De la misma imagen Sentinel-2A (ESA) de 19/02/2019 preprocesada a reflectancia superficial del capítulo 2, se seleccionó una matriz de 400×400 píxeles, en la que las clases horticultura, agricultura extensiva, construcción, arboleda, tierra arada y tierra de uso mixto, se encontraron presentes. Se calculó el NDVI y la matriz de varianza del mismo en una ventana de 3×3, sobre la misma se digitalizaron manualmente 2 zonas (Figura 3.3). La zona 1, con un uso de suelo mayormente hortícola, y la zona 2 con un uso de suelo predominantemente agricultura extensiva (maíz y soja). En este capítulo se realizó manualmente, dado que el foco estaba en optimizar parámetros de la segmentación a nivel de lote diferenciando morfología de aquellos con agricultura extensiva de los hortícolas.

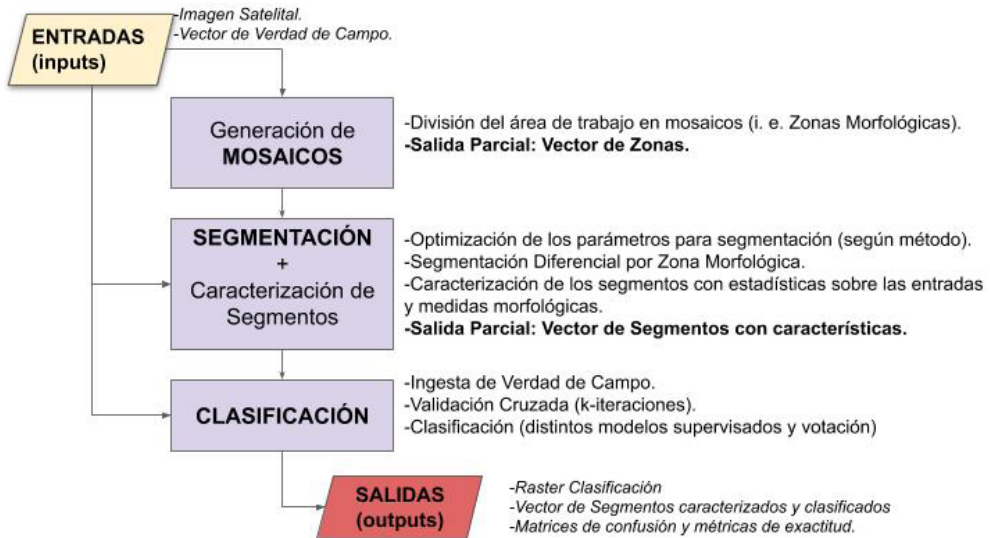


Figura 3.2: Arquitectura modular propuesta para GEOBIA.



Figura 3.3: Zonas Morfológicas: divide la optimización de parámetros en el proceso de segmentación.

3.2.1.2. Módulo: Segmentación

3.2.1.2.1. Fuente de datos para segmentación La discusión en torno a qué dato o conjunto de datos fueron útiles para segmentar, dependió de dos grandes cuestiones: disponibilidad y utilidad. Respecto a una amplia disponibilidad de datos (espectrales o texturales) se observó cuales son más útiles según los algoritmos usados (más o menos eficientes con imágenes VHR o multiespectrales, ruidos, datos con poco contraste, etc.) y junto al trabajo realizado de

optimización de parámetros para poder mejorar los resultados. En esta tesis, se trabajó con información multispectral y textural. Se realizaron pruebas con distintas combinaciones de datos disponibles, hasta encontrar aquella que resultó más útil a los fines de identificar lotes hortícolas.

La textura es una característica de las clases específicas de cobertura del suelo en imágenes satelitales. Es particularmente útil en situaciones en las que las diferencias espectrales entre clases son pequeñas, y las mismas se distinguen por su organización sobre el terreno. Esta organización se muestra diferente u opuesta entre espacios naturales y artificiales; o campos de cultivos extensivos frente a campos hortícolas con diversidad; ciudad frente a arboledas de ribera, entonces todas estas coberturas presentan distintos valores de textura. La utilidad y el uso de la textura depende en gran medida de la resolución de las imágenes satelitales y de la escala de la intervención humana o del fenómeno que creó la textura [145]. El usuario debe observar el fenómeno visualmente para determinar un ajuste adecuado del parámetro de tamaño. Dado que las texturas dependen en gran medida de la escala, pueden aparecer texturas jerárquicas. El módulo usado fue *r.texture* de GRASS, que utiliza un modelo de textura común basado en la *matriz de co-ocurrencia de nivel de gris* descrita por Haralick et al. (1973) (citado por [85]). Las medidas texturales disponibles en el módulo *r.texture* [145], y aplicadas a la escena, sobre el NDVI, fueron:

Texturas estadísticas de primer orden en el ámbito espacial:

- Suma Promedio (SA) (Figura 3.4 (a)).
- Entropía (ENT): Esta medida analiza la aleatoriedad. Es alta cuando los valores de la ventana en movimiento son disímiles. Es baja, cuando los valores están cerca de 0 o 1, es decir, cuando los píxeles de la ventana local son uniformes (Figura 3.4 (b)).
- Diferencia de Entropía (DE) (Figura 3.4 (c)).
- Suma de Entropía (SE) (Figura 3.4 (d)).
- Varianza (VAR): Una medida de la variación del tono gris dentro de la ventana en movimiento (momento de segundo orden alrededor de la media) (Figura 3.4 (e)).
- Diferencia de Varianza (VD) (Figura 3.4 (f)).
- Suma de Varianza (SV) (Figura 3.4 (g)).

Texturas estadísticas de segundo orden en el ámbito espacial:

- Segundo Momento Angular (ASM, también llamado Uniformidad): Esta es una medida de la homogeneidad local y lo opuesto de la Entropía. Los valores altos de ASM ocurren cuando los píxeles en la ventana en movimiento son muy similares (Figura 3.4 (h)).
- Momento de Diferencia Inversa (IDM, también llamado Homogeneidad): Esta medida se relaciona inversamente con la medida de contraste. Es una medida directa de la homogeneidad local de una imagen digital. Los valores bajos se asocian a una baja homogeneidad y viceversa (Figura 3.4 (i)).
- Contraste (CON): Esta medida analiza el contraste de la imagen (variaciones locales del nivel de gris) como la dependencia lineal de los niveles de gris de los píxeles vecinos (similitud). Típicamente alta, cuando la escala de la textura local es mayor que la distancia (Figura 3.4 (j)).
- Correlación (COR): Esta medida analiza la dependencia lineal de los niveles de gris de los píxeles vecinos. Típicamente alta, cuando la escala de la textura local es mayor que la distancia (Figura 3.4 (k)).
- Medidas de información de correlación (MOC) (Figura 3.4 (l)).
- Coeficiente de Correlación Máxima (CCM) (Figura 3.4 (m)).

Con el objetivo de lograr segmentar la imagen a nivel de lote, se probaron múltiples combinaciones de los datos espectrales de las bandas de Sentinel-2 y los índices de vegetación *NDVI*, *NDWI*, las medidas texturales calculadas sobre el *NDVI* y generadas por el *r.texture* y una medida llamada “varianza media-umbral”. Se muestran todas las características generadas en las Figuras 3.4 y 3.5. La *Varianza Media-Umbral* resulta de aplicar un filtro de media (con una ventana móvil de 21×21) a la varianza de *NDVI*, en la que se selecciona un umbral ($2000 > Var - Umbral > 200$) que resalta bordes internos en los lotes (“característica hortícola”), el resultado se puede observar en la Figura 3.5 (l). El umbral aplicado a la varianza surgió de pruebas preliminares que realzan el patrón de uso de suelo hortícola.

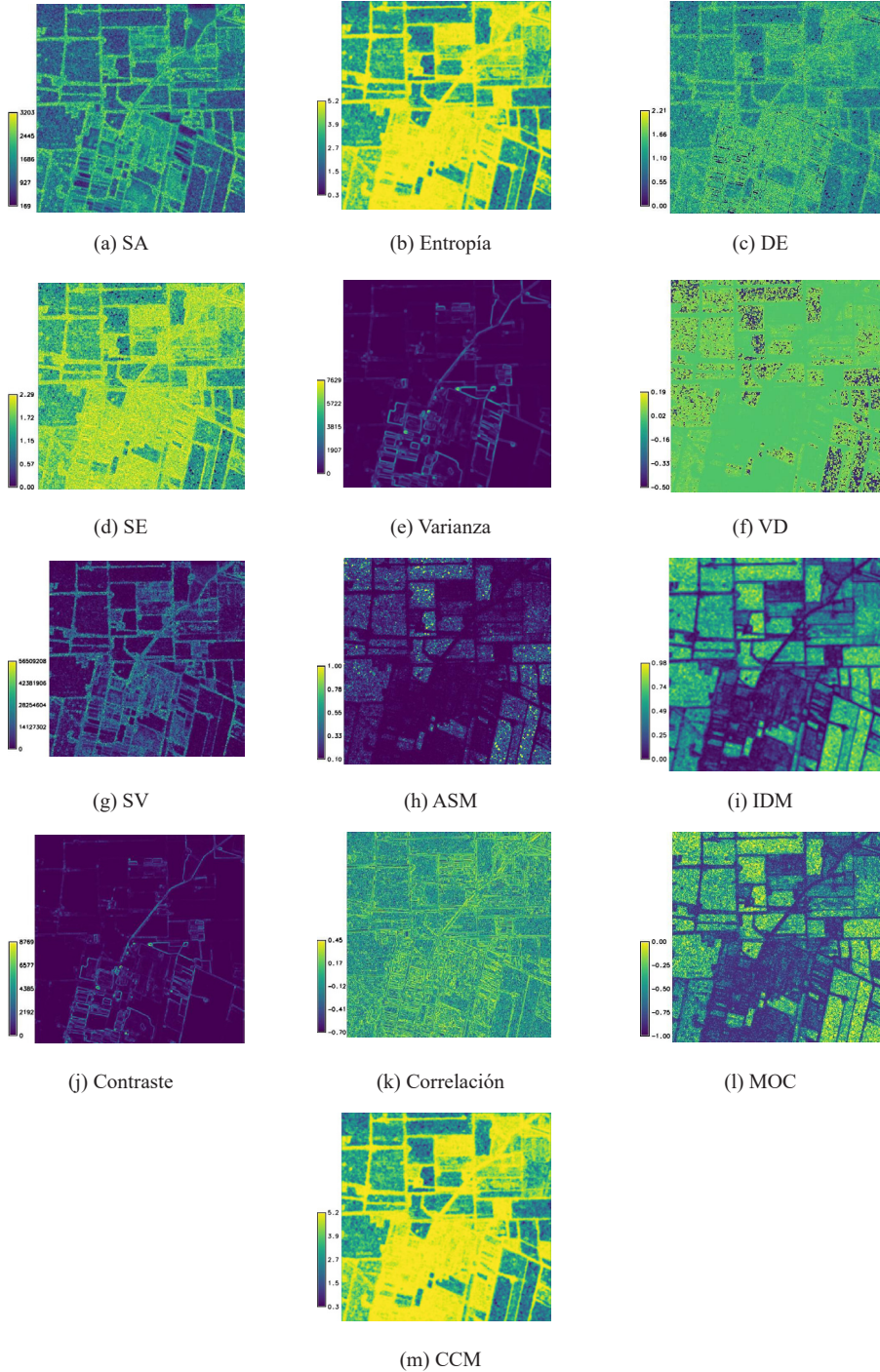


Figura 3.4: Medidas texturales disponibles en el módulo r.texture de GRASS sobre la zona de estudio, ventana 5×5 sobre NDVI

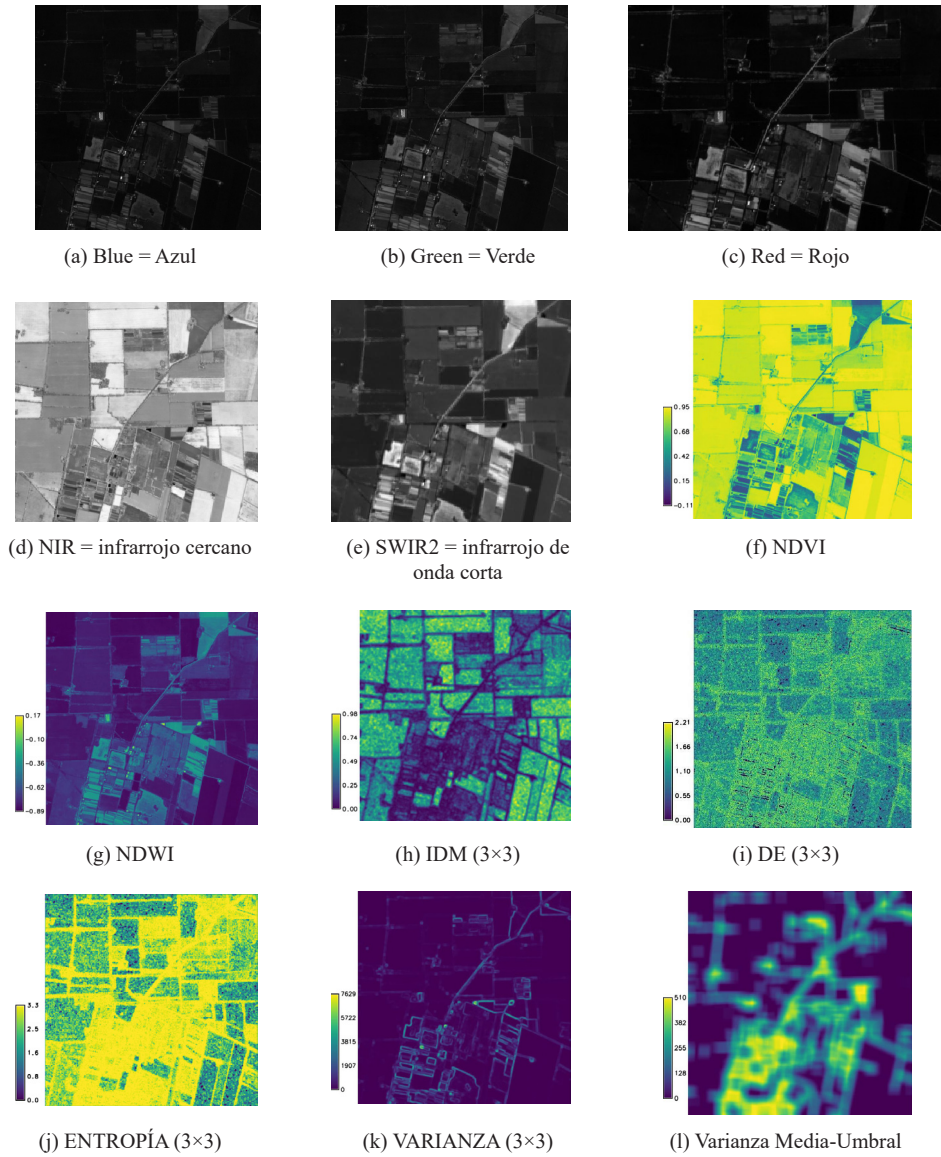


Figura 3.5: Capas de información espectral y textural generadas para segmentar y/o clasificar.

Las segmentaciones resultantes se evaluaron tanto visualmente, como analíticamente (valores de funciones de optimización F [143] o sum [141]) y fueron seleccionados para mostrar los resultados de 5 combinaciones de datos. En la Tabla 3.2 se enumeran las bandas espectrales, índices, y medidas texturales de cada conjunto de entrada de datos para segmentar.

Tabla 3.2: Combinación de datos de bandas espectrales, índices, y medidas texturales de cada conjunto de entrada de datos usados en los mejores resultados de segmentación.

Entradas/Datos	Sentinel-2A	índices de vegetación	medidas texturales (de NDVI)
1	10 bandas*		
2	10 bandas	NDVI y NDWI	13 medidas texturales+
3	B, G, R y NIR**		
4	G, R, NIR y SWIR2		
5	B, G, R y NIR	NDVI y NDWI	IDM (Figura 3.5(h)) VARIANZA (Figura 3.5(k)) Varianza Media-Umbral (Figura 3.5(l))

* Se utilizaron las bandas azul (B ~ 493 nm), verde (G ~ 560 nm), rojo (R ~ 665 nm), infrarrojo cercano (NIR ~ 833 nm) con 10 m de resolución espacial, y 4 bandas estrechas en el dominio espectral de borde rojo de vegetación VNIR (B5 ~ 704 nm, B6 ~ 740 nm, B7 ~ 783 nm, B8a ~ 865 nm) y 2 bandas infrarrojas de onda corta, más anchas (SWIR1 ~ 1610 nm y SWIR2 ~ 2190 nm) remuestreadas a un píxel de 10 m.

** Sólo las bandas de mayor resolución espacial R, G, B, NIR de la imagen (10 m/px)

+ Las 13 medidas texturales calculadas sobre el NDVI con una ventana móvil de 5x5 (Figura 3.4)

3.2.1.2.2. Algoritmos, parámetros y optimización Con el fin de identificar lotes mayores a 1 ha sobre la imagen Sentinel-2 (10m/px), se trabajó con algoritmos de crecimiento y fusión de regiones aplicables a GEOBIA, *Mean Shift* y *Region Growing*. Se usó la implementación de los mismos en el módulo *i.segment*[140] de GRASS [146].

Se establecieron los parámetros mínimos y máximos de cada algoritmo (tamaño mínimo de objeto, umbral de similitud, *r*) para fijar los rangos de valores de prueba, y luego optimizar. El mínimo se fijó tras un primer resultado que arrojó una sobre-segmentación (comparada a la deseada por el usuario), el máximo en cambio, es aquel que genera una sub-segmentación (objetos de distinta clase en un mismo segmento). Dichas pruebas empíricas se siguieron bajo las sugerencias de Momsen [147] que recomienda iniciar con un valor bajo (ej. 0,01) para el *umbral*. El tamaño mínimo se estableció del mismo modo, añadiendo al criterio del tamaño de objeto que se pretende y la capacidad de cómputo disponible que fue medida por el tiempo que tomó el procesamiento. Por ende se trabajó con un tamaño mínimo de entre 10 píxeles (1 ha) y 100 píxeles (10 ha), para las distintas pruebas. Establecer el tamaño mínimo es muy importante dado que los algoritmos de región creciente iteraron hasta alcanzarlo, y siguieron si otro parámetro lo permitía.

Lennert [148]] menciona que el *umbral* se escala al rango de todos los datos de entrada (i.e., toda la imagen), no a la región donde opera el algoritmo (que puede ser sólo una porción de estudio o interés⁷). Esto permitió la aplicación del mismo umbral a las subregiones o zonas morfológicas, asegurando que este umbral tenga el mismo significado en toda la imagen [148]. El umbral entonces definió esa distancia máxima (que asignó al valor umbral = 1) a partir los datos de toda la imagen de 400×400 píxeles. Luego, el algoritmo segmenta las zonas morfológicas (reinterpretadas como región computacional) parametrizado con ese umbral. Esto hace al *umbral* comparable entre zonas, y permite la optimización (con las *Sum* de Espíndola o la *F* de Johnson) del mismo por zona.

En el caso del algoritmo MeanShift, además de umbral y tamaño mínimo de objeto, se establecieron los parámetros de radio *r*, definiendo las “zonas de interés”: *hrs* (espectral) y *radius* (espacial). Se comenzó con valores bajos de *hrs* entre 0 y 0,1. Para la zona de interés especial, i.e., parcelas hortícolas, se probaron los siguientes radios espectrales 1,5; 2,5 y 3,5.

Basado en el rango que se estableció para cada parámetro, la herramienta **i.segment.uspo** [125] crea una serie de segmentaciones que luego son evaluadas usando las funciones de optimización *F* [143] o *sum* [141]. Esta herramienta permite una optimización por zonas morfológicas; dado que “la presencia de valores extremos atípicos entre los parámetros de segmentación optimizados es un indicador de que no se debe utilizar un único parámetro para toda la escena” [125].

Las semillas con las que inician los algoritmos usados en esta tesis fueron todos los píxeles de la imagen usado en esta tesis, inicia con todos los píxeles de la imagen como semillas. Asimismo el módulo de GRASS [140] permitió la posterior adición de semillas por el usuario, como método alternativo de siembra, dando inicio a las pruebas de segmentaciones jerárquicas. El uso de *i.segment.uspo*, permitió automáticamente guardar la primera segmentación optimizada, y reinició el ciclo de fusión de segmentos, en base a los segmentos de esta primera.

⁷ En GRASS GIS se trabaja con *región computacional*, es decir los rasters son ingresados de manera completa, pero solo realiza las operaciones en la región computacional definida por el usuario, que puede ser considerada como la región de interés (ROI). Las salidas de GRASS tienen su contorno y resolución igual a la de la región computacional usada. (Notas propias de clase del curso de GRASS Dra. Veronica Andreo).

3.2.1.3. Módulo: Clasificación de segmentos

3.2.1.3.1. Fuente de datos para clasificación Se generaron tres conjuntos de datos para clasificar fruto de la evaluación de las características, se presentan en la Tabla 3.3.

3.2.1.3.2. Caracterización de segmentos Para caracterizar los segmentos se trabajó con las medidas morfológicas de tamaño y forma: área, perímetro, compacidad en comparación a un cuadrado⁸, compacidad en comparación a un círculo⁹ y dimensión fracta¹⁰. Se extrajeron además estadísticos de los datos espectrales y texturales seleccionados (Tabla 3.3) para cada clasificación: mínimo, máximo, rango, media, desvío estándar, coeficiente de varianza, mediana, primer cuartil, tercer cuartil, percentil 90.

Tabla 3.3: Combinación de datos de bandas espectrales, índices, y medidas texturales de cada conjunto de entrada de datos usados en los mejores resultados de clasificación. Se utilizaron las bandas azul (B ~ 493 nm), verde (G ~ 560 nm), rojo (R ~ 665 nm), infrarrojo cercano (NIR ~ 833 nm), índices de vegetación NDVI y NDWI, Suma Promedio (SA), Suma de Varianza (SV), Suma de Entropía (SE), Diferencia de Entropía (DE), Segundo Momento Angular (ASM), Momento de Diferencia Inversa (IDM), Correlación, Varianza, Entropía, Varianza Media-Umbra y Contraste.

Entradas\Datos	Sentinel-2A	Índices de vegetación	Medidas texturales (de NDVI)
i	10 bandas	NDVI y NDWI	
ii	G, R y NIR	NDVI y NDWI	DE, Varianza (3×3), Varianza Media-Umbra (21×21)
iii	B, G, R y NIR	NDVI y NDWI	SA, SV, SE, DE, ASM, Contraste, Correlación, Varianza, IDM, Entropía (3×3) y Varianza Media-Umbra (21×21)

3.2.1.3.3. Verdad de campo y validación cruzada La verdad de campo usada es aquella descrita en el capítulo 2. La misma fue re-etiquetada a los fines de obtener clases diferenciables y de interés para la resolución del problema de “identificar lotes hortícolas”, en 6 clases: Hortícola, Agricultura extensiva, Construcción, Arboleda, Tierra de uso mixto y Tierra arada. Dicha verdad de campo fue tratada con un método de validación cruzada de k-iteraciones. Se utilizó el paquete *caret* del software *R* para dividir los datos en conjuntos

8 $compact-square = 4 \# \sqrt{\frac{area}{perimetro}}$

9 $compact-circle = \frac{perimetro}{\sqrt{2}} \# \sqrt{r \# area}$

10 $fd = 2 \# (\log(perimetro) / \log(area + 0,001))$

de entrenamiento y validación. Se seleccionaron 5 subconjuntos (llamados *folds*) y 10 iteraciones.

La idea básica de la validación cruzada de k -iteraciones es que los datos de muestra se dividen en k subconjuntos. Uno de los subconjuntos se utiliza como datos de prueba y el resto ($k-1$) como datos de entrenamiento. El proceso de validación cruzada es repetido durante k iteraciones, con cada uno de los posibles subconjuntos de datos de prueba. Finalmente se realiza la media aritmética de los resultados de cada iteración para obtener un único resultado. La validación es la parte más importante de la construcción de un modelo supervisado. Para construir un modelo con un buen poder de generalización, se debe tener una estrategia de división de datos razonable, y esto es crucial para la validación del modelo [149].

3.2.1.3.4. Modelos de clasificación y esquema votación Para la clasificación se modificó el código en software R del módulo *v.class.mlR* [150], según los requerimientos de software, que respondieron mejor a los requerimientos de usuario. Los modelos de clasificación usados fueron no paramétricos: RF y SVM (ver capítulo 2). Con los resultados de las clasificaciones de los dos modelos se empleó un sistema de votación, ponderado por el *kappa* de cada modelo y por su desempeño en la identificación de clase hortícola. El módulo generó un dato tipo columna de probabilidad de la clase elegida, se modifica el código para rasterizar dicha columna y que la medida de *bondad o goodness* de clasificación acompañe el raster de segmentos etiquetados con cada clase.

3.3. Resultados y Discusiones

3.3.1. Resultado de segmentaciones

Los resultados sobresalientes del esquema de pruebas realizadas son 8. En la Tabla 3.4 se presenta un resumen de la evaluación analítica de estos resultados, donde se describen los diferentes parámetros optimizados (tamaños mínimos y umbrales); las funciones de optimización usadas; y los distintos datos de entrada (Tabla 3.2); así como también los resultados de pruebas con distintos algoritmos (RG o MS) y los obtenidos de segmentaciones jerárquicas. Los resultados son presentados también para su análisis visual, en comparación con una digitalización de segmentos sobre la imagen Sentinel-2A (Fi-

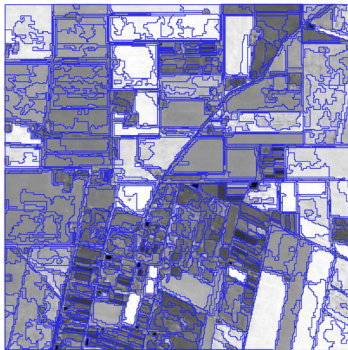
gura 3.6). Dicha segmentación estima 215 objetos en la escena: en la zona morfológica 1 se identificaron 130 objetos (tamaño mínimo 40 px, medio 2.820 px, mediana 1.581 px) y en la zona 2, 75 objetos (tamaño mínimo 530 px, medio 15.672 px, mediana 10.804 px). La nomenclatura de las pruebas tiene el siguiente formato:

<algoritmo>_<conjunto de parámetros>_<entradas de segmentación>_<condición jerárquica>_
algoritmo: Region Growing (RG) o Mean Shift (MS).
conjunto de parámetros: Indicado por un número (1, 2, 3, 4 o 5) en ellos varía algún parámetro del ajuste de algoritmo de segmentación (α , tamaño mínimo de segmento, umbral) o la función de optimización elegida (F o sum).
entradas de segmentación: Diferentes fuentes de datos para segmentar (Tabla 3.2).
condición jerárquica: Identificador de esquema de segmentación jerárquica o no: **-H**

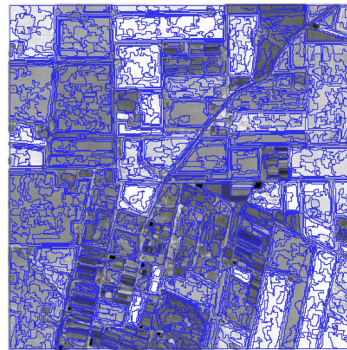
De las pruebas de segmentación realizadas las que dieron mejores resultados (visuales y analíticos) son las últimas tres: **RG4 Entrada 3-H** (Figura 3.7(f), Tabla 3.4); **RG4 Entrada 4** (Figura 3.7(g), Tabla 3.4) y **RG4 Entrada 5-H** (Figura 3.7(h), Tabla 3.4). Fueron analizadas visualmente comparando las segmentaciones en forma, cantidad y sentido geográfico de los segmentos (con énfasis en la zona 1 de lotes hortícolas) respecto de la laboriosa segmentación manual. La evaluación analítica se realizó mediante la comparación de los resultados de la función de optimización F de Jonhson. Para ello se modificó umbrales y tamaños mínimos en reiteradas ocasiones, a fin de obtener resultados de F superiores a 0,85. Además se evaluó los tiempos de ejecución. Resultaron tres las pruebas de segmentación elegidas, con las que se continuó el esquema de pruebas de clasificación.



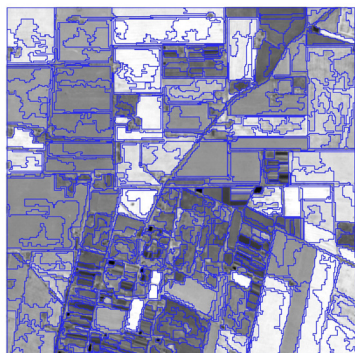
Figura 3.6: Segmentación Manual sobre imagen Sentinel-2A. Identifica 215 objetos totales. Zona 1: 130 objetos (tamaño mínimo 40 px, medio 2.820 px, mediana 1.581 px). Zona 2: 75 objetos (tamaño mínimo 530 px, medio 15.672 px, mediana 10.804 px).



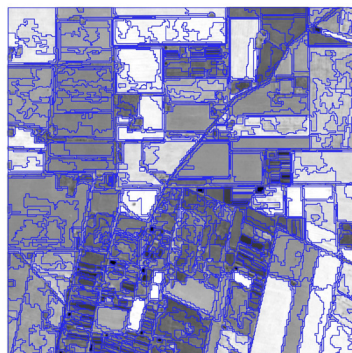
(a) RG 1 Entrada 1



(b) RG 1 Entrada 2



(c) RG 1 Entrada 2



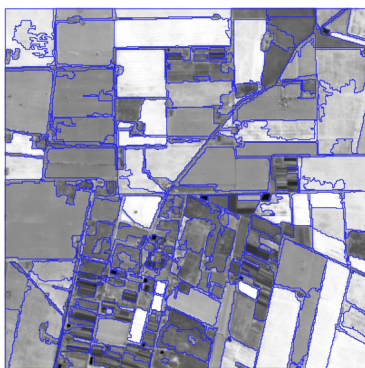
(d) RG 1 Entrada 2



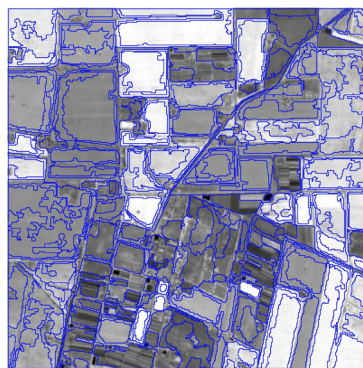
(e) MS 1 Entrada 1



(f) RG 4 Entrada 3 -H (191 objetos)



(g) RG 4 Entrada 4 (245 objetos)



(h) RG 4 Entrada 5 -H (275 objetos)

Figura 3.7: Resultado de pruebas de segmentación, sobre escena Sentinel-2 en periurbano de Córdoba.

Tabla 3.4: Resumen analítico y visual de los resultados destacados de las pruebas de segmentación, sobre imagen Sentinel-2 del periurbano de Córdoba. Detalle de método, parámetros, número de objetos y medidas de optimización. Objetos: número de objetos; OPT: función de optimización, precisión general; Umbral: umbral optimizado por zona morfológica; Minsize: tamaño mínimo de objeto (píxeles) optimizado; α : parámetro de peso en función F de Jonhson o sum para la función de optimización de Espíndola.

Prueba	Método	Objetos	OPT	Umbral	Minsize	α /Sum	Análisis Visual y Observaciones
RG1 Entrada 1 (Figura 3.7(a))	RG	1172	0,61	0,04	20	1	ALTA SOBRE-SEGMENTACIÓN Se ajustó mejor en la zona 2, y las formas de los segmentos tienen sentido geográfico (lote agrícola extensivo), los objetos en la zona 2 son internamente e más homogéneos y disímiles de los otros, en términos espectrales. En la zona 1, también sobre-segmentó los lotes hortícolas y encuentra objetos sin sentido (unió parcelas hortícolas con arboledas o bordes de caminos).
			0,72	0,04	20		
RG1 Entrada 2 (Figura 3.7(b))	RG	1144	0,52 0,55	0,04 0,02	50 50	1	ALTA SOBRE-SEGMENTACIÓN Los objetos estos son de menor tamaño y con menor sentido geográfico. Las OPT en zona 1 y 2, son los valores más bajos de las pruebas mostradas, es una segmentación de baja calidad.
RG2 Entrada 1 (Figura 3.7(c))	RG	557	0,91 0,93	0,04 0,04	80 60	0,01	SOBRE-SEGMENTACIÓN Con un α = 0,1 se otorgó más peso a la heterogeneidad entre segmentos que a la homogeneidad intra-segmen- to, se obtienen objetos más grandes y heterogéneos que los precedentes.
RG3 Entrada 1 (Figura 3.7(d))	RG	1327	1,26 1,44	0,04 0,04	20 20	Sum	ALTA SOBRE-SEGMENTACIÓN Si bien los valores de umbral, núm. objetos y tamaño mínimo son similares a RG1 Entrada 1, resultó aún más segmentada. Los valores sum de zona 1 y 2 son buenos, comparados con los reportados por Espíndola (47). Entonces, es una medida de precisión que resulta escasa como OPT para ponderar uno de los dos factores de la ecuación, sobre el otro
MS1 Entrada 1 (Figura 3.7(e))	MS	60	0,93 0,97	0,01 0,04	60 50	0,01	ALTA SUB-SEGMENTACIÓN El costo computacional fue mayor en decenas de minutos de proce- samiento respecto de las otras pruebas. Encuentra 60 objetos altamente disímiles de sus vecinos .
RG4 Entrada 3-H (Figura 3.7(f))	RG	191	0,91 0,92	0,18 0,18	70 90	0,01	Leve SUB-SEGMENTACIÓN Muy buena segmentación, pero subdivide los lotes hortícolas. La función de optimización supera el 0,9 para ambas zonas. Se optimizó con F y un $\alpha = 0,1$ en buscó objetos más heterogéneos internamente. Se trabajó con una segmentación JERÁRQUICA.
RG4 Entrada 4 (Figura 3.7(g))	RG	241	0,88 0,96	0,14 0,19	70 30	0,01	BUENA SEGMENTACIÓN Los resultados visuales y analíticos son muy buenos. Visualmente es mejor que la anterior en la zona 1 (hortícola), definiendo mejor los límites de los lotes y acercándose al número de objetos identificados manualmente (215, Figura 3.6).
RG4 Entrada 5-H (Figura 3.7(h))	RG	275	0,93 0,96	0,17 0,19	80 45	0,01	Leve SOBRE-SEGMENTACIÓN Se realiza una segmentación jerárquica. Genera sobre-segmenta- ción de los bordes de los lotes. Buena definición de lotes en la zona 1.

3.3.2. Resultados de clasificaciones

Sobre las tres segmentaciones elegidas se aplicaron los tres conjuntos de datos para clasificar (sección 3.2.1.3.1). Generando así 9 resultados de clasificación (Tabla 3.5). Cabe recordar que cada clasificación es resultado de la aplicación de una votación entre algoritmos RF y SVM. Cada clasificación es nominalizada como:

<clasificador (RF o SVM)>_<conjunto de clasificación (i, ii o iii)>_<segmentación base (f, g o h)>.

De las 9 clasificaciones, las que mejores resultados fueron *SVM i f*, *RF ii g* y *RF i h*, por lo que sólo se analizarán y/o mostrarán resultados de estas. Las restantes 6 clasificaciones resultantes no fueron satisfactorias, dado que sus desempeños fueron inferiores y los resultados visuales no coinciden con los esperados. Del resultado *SVM i f*, se observa en la Figura 3.8 (a), una clasificación que tiene sentido respecto de lo esperado en la escena. En la matriz de confusión (Tabla 3.6), las métricas de precisión general son buenas y similares a las otras clasificaciones ($OA = 0,73$ y $\kappa = 0,66$), pero el desempeño respecto a la identificación de la clase Hortícola es el más bajo de las tres clasificaciones ($FIScore = 0,65$).

La segunda clasificación exhibida: *RF ii g*, en la Figura 3.8 (b), resulta la más consistente en cuanto al análisis visual (respecto de lo esperado en la escena, Figura 3.6). Presenta un mayor equilibrio entre las métricas de precisión general ($OA = 0,75$ y $\kappa = 0,64$) y la identificación de la clase Hortícola. Esta última medida es la mayor de las tres clasificaciones, reportando un $FIScore = 0,77$ (Tabla 3.7). Resulta interesante rescatar que el mejor desempeño de la clasificación se obtuvo al incorporar las capas de textura (DE, Varianza y Varianza Media-Umbra) y reducir las bandas espectrales sólo a Verde, Rojo, Infrarrojo Cercano (NIR), respecto de los otros conjuntos de datos (sección 3.2.1.3.1).

La tercer clasificación mostrada: *RF i h*, (Figura 3.8 (c)) es la peor en el análisis visual de las tres elegidas (siendo la mejor de las tres obtenidas a partir de la segmentación *RG4 Entrada 5-H*). Si bien la métrica de precisión global κ es la mejor, y su OA levemente menor a *RF ii g* ($OA = 0,74$ y $\kappa = 0,67$), lo que se tomó en cuenta es la específica de clase Hortícola $FIScore = 0,76$ (Tabla 3.8).

Tabla 3.5: Esquema de pruebas de 9 clasificaciones, a partir de 3 segmentaciones seleccionadas (f, g, h) y 3 conjuntos de fuentes de datos de clasificación (i, ii, iii).

		Segmentaciones		
		RG4 Entrada 3-H (f)	RG4 Entrada 4 (g)	RG4 Entrada 5-H (h)
Datos de entrada para clasificar	i	SVM i f	RF i g	RF i h
	ii	RF ii f	RF ii g	RF ii h
	iii	RF iii f	RF iii g	RF iii h

Tabla 3.6: Matriz de confusión de clasificación SVM i f de Figura 3.8(a) (25 remuestreos, repetidos 10 veces). Las entradas son recuentos porcentuales promedio de clases a través de las re-muestras. OA = 0,73 ; kappa κ 0,66

Predicción	Referencia						
	Tierra Arada	Arboleda	Construcción	Extensivo	Hortícola	Mixto	Total Predicción
Tierra Arada	2,7	0	0,1	0	1,3	0	4,1
Arboleda	0	11,5	2,2	0,9	0,2	1,2	16
Construcción	0	1,3	10,2	0	3,1	2,6	17,2
Extensivo	0	0	0	27,9	0	0	27,9
Hortícola	4,3	0	3	0	17,7	2,8	27,8
Mixto	1,2	0	1,9	0,2	0,9	2,8	7
Total Referencia	8,2	12,8	17,4	29	23,2	9,4	100
Recall	0,33	0,9	0,59	0,96	0,76	0,3	
Precisión	0,66	0,72	0,59	1	0,64	0,4	
F1 Score	0,44	0,8	0,59	0,98	0,65	0,34	

Tabla 3.7: Matriz de confusión de clasificación RF ii g de Figura 3.8(b) (25 remuestros, repetidos 10 veces). Las entradas son recuentos porcentuales promedio de clases a través de las re-muestras. OA = 0,75 ; kappa κ 0,64

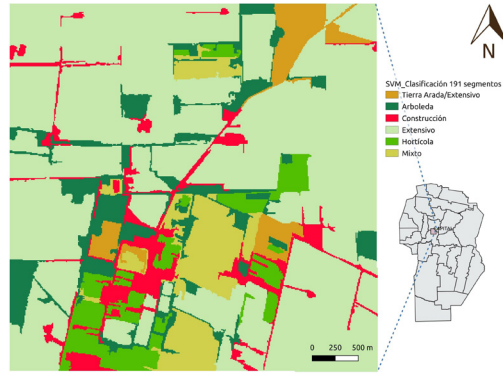
<i>Predicción</i>	<i>Referencia</i>						
	Tierra Arada	Arboleda	Construcción	Extensivo	Hortícola	Mixto	Total Predicción
Tierra Arada	4,8	0	0	0	1,7	1,1	7,6
Arboleda	0	9,1	3,6	0	0,3	1,3	14,3
Construcción	0	1,3	11	0	2,3	1,7	16,3
Extensivo	1,1	0	0	25,6	0,2	0	26,9
Hortícola	3,3	0,9	3	0	19,9	0,6	27,7
Mixto	0,8	0,9	0,2	1,1	0	4,2	7,2
<i>Total Referencia</i>	10	12,2	17,8	26,7	24,4	8,9	100
<i>Recall</i>	0,48	0,75	0,62	0,96	0,82	0,47	
<i>Precisión</i>	0,63	0,64	0,67	0,95	0,72	0,58	
<i>F1 Score</i>	0,54	0,69	0,64	0,95	0,77	0,52	

Tabla 3.8: Matriz de confusión de clasificación RF i h de Figura 3.8(c) (25 remuestros, repetidos 10 veces). Las entradas son recuentos porcentuales promedio de clases a través de las re-muestras. OA = 0,74 ; kappa κ 0,67

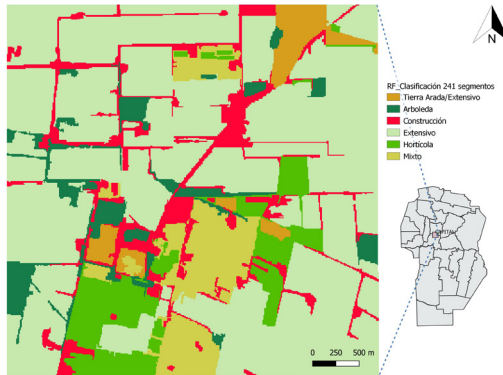
<i>Predicción</i>	<i>Referencia</i>						
	Tierra Arada	Arboleda	Construcción	Extensivo	Hortícola	Mixto	Total Predicción
Tierra Arada	6,5	0	0,3	0	1	0	7,8
Arboleda	0	6,9	3,8	2,2	0	0,5	13,4
Construcción	0,8	2,9	8,7	0	0,3	0,4	13,1
Extensivo	0	2,1	0	29,8	0	0	31,9
Hortícola	3,3	0,1	2,4	0	17,3	2,1	25,2
Mixto	0	1,2	0,1	1	1,7	4,6	8,6
<i>Total Referencia</i>	10,6	13,2	15,3	33	20,3	7,6	100
<i>Recall</i>	0,61	0,52	0,57	0,9	0,85	0,61	
<i>Precisión</i>	0,83	0,51	0,66	0,93	0,69	0,53	
<i>F1 Score</i>	0,7	0,51	0,61	0,91	0,76	0,57	

3.3.3. Discusión de resultados

La mejor clasificación GEOBIA resulta **RF ii g**, Figura 3.8(b). Ésta se compara con la metodología de clasificación del capítulo 2 “clasificación espectral-textural por píxel” (sobre esta porción de la escena y con el re-etiquetado de la verdad de campo), presentada en la Figura 3.9. Ambas se evalúan visual y analíticamente, los resultados visuales con GEOBIA responden mejor a identificar lotes hortícolas y disminuyen la sobre-estimación de la clase Hortícola respecto de la clasificación por píxeles siguiendo el método del capítulo 2. Al comparar los resultados analíticos generales de las clasificaciones (GEOBIA versus Espectral Textural de PÍXEL) el mayor κ fue de GEOBIA con 0,64 y la otra obtuvo $\kappa = 0,6$. Si se analizan y comparan las métricas de exactitud sólo de clase **Hortícola**, la clasificación GEOBIA resulta $FIScore = 0,77$ frente a Espectral-Textural de Píxel $FIScore = 0,79$ (Tabla 3.7 y Tabla 3.9). Vlachopoulos *et al.* (153) clasificaron avena, maíz y cebada comparando GEOBIA y clasificación basada en píxeles (PBI). En promedio, la precisión de la clasificación general que alcanzaron con PBI fue ligeramente inferior a aquella con GEOBIA (97,06% frente a 97,49%).



(a) SVM i f (191 objetos)



(b) RF ii g (241 objetos)



(c) RF i h (275 objetos)

Figura 3.8: Clasificaciones votadas según cada segmentación elegida y distintas fuentes de datos. En base a imagen Sentinel-2A del 19/02/2019 de Córdoba.

Tabla 3.9: Matriz de confusión de la clasificación Random Forest (RF) espectral-textural per pixel sobre muestra 400×400 px de Sentinel-2. Las entradas son recuentos porcentuales promedio de clases a través de las re-muestras (25 remuestros, repetidos 10 veces). OA = 0,69 ; kappa κ 0,6

Predicción	Referencia						
	Tierra Arada	Arboleda	Construcción	Extensivo	Hortícola	Mixto	Total Predicción
Tierra Arada	2,86	0	0	0	2,86	0	5,70
Arboleda	0	8,57	0	0	0	2,86	11,43
Construcción	0	5,71	8,57	0	2,86	0	17,14
Extensivo	0	2,86	0	17,14	0	2,86	22,86
Hortícola	0	0	2,86	0	28,57	2,86	34,29
Mixto	2,86	2,86	0	0	0	2,86	8,58
Total Referencia	5,70	20	11,43	17,14	34,29	11,44	100
Recall	0,5	0,43	0,75	1	0,75	0,25	
Precisión	0,5	0,75	0,5	0,75	0,83	0,33	
F1 Score	0,5	0,55	0,6	0,86	0,79	0,28	

Llaman la atención los resultados analizados solo por las métricas. El κ obtenido (0,64) es inferior a los reportados por la clasificación del capítulo 2 (κ = 0,86, en Tabla 2.9), y levemente superior al κ = 0,6 de la clasificación de pixel realizada en esta escena. Esta última clasificación surge del re-etiquetado en 6 clases, siguiendo la metodología propuesta en el capítulo 2. Sin embargo, estas métricas deben ser analizadas a la luz de la relación muestra/población sobre la que se evalúan en cada clasificación. Es decir, que esos κ están reportados sobre una proporción de análisis de la imagen, y estas nos dan mayor representatividad en el esquema GEOBIA dado que trabaja con objetos y no píxeles, como se ve en la Tabla 3.10.

Tabla 3.10: Relación píxeles de la verdad de campo (muestra) y el total de la escena (población) en las clasificaciones GEOBIA (A), Espectral-Textural de Pixel (escena 400 x 400 píxeles) capítulo 3 (B) y Espectral-Textural de Pixel del total de Interfaz Rural-Urbana capítulo 2 (C).

	Clasificaciones		
	GEOBIA (A)	Pixel Escena 400 x 400 (B)	Pixel Interfaz Rural-Urbana (C)
Muestra (m)	95.540 px	125 px	1886 px
Población (p)	160.000 px	160.000 px	58.912.000 px
Relación m/p	60%	0,08%	0,003%

Vale la pena resaltar que al evaluar conjuntamente las métricas de precisión global, las particulares de precisión hortícola y un análisis visual de la clasificación resulta en un mayor equilibrio de desempeño la clasificación por **GEOBIA RF ii g**.

El resultado elegido *RF ii g*, obtenido mediante la cadena de procesamiento (Anexo B) es fruto de una segmentación de región creciente a partir de 4 bandas espectrales de Sentinel-2, que luego fue clasificado por un modelo *Random Forest* alimentado con 3 bandas espectrales, 5 índices y texturas. Dicha clasificación obtuvo una precisión general de $OA = 0,75$ y un $\kappa = 0,64$; específicamente la clase Hortícola alcanzó un $FIScore = 0,77$. Dichos resultados se comparan con trabajos recientes similares. El trabajo referente de Grippa *et al.* [125] en Burkina-Faso obtuvo una precisión general de $OA = 0,81$ y $\kappa = 0,79$; trabajaron con la clase ‘vegetación y suelo desnudo mixto’ similar a la hortícola buscada en esta tesis, ellos lograron un $FIScore = 0,56$. Específicamente sobre vegetales, Lebourgeois *et al.* [152] obtiene un $FIScore = 0,57$, mediante un esquema combinado de clasificación de RF y GEOBIA para el mapeo de la agricultura en pequeña escala sobre una serie temporal simulada Sentinel-2, VHRS y DEM. Recientemente Su & Zhang (2020) [122] clasificaron una Worldview-2 de resolución espacial 2m, con RF y GEOBIA e identificaron una clase ‘vegetales’ (compuesta por campos con zucchini o tomate o repollo) alcanzaron una $OA = 0,91$ y un $FIScore = 0,61$ para vegetales. Además aportan una discusión interesante entorno a la ampliación automática de muestras de entrenamiento, señalándola como una limitante de las clasificaciones GEOBIA. Tian *et al.* (146) aplicaron GEOBIA a una serie temporal de ALOS PALSAR (banda L, alta resolución) alcanzaron una $OA = 0,76$ y específicamente en la clase vegetales una precisión de 0,7. Siguiendo la propuesta de este estudio, a futuro los datos de SAOCOM (Banda L) podrían sumar información de relevancia para clasificaciones hortícolas. Si bien los resultados de precisión general de la clasificación son inferiores a las referencias, la precisión alcanzada al identificar horticultura es prometedora y muestra un adelanto respecto de los trabajos de GEOBIA actuales.

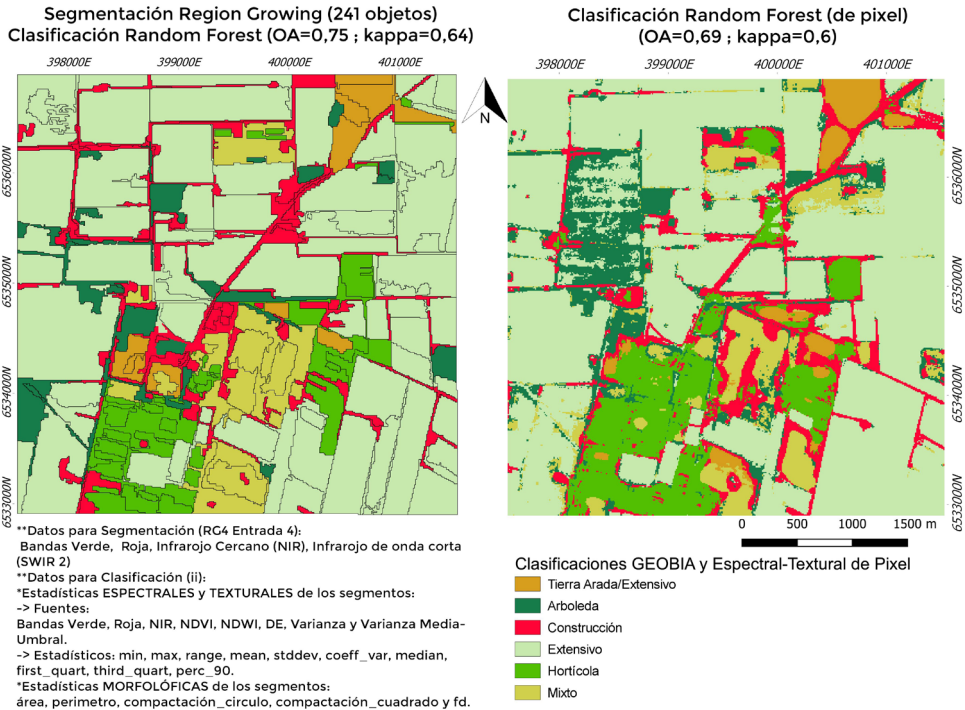


Figura 3.9: Clasificación GEOBIA versus clasificación Espectral-Textural de Píxel. Sobre escena Sentinel-2A (19/02/2019), del periurbano de Córdoba Capital (zona de El Quebrachal).

Se discute a futuro la posibilidad de usar métodos automatizados de ampliación de muestras de entrenamiento para el abordaje de toda la interfaz rural-urbana, así como la necesidad de selección de variables y sobre todo la clasificación GEOBIA sobre una serie temporal de Sentinel-2, para alcanzar mayores exactitudes en la predicción. Estos estudios indican que RF tiene un rendimiento de clasificación de vanguardia para el mapeo de campos de cultivo basado en objetos [125; 122; 151; 152]. Un campo no explorado y limitante en la cadena de procesamiento seguida es que RF es capaz de cuantificar la importancia de las características, lo cual es crucial en OBIA, ya que existen varios tipos de características basadas en objetos que contienen no solo información espectral y de textura, sino también señales geométricas y contextuales [122].

3.4. Conclusión

El resultado elegido de clasificación fue **RF ii g**, partiendo de una segmentación *RG4 Entrada 4*, es una interpretación de la escena con un muy buen desempeño analítico y visual. Dicho resultado de clasificación *Random Forest* requirió 3 bandas más 5 índices espectrales y texturales (ver sección 3.2.1.3.1, **ii**). La misma surgió de la segmentación de 241 objetos a partir de *region growing* sobre las entradas de 4 bandas de Sentinel-2 (Entrada 4: G, R, NIR y SWIR2, ver sección 3.2.1.2.1) y con parámetros $\alpha = 0,01$, tamaño mínimo y umbral optimizados. La eficiencia de uso de los datos de entrada y el tiempo de ejecución del proceso de segmentación, es también evaluado positivamente. Del análisis visual se desprende que con la aproximación GEOBIA se logran mejoras importantes en cuanto a la identificación de lotes (especialmente hortícolas) y a la eliminación del “efecto sal y pimienta” al interior de los mismos.

Se alcanzó el objetivo de explorar la herramienta GEOBIA y se disponibilizó un producto operacional que responde con mayor precisión a *¿Dónde?* se producen las hortalizas. La sucesión automatizada parte de una imagen satelital, genera índices espectrales y filtros texturales. Se definió los algoritmos, los datos fuente para cada instancia y parámetros óptimos. Se logra segmentar la imagen en lotes agrícolas en una zona periurbana, con espacial énfasis en los campos hortícolas, lo que permite su identificación y caracterización.

Vale la pena resaltar que no solo deben evaluarse métricas generales y particulares de precisión, sino también los resultados esperados (el patrón de distribución de clases), mediante el análisis visual. Asimismo el resultado GEOBIA, claramente mejora la representatividad de las métricas de precisión al incrementar la relación muestra/población, sin necesidad de incrementar las (costosas y poco accesibles) salidas a campo.

Al comparar la precisión, de las clasificaciones de cobertura y uso de suelo, alcanzada con GEOBIA frente a las clasificaciones basadas en Píxel, se observa una tendencia positiva que indica a GEOBIA como un método innovador con mucho futuro para la clasificación de cultivos hortícolas, especialmente en el Instituto “Mario Gulich”. Hoy ambos abordajes alcanzan la misma pre-

cisión evaluados en la detección de horticultura, teniendo en cuenta que el abordaje de pixel cuenta con ya más de cincuenta años y GEOBIA apenas unos veinte. Este futuro de GEOBIA implica más pruebas, uso de otras plataformas de datos y mayor capacidad de cómputo. Esta tesis apuesta al futuro de las clasificaciones GEOBIA y logra ubicarlas competitivamente respecto a las clasificaciones basadas en pixel, que se encuentran en su mejor momento.

Capítulo 4

Clasificación hortícola en base a imágenes de alta resolución espacial y uso de Vehículo Aéreo No Tripulado (VANT)

4.1. Introducción

En este capítulo se exploran datos de sensado remoto de alta resolución espacial a los fines de responder a la pregunta de investigación **¿Qué se produce?**, es decir una clasificación de cobertura de detalle que indique especies hortícolas en producción. Para ello se usa un *vehículo aéreo no tripulado* (VANT) con una cámara multispectral de resolución submétrica (Parrot Sequoia+), también llamadas VHRI por su nombre en inglés *very high resolution images*. La utilización de VANT o drones, con sensores a bordo y el procesamiento de las imágenes adquiridas representan un nuevo desafío técnico/científico que será abordado en este capítulo. A fin de caracterizar el uso y cobertura de suelo intra-lote hortícola se planifica y ejecutan vuelos programados, y se procesan las imágenes submétricas multispectrales obtenidas.

4.1.1. VANT en Agricultura

Los Vehículos Aéreos No Tripulados (VANT) son vehículos aéreos pilotados de forma remota, que tienen un papel importante en los sectores de defensa y actualmente en el ámbito civil y comercial. Se utilizan en una serie de aplicaciones que están revolucionando muchos servicios públicos, incluyendo monitoreo en tiempo real, búsqueda y rescate, encuestas de vida silvestre, servicios de entrega, cobertura inalámbrica y agricultura de precisión, topografía y mapeo, teledetección aérea, entre otras para diversas industrias, como petróleo y gas, construcción, plantas de energía y ferrocarriles.

Los VANT, (o UAV por sus siglas en inglés), también denominados comúnmente drones, son la plataforma sobre la que se puede montar un sensor para

observación remota (cámara). En el ámbito académico, un perfil bibliográfico del Remote Sensing Open Access Journal, reporta Vehículo Aéreo No Tripulado (VANT)¹ como la palabra clave más influyente entre 2009 y 2018 [154]. En los últimos tiempos, los VANT se hicieron cada vez más populares en la agricultura de precisión (131). Se estima que el 80% del uso doméstico de los VANT en los Estados Unidos se destina a la teledetección de cultivos y a la agricultura de precisión [156]. La tecnología de teledetección basada en los vehículos aéreos no tripulados está siendo utilizada para la predicción del rendimiento [157], en especial las imágenes de infrarrojo cercano (NIR) en detalle, ayudan a determinarlo y alertar las situaciones de estrés del cultivo [156]. En especial se detectan aquellos estrés debidos a deficiencias de agua y nutrientes, enfermedades y plagas [158]. El objetivo de la agricultura de precisión es optimizar el uso del agua, los nutrientes, los plaguicidas y los herbicidas mediante la gestión de los campos de cultivo en lugares específicos (i.e. *sitio específicos*). Proveer dicha información a nivel cultivo (en detalle y en el momento), en teoría, ayuda a disminuir el uso de insumos, reduciendo así el costo asociado con la producción, optimizando el agua y reduciendo los impactos ambientales de la sobre-fertilización y la amplia aplicación de pesticidas.

Los VANT tienen la capacidad de realizar algunas tareas difíciles o peligrosas, con alta movilidad, seguridad y bajo costo (comparativo). Los VANT para la teledetección es son mucho más baratos que las aeronaves tripulados y los satélites, y proporcionan datos de mayor resolución debido a la baja altitud a la que pueden operar [159]. Estos vehículos pueden operar bajo nubes, y por lo tanto proporcionan datos de alta resolución espacial y con una revisita a demanda del usuario, con una disminución significativa del tiempo de respuesta respecto de los aviones tripulados y los satélites.

A pesar de sus potencialidades, los VANT aún no han visto un uso generalizado por parte de los agricultores y los usuarios finales para el uso en agricultura de precisión de forma rutinaria. Muchos de los proyectos en curso se limitan a una prueba de concepto o demostración de tecnología. Capacidades para su adopción por la industria agrícola y los agricultores para el monitoreo sanitario de los cultivos y el diseño de planes de tratamiento óptimos y específicos para cada lugar no se han desarrollado completamente. Esto se debe a dos razones principales: en primer lugar, no se ha abordado suficientemente

¹ El reporte menciona la palabra clave como "UAV", por sus siglas en inglés.

la cuestión de la exactitud y fiabilidad de los datos de teledetección para determinar el estrés de las plantas debido a la deficiencia de agua y nutrientes. En segundo lugar, el sistema general es muy complicado y costoso para los agricultores o usuarios finales, ya sea para su uso diario o rutinario [156]. Los VANT pueden equiparse con una amplia gama de sensores de imagen útiles para describir de forma fiable la variabilidad espacial en los diferentes cultivos. Las cámaras RGB y multiespectrales son los sensores más utilizados en aplicaciones agrícolas de precisión. La cámara Parrot Sequoia+, usada en esta tesis, adquiere imágenes en el espectro visible (RojoVerde-Azul (RGB)) con un sensor, y también con otros sensores multiespectrales, capta el infrarrojo cercano (NIR), borde rojo (Red Edge), rojo (Red) y verde (Green). Las imágenes multiespectrales de teledetección en el infrarrojo cercano son útiles para caracterizar el estado vegetativo de los cultivos a través del cómputo de índices espectrales, tales como el Índice de Diferencia de Vegetación Normalizada (NDVI) ampliamente utilizado y correlacionado con la biomasa y la actividad fotosintética [157].

En este estudio se operativiza un VANT, se investiga el uso de las imágenes provenientes del sensor montado en el VANT (Parrot Sequoia+). Se determina especies hortícolas cultivadas a partir de la segmentación y clasificación de imágenes agrícolas. Generando la línea base para una futura estimación de biomasa. Se detalla el instrumento (VANT) y la imagen empleada, así como la metodología adoptada y los análisis realizados en el presente estudio.

4.2. Materiales y Métodos

El sistema de sensado remoto basado en VANT debe tener características tales como costo efectivo, producción rápida, operación fácil por parte del personal local y buena precisión geométrica [156; 160]. El sistema constó de:

- una plataforma aérea (eBee SQ), tipo ala fija,
- una cámara digital con 5 sensores para la captura de la escena (Parrot Sequoia+),
- procesamiento de datos con fotogrametría digital: estructura de algoritmos de movimiento y SIG de código abierto gratuito para visualización.

Este sistema tiene la capacidad de producir imágenes con una resolución espacial <10 cm, medir el área de la parcela, evaluar el stock individual de plantas, entre otras aplicaciones. El producto básico del sistema es el ortomosaico de la escena. Los ortomosaicos permiten una interpretación visual, como la estructura de los árboles individuales, la densidad de la planta y el área límite de la parcela. Con otro equipamiento topográfico de campo se podría generar un DEM con el que se podría evaluar la información de altura de los árboles y la topografía del terreno con una alta precisión.

4.2.1. Instrumento y sensor: eBee SQ y Parrot Sequoia+

Las plataformas aéreas se dividen por su sustentación en: *ala fija* o *multirotor*. Se estudió la operación y puesta a punto del VANT de tipo ala fija para la teledetección hortícola. La aeronave está compuesta por un cuerpo central que tiene dos alas y una sola propela o hélice. Cuando está en el aire, las dos alas generan elevación que compensa el peso, permitiendo a la aeronave continuar en vuelo. Los VANT de ala fija son aerodinámicamente más estables, tienen mayor autonomía de la batería, que garantiza un tiempo de vuelo prolongado, permitiendo una cobertura de áreas de estudio extensas. El instrumento usado es *eBee SQ*, una plataforma de vuelo creada para portar una cámara multispectral *Parrot Sequoia+* (Figura 4.1).

La plataforma mide 110 cm, y pesa 1,1 kg.

La propulsión es una hélice eléctrica, que activa un motor sin escobillas. Tiene un radio de enlace de 3 km y las alas se desmontan para su transporte. La velocidad de crucero nominal es de 40 a 90 km/h, presenta una resistencia al viento de hasta 45 km/h, con una autonomía de vuelo de 55 minutos, lo que permite cubrir más de 200 hectáreas en un único

vuelo, a una altura superior a los 120 metros. Realiza multitud de actividades en vuelo, basadas en un esquema de decisiones e inteligencia artificial (IA) a bordo, analiza los datos mediante la unidad de medición inercial y el GPS de a bordo, que optimiza todos los aspectos del vuelo del eBee SQ (Figura 4.2).



Figura 4.1: Plataforma aérea de ala fija usada, eBee SQ.

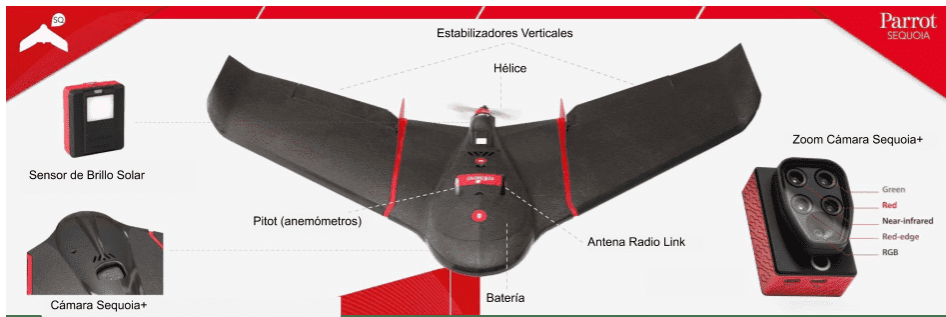


Figura 4.2: Plataforma eBee SQ en detalle.

La cámara *Parrot Sequoia+* consta de cuatro sensores: infrarrojo cercano, borde rojo, rojo y verde), más un sensor RGB. Con ello puede capturar imágenes multispectrales y RGB en un solo vuelo en forma simultánea. Cada sensor monocromático tiene una resolución de 1,2 megapixel (Mpx) en banda estrecha y sincronizada. Cada imagen multispectral tiene 1280×960 píxeles. Además, la cámara RGB captura fotos de 16 Mpx con una definición de 4608×3456 píxeles. Un resumen de sus características radiométricas se presenta en la Tabla 4.1, así como la intersección de las bandas espectrales con la reflectancia de la vegetación en la Figura 4.3. La plataforma tiene en su parte superior un sensor de brillo solar orientado hacia arriba, que permite la corrección de la señal para obtener imágenes precisas radiométricamente, independientemente de las condiciones de luz.

Tabla 4.1: Longitud de onda y ancho de banda de cada sensor multispectral de la cámara Parrot Sequoia+.

Banda	Verde	Rojo	Borde Rojo	Infrarrojo Cercano
<i>Longitud de onda</i>	550 nm	660 nm	735 nm	790 nm
<i>Ancho de banda</i>	40 nm	40 nm	10 nm	40 nm
<i>Definición</i>	1,2 Mpx	1,2 Mpx	1,2 Mpx	1,2 Mpx

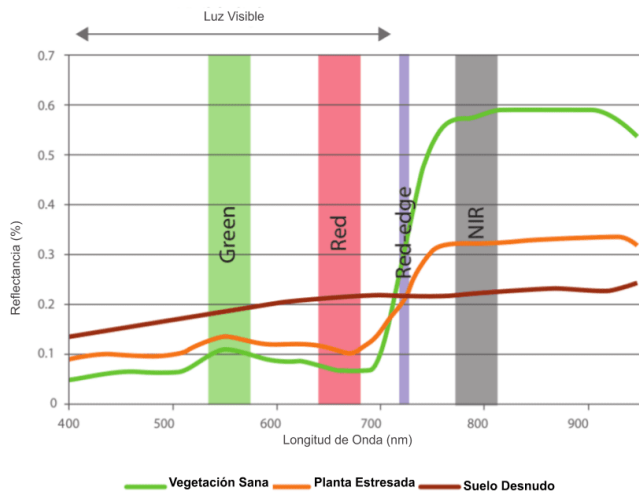


Figura 4.3: Gráfica representativa de reflectancia de vegetación e intersección con resolución radiométrica Parrot Sequoia+.

Se debe obtener un índice de solapamiento óptimo durante los vuelos a una determinada altitud, respetando un lapso de tiempo entre las capturas, se recomienda más de 70%. El sensor RGB puede hacer capturas con un intervalo mínimo de 1 segundo entre cada fotograma. Los sensores multi-espectrales pueden hacer capturas con un intervalo mínimo de 0,5 segundos entre cada fotograma.

4.2.2. Plan de vuelo y procesamiento

Como se mencionó, eBee SQ cuenta con un piloto automático programado con IA a bordo. Requirió realizar el plan de vuelo y cargarlo a la computadora de abordaje previamente. En este plan de vuelo, se establecieron parámetros como: altura máxima, radio de vuelo, centroide, se planificó el despegue y aterrizaje (dado que el VANT de ala fija precisa una plataforma lineal), y se resaltó la región que a mapear. Se realizaron misiones multi-vuelo. En las mismas, se importaron los datos de elevación y se generó un modelo digital de superficie (DSM), que permite realizar vuelos más seguros y precisos en el terreno. El programa con el que se diseña y opera la plataforma fue **eMotion**[®], es un software de gestión de vuelo (o estación en tierra) de un avión no tripulado, ver Figura 4.4.

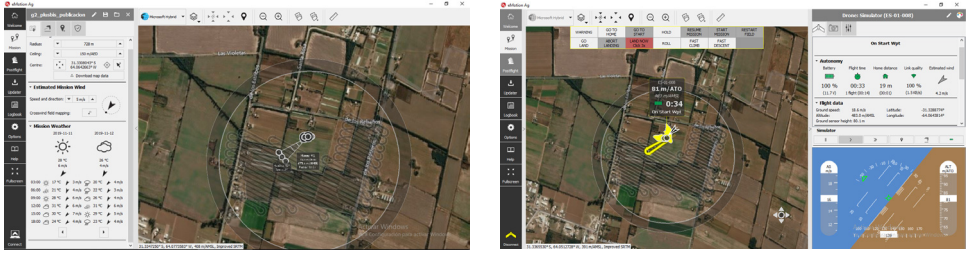


Figura 4.4: eMotion®, software de gestión de vuelo de VANT

Para establecer los parámetros se tuvieron en cuenta las restricciones legales de Argentina (cada país tiene la suya y siguiendo el marco internacional). La legislación vigente es Resolución ANAC 527/2015, Ley 25.326 – Disposición 20/2015 y Ley 27.161 – EANA (Empresa Argentina de Navegación Aérea S.E.). A nivel internacional el primer antecedente es el análisis circular 328 – anexo 190 de la Organización de Aviación Civil Internacional (OACI). Además de establecer los parámetros según las restricciones legales, el viento y las condiciones meteorológicas son un dato indispensable a la hora del diseño del vuelo. A los fines de prever la factibilidad del vuelo, se disponen de múltiples páginas web de consulta e incluso aplicaciones para móviles, como la que se ilustra en la Figura 4.5, que permite conocer condiciones meteorológicas (especialmente viento que influye sobre la dirección de vuelo, y orientación de la pista de despegue y aterrizaje), combinadas con las restricciones legales de espacio aéreo, que limita el área y altura de vuelo.

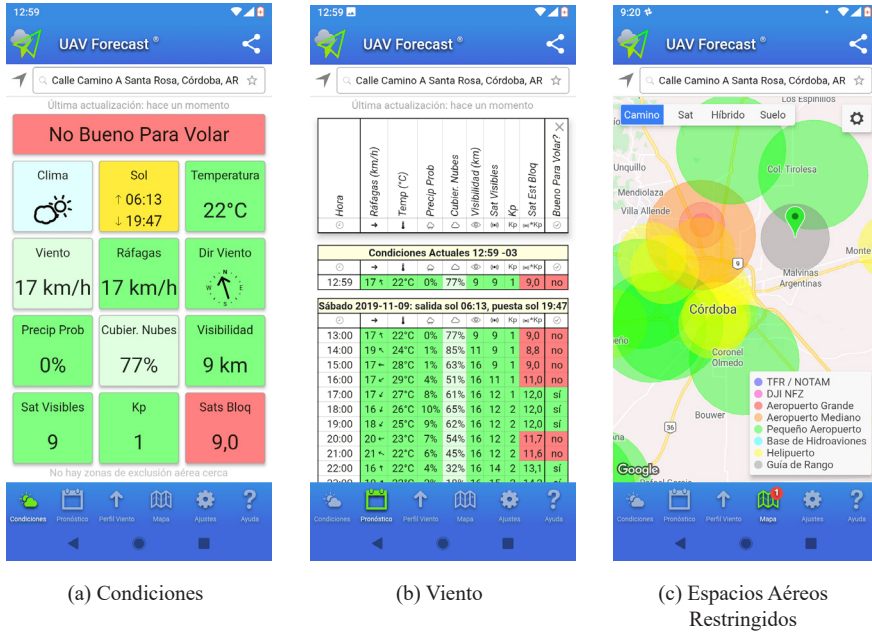


Figura 4.5: UAV Forecast, aplicación móvil para corroborar condiciones de vuelo con VANT.

En términos del equipamiento o hardware, se requirió una computadora con procesador intel core i7, con 8 GB o más de RAM, el software usado corre en Windows. El instrumental debe verificarse antes de comenzar el vuelo, con una lista de chequeos mínimos del buen funcionamiento y recepción de señal de las partes que componen al VANT (plataforma, cámara, conexión, estación terrena). En cuanto a los sensores multispectrales de *Parrot Sequoia+*, se calibraron radiométricamente para mediciones consistentes y son compatibles con RTK²/ PPK³ para resultados georreferenciados precisos. La calibración radiométrica se requiere cada vez que se inicia un vuelo, se realiza mediante tomas a unas tarjetas de calibración como se muestra en la Figura 4.6.

2 RTK (Real Time Kinematic) o navegación cinética satelital en tiempo real, es una tecnología de corrección GPS en tiempo real de los datos de ubicación cuando el VANT está capturando fotogramas del área de estudio, RTK necesita de una conexión base además de conexión WIFI.

3 PPK (Post Processed Kinematic) o navegación cinética satelital post proceso, es una técnica de corrección GPS de los datos de ubicación luego que el VANT ha capturado las fotografías. Es la usada en este estudio por ser un proceso más seguro de obtención de datos sin riesgo de perderlos.



Figura 4.6: Equipo de trabajo, ejecución de vuelo y calibración radiométrica de los sensores multispectrales Parrot Sequoia+.

4.2.3. Imágenes Parrot Sequoia+

Del vuelo se obtuvieron imágenes del sensor RGB en formato JPEG y de los sensores multispectrales (MSP) en formato TIFF. Mediante *eMotion*[®] se generó el *crudo* donde las imágenes fueron georreferenciadas (Figura 4.7) y se les adjuntó un metadato con la información de cámara, latitud, longitud, y los ángulos (alabeo, giro y cabeceo). Datos con los que el programa Pix4Dmapper[®] Ag/Pro (licencia INTA) realizó la corrección geométrica de las imágenes (Figura 4.8). Pix4D[®] es un software de procesamiento automático de imágenes terrestres y aéreas adquiridas por un VANT, que convierte los fotogramas en mosaicos georreferenciados y modelos 3D [161].

Las imágenes captadas por el eBee SQ con sus metadatos, fueron procesadas con el software Pix4D[®], y ofreció como resultado una nube de puntos (de uso interno del programa) y los ortomosaicos de cada banda MSP (Figura 4.9), RGB (Figura 4.10 (a)) e incluso el índice NDVI (Figura 4.10 (b)). La odometría del vehículo se utilizó para en el proceso automático para estimar la posición de cada nueva imagen que se incluye en el mosaico y, por lo tanto, hacer coincidir las características presentes en las áreas superpuestas de las imágenes. La navegación y la adquisición de imágenes se realizan de forma autónoma, con una ruta preestablecida en la que se define un conjunto de puntos de referencia (ver sección 4.2.2). Las imágenes se obtienen mientras el eBee SQ se mantiene flotando en cada punto de ruta, después de un giro de hélice, que le permite avanzar en la dirección y sentido preestablecido por el usuario. Los

mosaicos construidos, constan de la información proporcionada por un módulo GPS a bordo del VANT sobre la posición desde donde se adquiere cada imagen, como muestra la Figura 4.11.

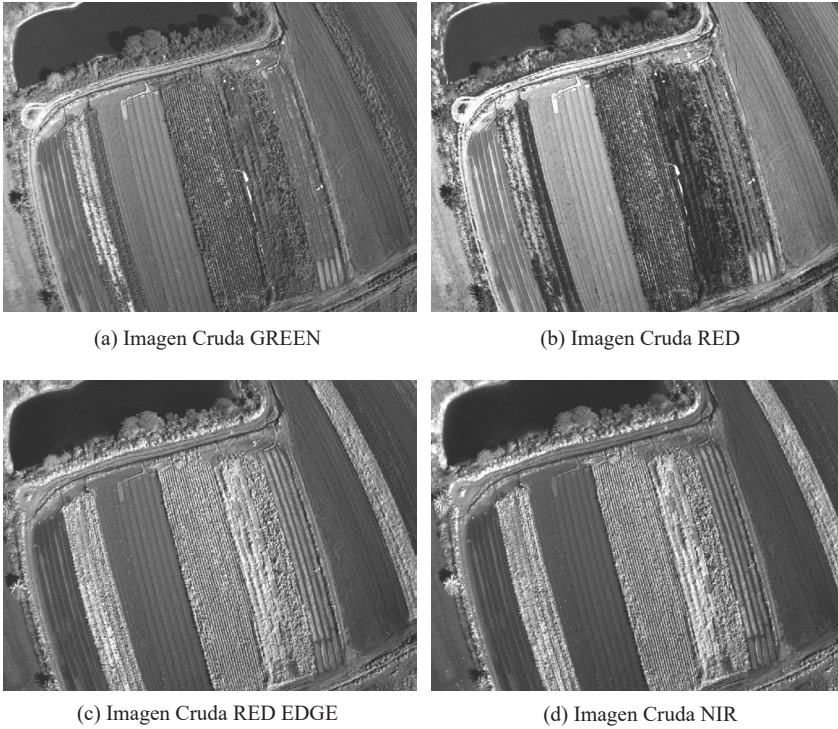


Figura 4.7: Imagen multispectral cruda por banda, tomada por Parrot Sequoia+, sobre campo hortícola en Córdoba Capital.

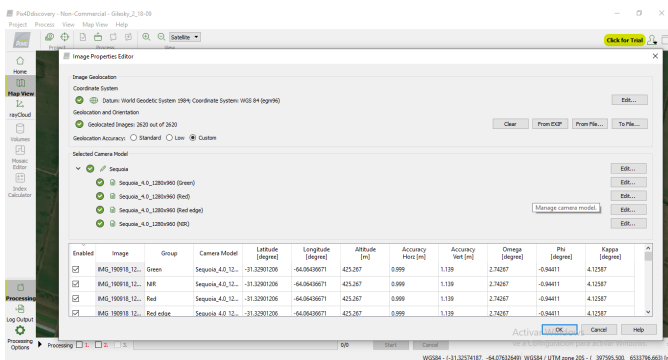


Figura 4.8: Observación de los metadatos de las imágenes del vuelo.

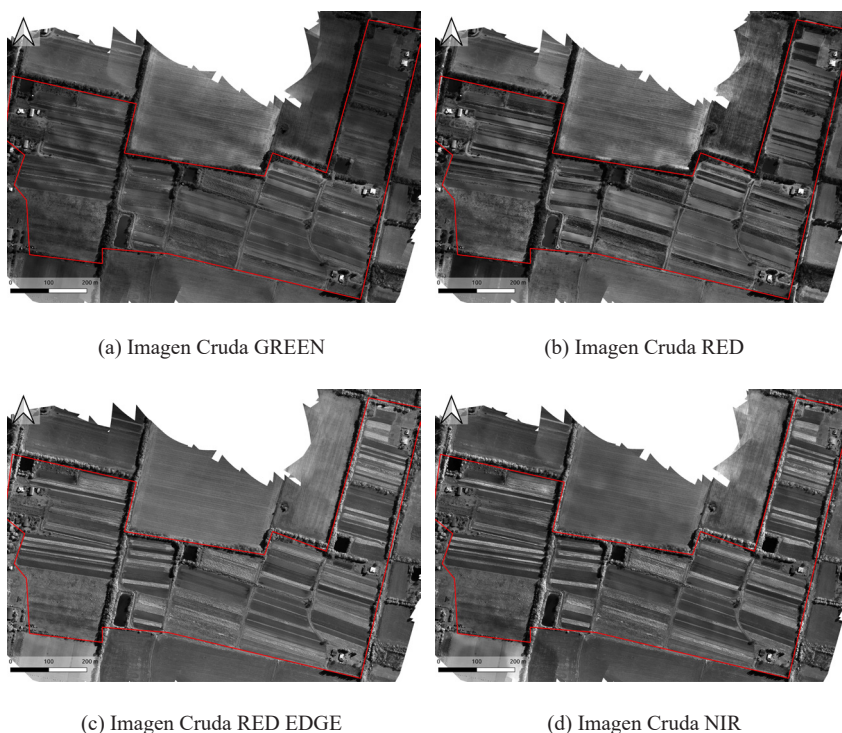


Figura 4.9: Ortomosaico del vuelo por banda, tomada por Parrot Sequoia+, sobre campos hortícola en Córdoba Capital.

4.2.4. Segmentación imagen multispectral Parrot Sequoia+

Sobre un vuelo que cubrió un área de 37,3 ha, se segmenta y clasifica un lote hortícola de 7,8 ha. Se trabajó con las imágenes ortorectificadas multispectrales (MSP) sobre la base de la arquitectura presentada en el capítulo 3. Es decir, se propone la identificación de objetos geográficamente relevantes; la caracterización de los mismos con sus estadísticos espectrales y variables morfológicas; y finalmente el etiquetado de los mismos con su clase.

El primer módulo ejecutado de la cadena de procesamientos fue la identificación automática de zonas morfológicas (a diferencia del capítulo 3, que fue manual), mediante el módulo *i.cutlines* de GRASS⁴, con un algoritmo de segmentación por detección de bordes, *Canny* (apartado 3.1.4.1.1). Este fue parametrizado de forma tal, que permitió obtener mosaicos de tamaños mayores a los tamaños de objetos que se pretendían identificar, y al mismo tiempo

⁴ <https://grass.osgeo.org/grass78/manuals/addons/i.cutlines.html>

que fuesen de un tamaño propicio para que el procesamiento de los mosaicos grandes no necesitara de mucho tiempo. Se encontraron 10 mosaicos o zonas morfológicas en el área de prueba (Figura 4.12).

Para la segmentación, a nivel de lote, se probaron distintas fuentes de datos y parámetros para optimizarla, se ejecutó un algoritmo de crecimiento y fusión de regiones Region Growing (RG) (apartado 3.1.4.1.2). Se optó por éste dado que Mean Shift (MS) presentó mayores costos computacionales en el capítulo 3, y con las imágenes obtenidas de altísima resolución espacial, la cantidad de datos era aún más grande. La optimización de los parámetros se realizó mediante el módulo *i.segment.uspo* de GRASS [148], basado en F (ver apartado 3.2.1.2.2).

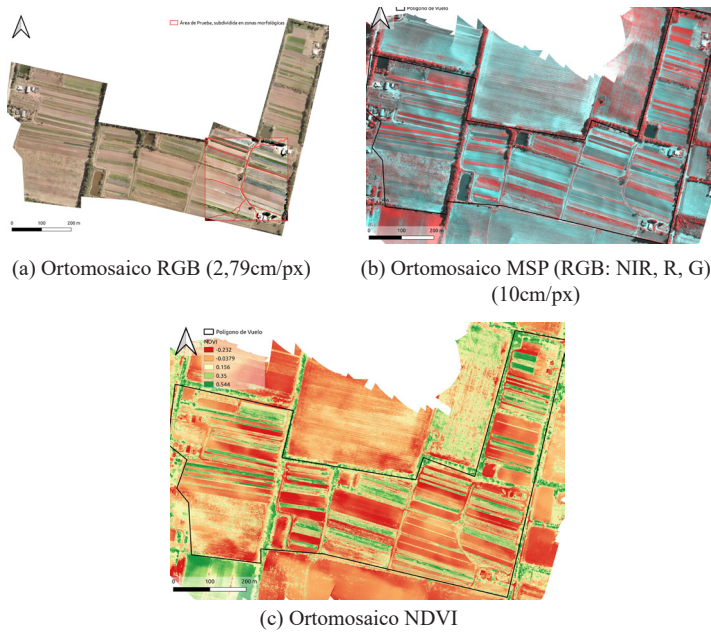


Figura 4.10: Ortomosaicos procesados en Pix4D®, de captura con Parrot Sequoia+, sobre campo hortícola en Córdoba Capital.
MSP = multiespectral; RGB=rojo, verde, azul.

(RG) (apartado 3.1.4.1.2). Se optó por éste dado que Mean Shift (MS) presentó mayores costos computacionales en el capítulo 3, y con las imágenes obtenidas de altísima resolución espacial, la cantidad de datos era aún más grande. La optimización de los parámetros se realizó mediante el módulo *i.segment.uspo* de GRASS [148], basado en F (ver apartado 3.2.1.2.2).

Los distintos grupos de fuentes de datos para segmentación fueron los siguientes:

- (a) G, R, RE, NIR
- (b) G, R, RE, NIR + NDVI + NDWI
- (c) G, R, RE, NIR + 5 medidas de textura: Contraste, Homogeneidad (IDM), Medidas de información de correlación (MOC), Suma Promedio (SA) y Suma de Varianza (SV)
- (d) RGB (sensor RGB)
- (e) G, R, RE, NIR + NDVI + NDWI (idem “b”) con menor tamaño mínimo a $30m^2$ y $80m^2$



Figura 4.11: Puntos georeferenciados de capturas de la cámara Parrot Sequoia+.

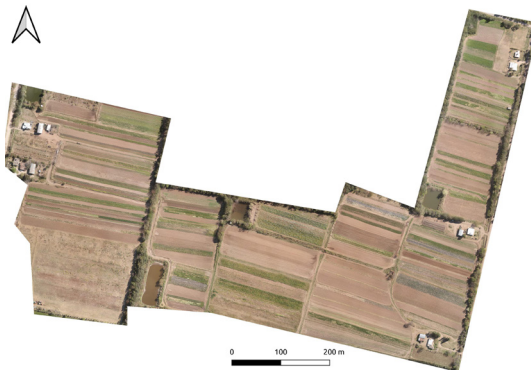


Figura 4.12: Área de vuelo sobre zona hortícola (Córdoba) y área de prueba (cuadrado rojo), subdividido en 10 zonas morfológicas. Sobre imagen RGB (2,79cm/px) de Parrot Sequoia+.

- (f) G, R, RE, NIR + NDVI + NDWI (idem “b”) con menor tamaño mínimo a $4m^2$ y $15m^2$

De las primeras pruebas preliminares realizadas con los 5 sensores, se concluyó que no pudieron usarse RGB y MSP juntas debido a una deformación geométrica de las RGB, que no permitieron una buena corrección. Por ello se procedió a trabajar con las imágenes MSP y se realizó sólo una prueba con la imagen RGB.

4.2.5. Clasificación imagen Parrot Sequoia+

4.2.6. Verdad de campo

Se realizó la georeferenciación de 691 puntos de cobertura del suelo como verdad de campo, el mismo día del vuelo. Se utilizó un GPS Garmin etrex 10 y la plataforma KoBo ToolBox⁵, mediante su aplicación KoBo Collect [162]. Los puntos, se descargaron e unificaron con la base de datos extraída de KoBo ToolBox, se digitalizaron y procesaron en QGIS 2.18.25 [102]. Se registraron todas las coberturas observadas del área de vuelo. Cada registro consta de una descripción del cultivo, especie, y condición (cosecha, estado vegetativo, porcentaje de cobertura).

4.2.7. Separabilidad de clases sobre imagen multiespectral de Parrot Sequoia+

En el área de prueba (lote hortícola) de 7,8 ha, se identificaron las coberturas de suelo presentes. La elección de clases a identificar sobre el recorte de vuelo, tiene como requisito diferenciar especies hortícolas presentes. Para verificar si las etiquetas de los puntos de verdad de campo correspondían a clases diferenciables y de relevancia para el objetivo, se analizó la separabilidad espectral de las clases de a pares, reportando las mediciones de *Jeffries-Matusita* y *Transformed Divergence* [104]. De los 121 pares de clases evaluados, solo 3 están entre 1,4 y 1,5; otros tres están entre 1,5 y 1,8; ocho entre 1,8 y 1,9 y el resto muestran una separabilidad superior a 1,9. Se establecieron 12 clases: Tierra Arada (tierra labrada); Siembras y Trasplantes; Invernadero; Construcción; Arboleda y Bordes; Caminos (suelo desnudo y planchado); Sombra; Acelga; Hinojo; Brócoli; Repollos y Poscosecha (lotes con rastrojo y restos de cosechas). En la Figura 4.13 se muestra un acercamiento en una escala 1:10 sobre la imagen RGB de cada clase planteadas y de otras especies que se observan en el vuelo, aunque no ingresen en el esquema de clasificación actual.

⁵ KoBo Toolbox es una aplicación de software de código abierto y gratuita, diseñada para realizar investigación con el apoyo de dispositivos móviles. KoBo Toolbox consta de varios programas: KoBo Form para el diseño de las encuestas, KoBo Collect para su aplicación mediante dispositivos móviles, KoBo Sync para la transferencia de las encuestas, una vez aplicadas, a la computadora para su procesamiento. Además, de otras herramientas que pueden ser consultadas en <https://www.kobotoolbox.org/>.

4.2.8. Caracterización de los segmentos

Para caracterizar los objetos de las segmentaciones elegidas, se sigue un esquema de pruebas con las siguientes fuentes de datos:

(I) G, R, RE y NIR

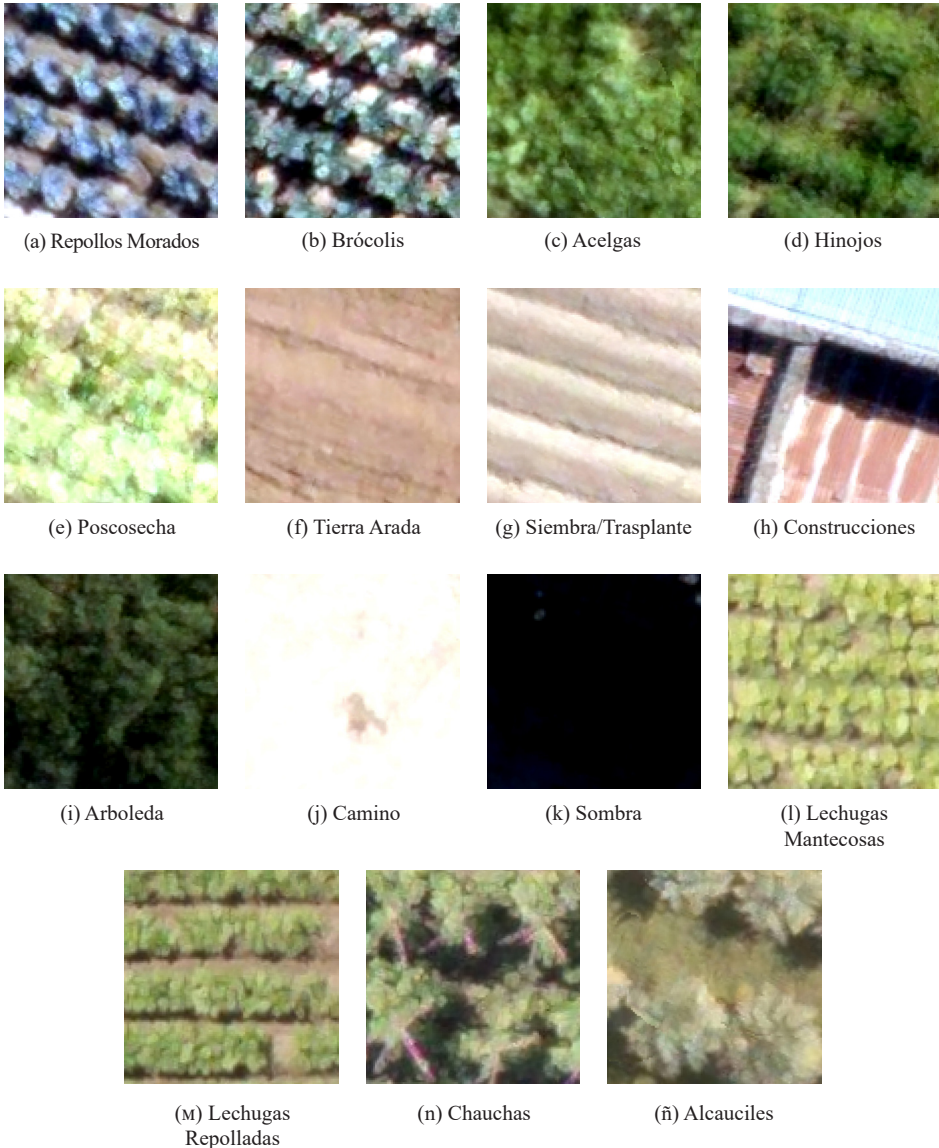


Figura 4.13: Imágenes sobre coberturas objetivos. Sensor RGB 2,79cm/pixel de cámara Parrot Sequoia+. Zoom escala 1:10. (a) *Brassica oleracea* var. *capitata* f. *rubra*, (b) *Brassica oleracea* var. *italica*, (c) *Beta vulgaris* var. *cicla*, (d) *Foeniculum vulgare* Mill. var. *dulce*, (e) Rastrojo, (f) Suelo labrado, (l) y (m) *Letuca sativa* L., (n) *Phaseolus vulgaris* L., (ñ) *Cynara scolymus* L.

(II) G, R, RE, NIR, NDVI y NDWI

(III) G, R, RE y NIR + 5 Texturas: Contraste, Homogeneidad (IDM), Medidas de información de correlación (MOC), Suma Promedio (SA) y Suma de Varianza (SV)

(IV) Aplicar Máscara de Tierra en G, R, RE, NIR. La máscara de tierra arada, se realizó mediante una clasificación no supervisada (kmeans) y se eliminaron píxeles de tierra desnuda, tierra arada, caminos, ver Figura 4.14.

La hipótesis sobre la importancia de enmascarar *tierra* (o suelo desnudo) para caracterizar los segmentos hortícolas, tiene relación con la escasa cobertura de suelo de los distintos cultivos, y la fuerte influencia espectral del suelo sobre la caracterización de los segmentos.

Sobre estas fuentes de datos se calculan los estadísticos espectrales y los descriptores morfológicos (ver apartado 3.2.1.3.2).

4.2.9. Clasificación de los segmentos

La clasificación es el último módulo de la cadena de procesamiento planteada en sección 3.2.1, en el esquema de votación planteado en el capítulo 3 se sumó *k vecinos más cercanos* (*knn* por sus siglas en inglés, k-nearest neighbors) a fin de probar otro algoritmo de aprendizaje automático no paramétrico. La mejor clasificación se eligió en función del valor de K obtenido por la validación cruzada de 10 iteraciones y 5 carpetas. Se recurrió a un esquema de segmentación en áreas pequeñas y posterior clasificación para evitar el problema de perder la representación de un objeto. Dado que para entrenar y luego validar el modelo, cada clase no puede ser representada por un objeto único. En términos generales, siempre es mejor un poco de sobre-segmentación que se corrige al clasificar, y no, sub-segmentación que no tiene corrección posible [163].

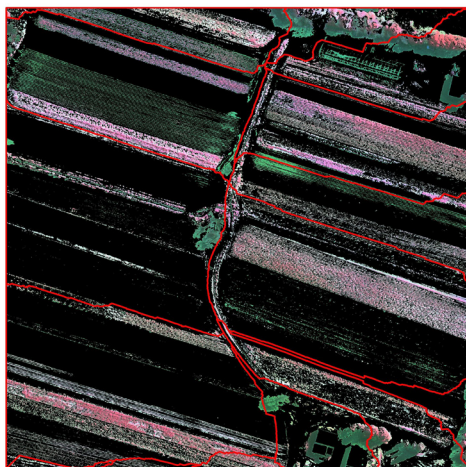


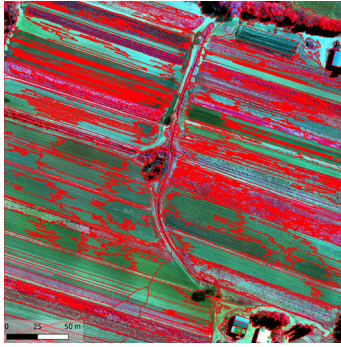
Figura 4.14: Máscara de suelo desnudo sobre imagen multiespectral Parrot Sequoia+. Composición RGB = Infrarrojo Cercano, Rojo y Verde. Zonas morfológicas demarcadas en rojo.

4.3. Resultados

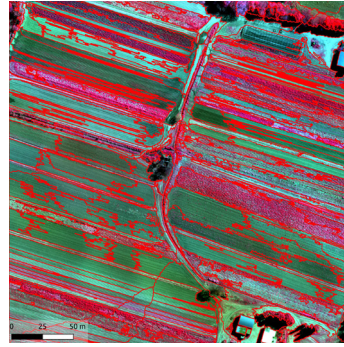
4.3.1. Segmentación imagen Parrot Sequoia+

Para la optimización de parámetros se realizaron múltiples pruebas, se presentan aquí aquellas que resultaron concluyentes y que acercan discusiones respecto de la cadena de procesamiento. Se presenta un resumen de la evaluación analítica y visual en la Tabla 4.2.

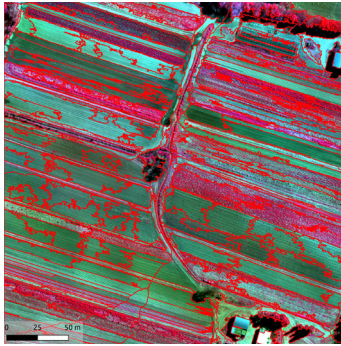
En las tres primeras pruebas presentadas, se optimizaron los parámetros de forma tal que se obtuvieran objetos de tamaños mínimos en un rango de entre $40m^2$ y $100m^2$ (Figura 4.15 (a, b y c)). En la prueba en RGB (d), se redujo el área de procesamiento y debió restringirse el tamaño de objeto requerido, dado que la capacidad de cómputo disponible era limitante, los resultados se muestran en la Figura 4.15 (d), obteniendo tamaños mínimos de $\sim 2m^2$. En las últimas dos pruebas presentadas, el tamaño mínimo de objeto se redujo, dada la necesidad de tener una cantidad estadísticamente suficiente de objetos para entrenar y validar. Resultó de trascendental importancia considerar que los objetos en la imagen no sean los únicos de su clase (ej. Invernadero y Acelga), dado que afectó a los modelos en la clasificación (que precisan objetos de entrenamiento y otros de validación).



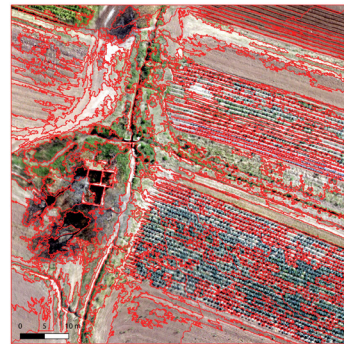
(a) 215 objetos



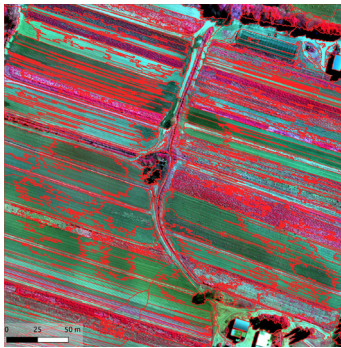
(b) 186 objetos (elegida)



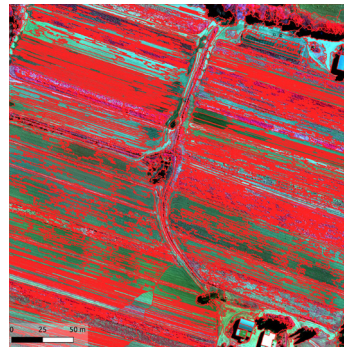
(c) 180 objetos



(d) RGB 2,79cm/px



(e) 287 objetos (elegida)



(f) 1556 objetos

Figura 4.15: Segmentaciones de imagen obtenidas por Parrot Sequoia+. Cada segmentación responde a un grupo de parámetros y fuentes de datos (inputs) diferente.

Según el análisis visual (segmentos obtenidos sobre los esperados), los resultados de la función de optimización (F de Jonhson), y especialmente del **tiempo empleado** por el ordenador en segmentar el área, se eligieron las segmentaciones **(b)** en base a [G, R, RE, NIR, NDVI, NDWI] (Figura 4.15 (b)), y la segmentación **(e)** con los mismos datos de entrada que (b) pero con un menor tamaño mínimo de objeto (Figura 4.15 (e)).

Tabla 4.2: Resultados de pruebas de segmentación con Parrot Sequoia+. Detalle de número de objetos, tiempo de ejecución, promedio de parámetros y medida de optimización. Objetos: número de objetos identificados; F de Jonhson: función de optimización y bondad general; Minsize: tamaño mínimo de objeto (m2).

Prueba	Objetos	F de Jonhson	Umbral	Minsize (m2)	Tiempo de ejecución (HH:mm:ss)	Valoración visual
RG (a)	215	0,98	0,36	55	05:23:57,5	<i>sub-segmenta</i>
RG (b)	186	0,99	0,35	60	06:27:56,2	<i>Buena</i>
RG (c)	180	0,96	0,36	53	04:40:28,3	<i>sub-segmenta</i>
RG (d)	278*	0,99	0,42	2	03:04:26	<i>sobre-segmenta</i>
RG (e)	287	0,97	0,25	39	02:22:39,2	<i>Buena</i>
RG (f)	1556	1	0,18	3,5	2:29:44,4	<i>sobre-segmenta</i>

*objetos identificados en un área de 5.000 m²

4.3.2. Clasificación imagen Parrot Sequoia+

Sobre los resultados de las segmentaciones elegidas **(b)** y **(e)**, se caracterizaron los segmentos con los 4 distintos conjuntos de datos. La primera observación que se puede hacer de las clasificaciones es que la segmentación **(b)** no es útil a la clasificación buscada, dado que se pierden dos clases presentes en la escena por el tamaño mínimo de objetos de la segmentación. Por lo que sólo se presenta una de las clasificaciones fruto de la segmentación (b) (con conjunto (I)), para su análisis visual y analítico (Figura 4.16(d) y Tabla 4.3). Las métricas globales de exactitud de las clasificaciones votadas, según las distintas fuentes para caracterizar objetos se muestran en la Tabla 4.3. Con estos resultados se descartan las clasificaciones RF(b)(I) y RF(b)(IV), por su inferior desempeño y por presentar fallas en los estadísticos de los segmentos. Las clasificaciones resultantes RF(e)(I), RF(e)(II) y RF(e)(III), se observan en la Figura 4.16 (a), (b) y (e).

Tabla 4.3: Métricas globales de precisión de las clasificaciones (más votados) según las distintas fuentes de datos de GEOBIA VANT.

Métricas de Precisión	RF(e)(I)	RF(e)(II)	RF(e)(III)	RF(e)(IV)	RF(b)(I)
OA	0,63	0,62	0,66	0,53	0,58
κ	0,56	0,55	0,6	0,45	0,48

Otra forma de evaluar las clasificaciones, fue en base métricas basadas en el poder de cada clasificación de acertar a los objetos de cada especie hortícola, es decir, no una métrica general de desempeño, sino una específica para Hortícola (OA hortícola y κ hortícola). Se presentan los resultados de las clasificaciones RF(e)I, RF(e)II, RF(e)(III) pre-seleccionadas, en la Tabla 4.4.

Tabla 4.4: Métricas de precisión hortícola, de las clasificaciones votados con distintas fuentes de datos de GEOBIA VANT.

Métricas de Precisión Hortícola	RF(e)(I)	RF(e)(II)	RF(e)(III)
OA hortícola	0,68	0,67	0,69
κ hortícola	0,57	0,56	0,58

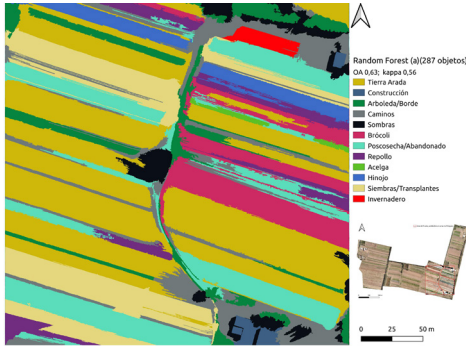
Las métricas, tanto globales como específicas de las clases hortícolas, dan mejores resultados en la clasificación **RF(e)(III)** es decir aquella surgida a partir de **G, R, RE, NIR + NDVI + NDWI** (segmentación de 287 objetos, con un tamaño mínimo de $39m^2$) y la *caracterización* de los objetos en base a **G, R, RE, NIR + 5 medidas de textura**: Contraste, Homogeneidad (IDM), Medidas de información de correlación (MOC), Suma Promedio (SA) y Suma de Varianza (SV). Se presenta la matriz de confusión de las 12 clases, surgida del re-muestreo del conjunto de entrenamiento/validación, Tabla 4.5. Los modelos de clasificación votados son RF, pero muestran resultados (OA y κ) competitivos con SVM, sin embargo las métricas de knn fueron inferiores a sus competidores en todas las pruebas (en Anexo 5 se muestran los detalles de los modelos y resultados).

Tabla 4.5: Matriz de confusión de clasificación elegida RF(e)(III). OA global 0,66; kappa 0,6. OA hortícola 0,69; kappa hortícola 0,58.

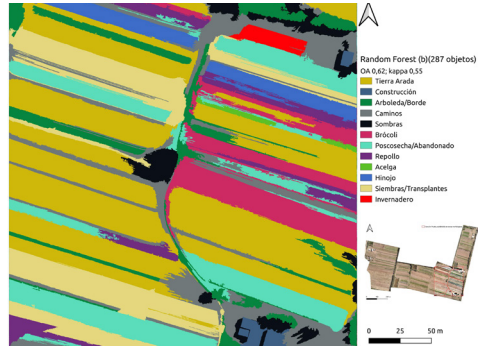
Predicción	Referencia (verdad de campo re-muestreadas)												Total Predicción
	Tierra Arada	Construcción	Arboleda	Caminos	Sombra	Brócoli	Pos-cosecha	Repollo	Acelga	Hinojo	Siembras	Invernadero	
Tierra Arada	23,6	0,4	0,6	2,3	0,3	0,0	0,0	0,0	0,0	0,0	1,2	0,7	29,1
Construcción	0,0	0,1	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2
Arboleda	0,4	0,1	6,9	2,3	0,4	0,0	0,9	0,0	0,4	0,4	0,0	0,0	11,8
Caminos	2,2	1,2	2,1	6,8	0,0	0,1	0,8	0,0	0,0	0,0	0,9	0,1	14,2
Sombra	0,0	0,0	0,0	0,4	5,5	0,0	0,0	0,0	0,0	0,0	0,3	0,0	6,2
Brócoli	0,0	0,0	0,0	0,0	0,0	4,9	0,8	1,7	0,0	0,7	0,6	0,0	8,7
Poscosecha	0,0	0,0	1,0	0,7	0,0	1,8	8,3	1,7	0,4	0,3	0,4	0,0	14,6
Repollo	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,6	0,0	0,5	0,1	0,0	1,7
Acelga	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Hinojo	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,8	0,0	0,0	0,8
Siembras	0,5	0,0	0,4	0,7	0,4	0,2	0,6	0,0	0,0	0,4	8,8	0,1	12,1
Invernadero	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Total Referencia	26,7	1,8	11,0	13,3	6,6	7,0	11,9	4,0	0,8	3,1	12,3	0,9	99,4
Recall	0,9	0,1	0,6	0,5	0,8	0,7	0,7	0,2	0,0	0,3	0,7	0,0	
Precisión	0,8	0,5	0,6	0,5	0,9	0,6	0,6	0,4	NAN	1,0	0,7	NAN	
F1 Score	0,85	0,17	0,6	0,5	0,85	0,65	0,65	0,27	NAN	0,46	0,7	NAN	

4.4. Discusión y Conclusiones

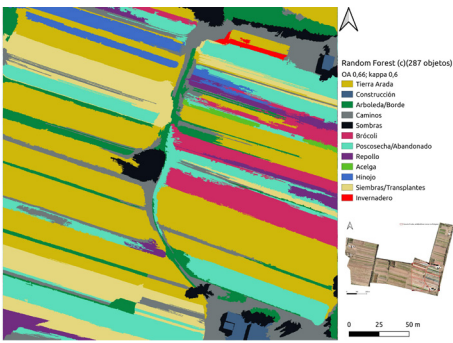
El uso de sensores de alta resolución espacial, montados en vehículos aéreos no tripulados, resulta una herramienta de gran utilidad y futuro en la identificación de especies hortícolas, y es un insumo trascendental de los inventarios de cultivos y los mapas de rendimiento. Este capítulo demuestra, que el abordaje de GEOBIA y la cadena de procesamiento presentada en el capítulo 3, es aplicable y eficiente para el procesamiento de imágenes de alta resolución espacial (VHRI), ya que disminuye el tiempo de cómputo de la clasificación, por disminución de datos a procesar. Sin embargo, debe tenerse en cuenta, el tiempo necesario para la segmentación. Éste depende directamente del tamaño de los objetos que se pretende generar, por lo que a mayor tamaño de objeto, mayor es el tiempo de cómputo necesario. Si los rangos de optimización o el conjunto de datos de entrada, son muy amplios, el tiempo de cómputo, también aumenta.



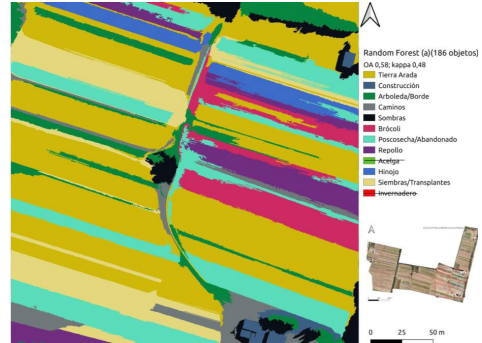
(a) RF(e)(I)



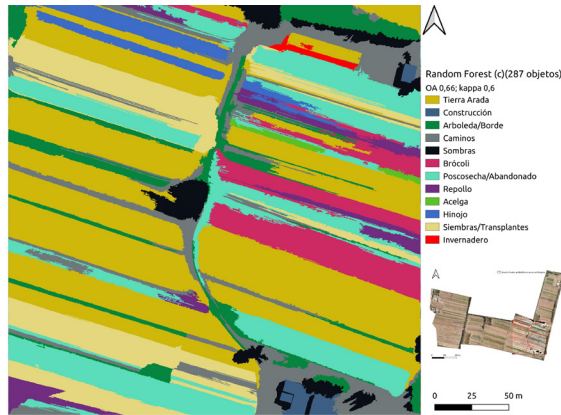
(b) RF(e)(II)



(c) RF(e)(IV)



(d) RF(b)(I)



(e) RF(e)(III) elegida

Figura 4.16: Mapas de cobertura y cultivos de lote hortícola en la ciudad de Córdoba, correspondientes a distintas clasificaciones de segmentos, en imágenes de alta resolución espacial (10cm/px), según distintas fuentes de datos.

En una búsqueda concerniente a procesamiento GEOBIA para imágenes VHRI (capturadas por un sensor en un VANT) con fines hortícolas⁶, *Scopus* arrojó sólo 14 resultados, 7 de los cuales fueron publicados en 2020, indicando un futuro promisorio de esta línea de trabajo. En el presente capítulo, se logra un desempeño de detección de cultivos hortícolas de OA Hortícola = 0,69. Específicamente, la identificación de un cultivo en período de cosecha, como Brócoli, alcanzó un $FIScore = 0,65$, operando con la una cámara Parrot Sequoia+ montada en un eBee SQ y operativizando una cadena de procesamiento basada en GEOBIA (sección 3.2.1). Con otros cultivos como Repollo, Hinojo o Acelga se debe trabajar aún más en la caracterización o correcta sincronización de fenología diferencial y sensado remoto. Sin embargo de Souza *et al.* [164], reportan una identificación en caña de azúcar, con un R^2 de 0,97 de precisión, contando con una plataforma y sensores de similares características, un esquema de OBIA en software eCognition[®] usado para diagramar las líneas de cultivos, en una etapa de pre-procesamiento; para evitar la “entre línea” no cultivada de suelo desnudo. Si bien la influencia del suelo desnudo o tierra, en la respuesta espectral de los cultivos hortícolas, es importante, el uso de una máscara de suelo (valor cero al píxel suelo desnudo), en este estudio no mejoró los resultados de segmentación ni de caracterización de los segmentos para su clasificación en clases de cultivos. Esto indica que el camino de mejora en detección de cultivos hortícolas y mapeo de rendimiento, debe continuar hacia la detección de líneas de cultivo y cálculo de biomasa.

El aporte de detección de líneas de cultivos específicos que se encuentran rodeadas de otros cultivos hortícolas, no se registra en la bibliografía. Tanto de Souza *et al.* [164], como Srivastava *et al.* [165], tienen estrategias de análisis de un solo cultivo (extensivo) [164] o de extracción y reporte de *stress* de la vegetación en general, de todo un lote [165]. Comparar las clasificaciones con métricas que evalúen la desempeño de las clasificaciones en *hortícola*, permite un avance en la identificación en contextos de horticultura de pequeña escala y diversificada. Otros, indican un contexto de horticultura, pero sólo reportan las precisiones alcanzadas de cultivos extensivos (lotes con un único cultivo) como arroz [166], o caña de azúcar [167] y rábano [168].

Si bien los datos aportados por las texturas y variables de contexto, resultaron obsoletos en el módulo de segmentación, estos arrojaron los mejores resul-

6 Búsqueda: (UAV OR DRONE) AND (GEOBIA OR OBIA) AND (horticultural OR horticulture OR vegetables).

tados al usarse para caracterizar los objetos. Es un aporte de las texturas que también se resalta en el trabajo de Dang *et al.* [168], quien con un esquema de CNN detecta daño de *Fusarium* en rábano a partir de imágenes VANT.

Las bases para seguir con las líneas de trabajo futuras están bien sentadas y las capacidades adquiridas. El camino hacia la segmentación de líneas de cultivo y plantas individuales de especies hortícolas en contextos de diversidad de variedades, junto con modelos de estimación de biomasa y la detección de stress de los cultivos, serán las bases para la construcción de mapas de rendimiento hortícola y del servicio de abastecimiento de alimentos frescos.

Capítulo 5

Conclusión

Este trabajo de tesis explora las herramientas de sensado remoto disponibles y de libre acceso para el estudio/caracterización de la actividad hortícola en la región alimentaria de proximidad a la ciudad de Córdoba (i.e., área periurbana), a fin de responder preguntas clave para el ordenamiento y planificación territorial, en donde la producción de alimentos es de trascendental importancia. La actividad hortícola como proveedora de alimentos y otros servicios ecosistémicos, es valorada en tanto se responde ¿Dónde?, ¿Qué? se produce; sentando las bases de información necesaria para responder a futuro el ¿Cuánto?, mediante modelos de rendimiento.

En primera instancia se expusieron los resultados de la búsqueda bibliográfica en torno al uso del sensado remoto para productos de interés hortícola. Se concluye que pese a que los sensores remotos tienen muchas ventajas para determinar y estudiar los principales cultivos, esos estudios no han incluido aún la mayoría de los cultivos hortícolas y el estudio e identificación de horticultura multi-específica (varias especies en un mismo lote). Es sorprendente, cómo a pesar de la importancia de la horticultura, en seguridad alimentaria, salud, los aspectos sociales y ambientales, las herramientas geoespaciales se utilizan sólo marginalmente en su gestión operativa.

Aún así, la GRS en horticultura tiene un futuro prometedor en cuanto a sensores y plataformas de mayor resolución, imágenes de libre acceso (ej. Sentinel-ESA y Landsat-NASA), sensores montados en vehículos aéreos no tripulados (VANT), la potencia del procesamiento computacional, fusión de datos, y mayor acceso a grandes bases de datos digitales (series temporales de rendimientos y condiciones meteorológicas). Sumado al contexto de la seguridad

humana con el enfoque de las necesidades alimentarias y la gestión territorial, el uso de la tecnología geoespacial en la horticultura aporta una nueva perspectiva para analizar y estudiar la horticultura.

El segundo capítulo se orientó a buscar herramientas de procesamiento de información espacial con las que responder ¿Dónde se produce?, se trabajó en particular con técnicas de clasificación para identificar uso hortícola. Este trabajo, a diferencia de los estudios locales [22] y [23], identifica y caracteriza la clase Hortícola desde un enfoque espectral, integrado a la información de contexto de cada píxel, y especialmente desde una perspectiva de detección automatizable de esta cobertura. La construcción de características de contexto con distintos tamaños de ventanas móviles, junto a la información espectral, se procesaron en un esquema de clasificación supervisada no paramétrica (RF) que arrojó un mapa de alta definición espacial (10m/px) con una precisión general de $OA = 0,89$ y un índice $\kappa = 0,86$. En dicha clasificación, se utilizó un análisis jerárquico para resolver la sobre-estimación de la superficie con cultivos hortícolas, logrando una precisión en la detección de clase Hortícola de $FIScore = 0,9$ con una superficie de 1649 ha y clase de Superficies cubiertas (uso hortícola) de $FIScore = 0,89$ en 839 ha. La sobre-estimación se debía, básicamente, a que las características espectrales y texturales de la horticultura que se mezclan con los bordes de otros tipos de coberturas y la edificación de baja densidad.

Como sub-producto de interés del capítulo 2, se deriva el proceso por el cual se identifica la zona de interfaz urbano-rural, que permite ser aplicado en otros periurbanos para detectar y cuantificar esa zona. *La Máscara de Escala de Interfaz Rural-Urbana*, permitió detectar y especializar el trabajo de identificación de los procesos que ocurren en el llamado periurbano de una ciudad. Se destaca que si bien las métricas analíticas resultan convincentes, la validación con otras fuentes e incluso la mirada experta sobre el mapa generado, es de carácter esencial para definir si una clasificación es útil o no. Este dominio experto es necesario en la evaluación de todos los productos generados.

El siguiente paso en esta trayectoria de mapeo hortícola es el uso del análisis de imágenes basado en objetos geográficos (GEOBIA) para la identificación y extracción de lotes hortícolas en la escena periurbana. En el tercer capítulo, se logró implementar una cadena de procesamiento completa, basada en software libre que incluye la segmentación, caracterización y clasifica-

ción de segmentos. Se detectaron los objetos **lotes hortícolas**, se describieron espectral, textural y morfológicamente, resultando en una clasificación con precisión general de $OA = 0,75$ y un $\kappa = 0,64$, sobre una relación muestra/población del 60%. Se concluyó, también, sobre los mejores resultados de los algoritmos de región creciente y fusión (RG) y la función de optimización (F de Jonhson). La cadena de procesamiento y su optimización resultan prometedores para la aplicación en todo el área y en otras zonas hortícolas, dado que todas las medidas de caracterización de los lotes permiten informar sobre el *parcelamiento* es decir la cantidad y tamaños de los lotes (y su modificación en el tiempo), de una zona. Este parcelamiento, es una medida que resulta importante en cuestión de distribución de recursos estratégicos y escasos como el agua de riego comunitario.

La arquitectura propuesta resulta de gran utilidad, dado que optimiza los recursos de acceso libre (imágenes y software) y es adaptable a diferentes zonas de estudio con apenas modificar algunos parámetros. La clasificación GEOBIA obtenida y su implementación automatizable, permite en principio mejorar los resultados de clasificación por pixel presentados en el capítulo 2, se compararon y el resultado visual de ambas metodologías sobre el mismo tamaño de escena. Dado que tanto OA y κ son similares, se selecciona la mejor clasificación por análisis visual. Sin embargo, los tiempos de ejecución (y capacidad de cómputo) deben ser considerados en la recomendación a futuro del mapeo de horticultura en otras regiones, así como la segmentación de los lotes con información de sensado remoto multitemporal.

En continuidad con el abordaje GEOBIA, se trabajó con imágenes de altísima resolución espacial (VHRI, con píxeles de 10 cm). A fin de aproximar una herramienta como respuesta a *¿Qué se produce?*, logrando identificar cultivos hortícolas con una precisión de 0,69 y un $\kappa_{horticola} = 0,58$. En especial Brócoli (en cosecha) alcanza un $FIScore = 0,5$. El capítulo 4 demuestra que el abordaje de GEOBIA es aplicable y eficiente en el procesamiento de VHRI y la identificación de coberturas hortícolas, disminuye el tiempo de cómputo de la clasificación por disminución de datos a procesar (paso del análisis de pixel a objeto). Sin embargo debe tenerse en cuenta que el tiempo de computo necesario para la segmentación es costoso y depende directamente del tamaño de los objetos que se pretende generar, es decir, a mayor tamaño de objeto, mayor es el tiempo de computo, lo mismo si los rangos de optimización son muy amplios.

Sobre los resultados y experiencia adquirida en las metodologías exploradas el camino hacia la segmentación de líneas de cultivo y plantas individuales de especies hortícolas, en contextos de diversidad de variedades, junto con modelos de estimación de biomasa y la detección de stress de los cultivos, serán las bases para la construcción de mapas de rendimiento hortícola y del servicio de abastecimiento de alimentos frescos. En trabajos futuros, en la aproximación al ¿qué se produce?, se incorporará la información hiperspectral para la identificación de especies hortícolas, es decir las firmas espectrales relevadas a campo de 27 especies hortícolas.

A lo largo del trabajo se han cumplimentado los objetivos específicos. Se generaron procedimientos automatizables para clasificaciones basadas en píxeles sobre la producción hortícola, incorporando características de contexto. Se logró implementar una cadena de procesamiento para generar una clasificación basada en objetos de un área hortícola con diversas especies. Finalmente se han programado vuelos con VANT, pre-procesado imágenes de altísima resolución espacial, explorado también información textural, morfológica de objetos. Así mismo, se logró un método para identificar de zona de interfaz rural-urbana (i.e., periurbanos), detección de lotes y especies de cultivos hortícolas.

Mas allá de los resultados obtenidos en términos cuantitativos, el desarrollo de la tesis ha significado un avance significativo de la autora en términos de cadenas de procesamiento automatizado con enfoques tanto de píxeles como de segmentos que serán de suma utilidad tanto en lo profesional como para las instituciones (INTA y CONAE) que dieron soporte a la realización de esta tesis.

- [1] M. E. Assessment *et al.*, *Ecosystems and human well-being*, vol. 5. Island press Washington,DC:, 2005.
- [2] Wikipedia©, “Joint photographic experts group (jpeg).” [https://es.wikipedia.org/wiki/Joint_Photographic_Experts_Group#:~:text=Joint%20Photographic%20Experts%20Group%20\(JPEG,de%20archivos%20e%20im%-C3%A1genes%20fjas.,28/06/2018](https://es.wikipedia.org/wiki/Joint_Photographic_Experts_Group#:~:text=Joint%20Photographic%20Experts%20Group%20(JPEG,de%20archivos%20e%20im%-C3%A1genes%20fjas.,28/06/2018). (Accessed on 30/10/2019).
- [3] I. S. H. S. International Society for Horticultural Science, “Defining horticulture.” <https://www.ishs.org/defining-horticulture,2019>. Accessed on 20/04/2019.
- [4] N. Ravichandra, *Horticultural nematology*, vol. 2014. Springer, 2014.
- [5] M. Altieri and C. I. Nicholls, “Teoría y práctica para una agricultura sustentable,” *Serie Textos Básicos para la Formación Ambiental*, vol. 1, 2000.
- [6] M. Alfie Cohen, “Planeación urbana y medio ambiente: los cinturones verdes,” 2011.
- [7] M. Frank, D. Drikakis, and V. Charissis, “Machine-learning methods for computational science and engineering.” *Computation*, vol. 8, no. 1, p. 15, 2020.
- [8] M. Wilmanski, C. Kreucher, and J. Lauer, “Modern approaches in deep learning for SAR ATR,” in *Algorithms for synthetic aperture radar imagery XXIII*, vol. 9843, p. 98430N, International Society for Optics and Photonics, 2016.
- [9] L. Deng, “Artificial intelligence in the rising wave of deep learning: The historical path and future outlook [perspectives],” *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 180–177, 2018.
- [10] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, “A survey of deep neural network architectures and their applications,” *Neurocomputing*, vol. 234, pp. 11–26, 2017.
- [11] E. Christophe, J. Michel, and J. Inglada, “Remote sensing processing: from multicore to GPU,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, no. 3, pp. 643–652, 2011.
- [12] M. V. Marinelli, C. M. Scavuzzo, B. L. Giobellina, and C. M. Scavuzzo, “Geoscience and remote sensing on horticulture as support for management and planning,” *Journal of Agronomy Research*, vol. 2, no. 2, p. 43, 2019.
- [13] F. A. O. of the United Nations, “Faostat database,” 2019.
- [14] Corporación del Mercado Central de Bs. As., “Mercado central de buenos aires. El mercado en números.” <http://www.mercadocentral.gob.ar/>, 2019. Actualización 2019.
- [15] Consejo Federal de Inversiones (CFI), “Informe Sector Hortícola.” <http://biblioteca.cfi.org.ar/wp-content/uploads/sites/2/2016/07/informe-horticola.pdf>, 2016.
- [16] W. H. O. World Health Organization *et al.*, “Estrategia mundial sobre régimen alimentario, actividad física y salud,” 2004.
- [17] O. P. S. Organización Panamericana de la Salud, “Plan de acción para la prevención y el control de las enfermedades no transmisibles en las americas 2013-2029,” *OPS publication*, vol. WDC USA, 2013.

- [18] I. N. D. E. C. Instituto Nacional de Estadística y Censos and S. de Gobierno de Salud de la Nación, “4^o Encuesta nacional de factores de riesgo. Resultados definitivos,” tech. rep., Gobierno de la República Argentina, Octubre de 2019.
- [19] B. L. Giobellina, “Dinámicas territoriales del cinturón verde de Córdoba,” Giobellina, B. (comp.) *La alimentación de las ciudades: transformaciones territoriales y cambio climático en el Cinturón Verde de Córdoba*. 1a ed. - Ciudad Autónoma de Buenos Aires: Ediciones INTA, 2018.
- [20] B. L. Giobellina, Giobellina, B. (comp.) *La alimentación de las ciudades: transformaciones territoriales y cambio climático en el Cinturón Verde de Córdoba*. 1a ed. - Ciudad Autónoma de Buenos Aires: Ediciones INTA, 2018.
- [21] A. Barsky, “El periurbano productivo, un espacio en constante transformación. introducción al estado del debate, con referencias al caso de Buenos Aires,” *Scripta Nova*, vol. 9, no. 194, p. 36, 2005.
- [22] A. Barchuk, L. Suez, and L. Locati, “Cobertura y uso de la tierra en el área periurbana de la ciudad de Córdoba, Argentina. aportes a la planificación territorial,” *Revista de la asociación Argentina de ecología de paisajes*, vol. 7, pp. 15–36, 2017.
- [23] N. Mari, B. Giobellina, A. Benitez, and V. Marinelli, “Mapping and characterizing the green belt of Córdoba: Land dynamics and the urban-rural transformation process,” *Journal of Agronomy and Research*, vol. 2, no. 1, pp. 29 – 46, 2019.
- [24] L. E. Lanfranconi, E. A. Tuda, M. I. Buteler, H. M. Fontán, R. Beretta, and C. W. Robledo, “Situación de contexto del área central bajo riego de la provincia de Córdoba,” 1987.
- [25] J. Fernández Lozano, “La producción de hortalizas en Argentina,” *Gerencia de Calidad y Tecnología. Mercado Central de Bs. As.*, 2012.
- [26] D. R. Easterling, G. A. Meehl, C. Parmesan, S. A. Changnon, T. R. Karl, and L. O. Mearns, “Climate extremes: observations, modeling, and impacts,” *Science*, vol. 289, no. 5487, pp. 2068–2074, 2000.
- [27] A. E. Brown, L. Zhang, T. A. McMahon, A. W. Western, and R. A. Vertessy, “A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation,” *Journal of hydrology*, vol. 310, no. 1-4, pp. 28–61, 2005.
- [28] N.W. Arnell, “Climate change and global water resources: Sres emissions and socio-economic scenarios,” *Global environmental change*, vol. 14, no. 1, pp. 31–52, 2004.
- [29] A. M. Thomson, R. A. Brown, N. J. Rosenberg, R. Srinivasan, and R. C. Izaurralde, “Climate change impacts for the conterminous USA: an integrated assessment,” *Climatic change*, vol. 69, no. 1, pp. 67–88, 2005.
- [30] C. Rosenzweig and D. Hillel, *Climate variability and the global harvest: Impacts of El Niño and other oscillations on agro-ecosystems*. Oxford University Press, 2008.
- [31] I. Serageldin, “Nurturing and nourishing the world’s poor: Important roles for horticulture in sustainable development,” in *XXVI International Horticultural Congress: Horticulture, Art and Science for Life-The Colloquia Presentations 642*, pp. 25–34, 2002.
- [32] J. M. Paruelo, M. Texeira, L. Staiano, M. Mastrángelo, L. Amdan, and F. Gallego, “An integrative index of Ecosystem Services provision based on remotely sensed data,” *Ecological Indicators*, vol. 71, pp. 145–154, 2016.
- [33] C. Baldini, *Territorio en movimiento: las transformaciones territoriales del Cinturón Hortícola Platense en los últimos 30 años*. PhD thesis, Universidad Nacional de La Plata, 2020.
- [34] S. J. Sarandón and C. C. Flores, “Agroecología: bases teóricas para el diseño y manejo de agroecosistemas sustentables,” *Colección libros de cátedra. Editorial de la Universidad Nacional de La Plata. Capítulo*, vol. 5, pp. 131–158, 2014.
- [35] N. Pérez and L. L. Vázquez, “Manejo ecológico de plagas,” CEDAR: *La Habana. Cuba*, vol. 296, 2004.

- [36] C. I. Nicholls, *Control biológico de insectos: un enfoque agroecológico*. Universidad de Antioquia, 2008.
- [37] M. Pérez and M. E. Marasas, “Servicios de regulación y prácticas de manejo: aportes para una horticultura de base agroecológica,” *Revista Ecosistemas*, vol. 22, no. 1, pp. 36–43, 2013.
- [38] O. Dubois et al., *The state of the world's land and water resources for food and agriculture: managing systems at risk*. Earthscan, 2011.
- [39] H. Elena, Y. Noé, F. Ledesma, and S. Banchemo, “Infraestructura de datos espaciales como entorno de trabajo colaborativo para caracterizar áreas de riego y sus obras afines,”
- [40] M. V. Marinelli, N. A. Mari, and C. M. Scavuzzo, “Geomatic tools for water management in a community irrigation system, Cruz del Eje, Córdoba,” *Journal of Agronomy Research*, no. 3, 2019.
- [41] R. Revelle and V. Lakshminarayana, “The ganges water machine,” *Science*, vol. 188, no. 4188, pp. 611–616, 1975.
- [42] D. R. Ponvert-Delisle Batista, “Algunas consideraciones sobre el comportamiento de la sequía agrícola en la agricultura de Cuba y el uso de imágenes por satélites en su evaluación,” *Cultivos Tropicales*, vol. 37, no. 3, pp. 22–41, 2016.
- [43] L. Meza and M. Gonzáles, “Herramientas para la adaptación y mitigación del cambio climático en el sector agropecuario, resultados del taller práctico,” tech. rep., Organización de las Naciones Unidas para la Agricultura, 2012.
- [44] Moreno, J.M., C. Laguna-Defior, V. Barros, E. Calvo Buendía, J.A. Marengo y Ú. Oswald Spring (eds.), *2020: Adaptación frente a los riesgos del cambio climático en los países iberoamericanos – Informe RIOCCADAPT*. McGraw-Hill, Madrid, España, 2020.
- [45] C. Boryan, Z. Yang, R. Mueller, and M. Craig, “Monitoring us agriculture: the us department of agriculture, national agricultural statistics service, cropland data layer program,” *Geocarto International*, vol. 26, no. 5, pp. 341–358, 2011.
- [46] V. Crnojević, P. Lugonja, B. N. Brkljać, and B. Brunet, “Classification of small agricultural fields using combined Landsat-8 and RapidEye imagery: case study of northern Serbia,” *Journal of Applied Remote Sensing*, vol. 8, no. 1, p. 083512, 2014.
- [47] V. P. Gómez, V. D. B. Medina, J. L. Bengoa, and D. A. N. García, “Accuracy assessment of a 122 classes land cover map based on Sentinel-2, Landsat 8 and Deimos-1 images and ancillary data,” in *JGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pp. 5453–5456, IEEE, 2018.
- [48] M. V. Marinelli, N. A. Mari, D. H. Pons, B. L. Giobellina, and C. Marcelo, “Spatial and spectral features for horticulture mapping,” in *III International Conference on Agro BigData and Decision Support Systems in Agriculture*, p. 37.
- [49] P. N. Rao, H. Ravishankar, U. Raj, and K. Nagajothi, “Production estimation of horticultural crops using IRS-1d Liss-III data,” *Journal of the Indian Society of Remote Sensing*, vol. 32, no. 4, pp. 393–398, 2004.
- [50] K. Usha and B. Singh, “Potential applications of remote sensing in horticulture—a review,” *Scientia Horticulturae*, vol. 153, pp. 71–83, 2013.
- [51] C. Baldini, M. E. Marasas, and A. A. Drozd, “Entre la expansión urbana y la producción de alimentos,” *Revista de la Facultad de Agronomía*, vol. 118, no. 2, pp. 031–031, 2019.
- [52] J. G. Zaehring, G. Wambugu, B. Kiteme, and S. Eckert, “How do large-scale agricultural investments affect land use and the environment on the western slopes of mount Kenya? empirical evidence based on small-scale farmers’ perceptions and remote sensing,” *Journal of environmental management*, vol. 213, pp. 79–89, 2018.
- [53] L. Johnson, D. Roczen, S. Youkhana, R. Nemani, and D. Bosch, “Mapping vineyard leaf area with multispectral satellite imagery,” *Computers and electronics in agriculture*, vol. 38, no. 1, pp. 33–44, 2003.

- [54] M. Min, W. S. Lee, T. F. Burks, J. D. Jordan, A. W. Schumann, J. K. Schueller, and H. Xie, "Design of a hyperspectral nitrogen sensing system for orange leaves," *Computers and electronics in agriculture*, vol. 63, no. 2, pp. 215–226, 2008.
- [55] H. Okamoto, Y. Suzuki, T. Kataoka, and K. Sakai, "Unified hyperspectral imaging methodology for agricultural sensing using software framework," in *International Symposium on Application of Precision Agriculture for Fruits and Vegetables 824*, pp. 49–56, 2008.
- [56] S. S. Panda, G. Hoogenboom, and J. O. Paz, "Remote sensing and geospatial technological applications for site-specific management of fruit and nut crops: a review," *Remote Sensing*, vol. 2, no. 8, pp. 1973–1997, 2010.
- [57] R. Smart, J. K. Dick, I. M. Gravett, and B. Fisher, "Canopy management to improve grape yield and wine quality-principles and practices," *South african journal of Enology and Viticulture*, vol. 11, no. 1, pp. 3–17, 1990.
- [58] N. Parsons, R. Edmondson, and Y. Song, "Image analysis and statistical modelling for measurement and quality assessment of ornamental horticulture crops in glasshouses," *Biosystems engineering*, vol. 104, no. 2, pp. 161–168, 2009.
- [59] M. O'Connell, D. Whitfield, M. Abuzar, K. Sheffield, L. McClymont, and A. McAllister, "Satellite remote sensing of crop water use in perennial horticultural crops," in *Program and abstracts Australian irrigation conference held in Sydney*, pp. 129–130, 2010.
- [60] A. Alvino and S. Marino, "Remote sensing for irrigation of horticultural crops," *Horticulturae*, vol. 3, no. 2, p. 40, 2017.
- [61] A.-K. Mahlein, E.-C. Oerke, U. Steiner, and H.-W. Dehne, "Recent advances in sensing plant diseases for precision crop protection," *European Journal of Plant Pathology*, vol. 133,
- [62] H. Nilsson, "Remote sensing and image analysis in plant pathology," *Annual review of phytopathology*, vol. 33, no. 1, pp. 489–528, 1995.
- [63] H. K. Lichtenthaler *et al.*, "Applications of chlorophyll fluorescence in stress physiology and remote sensing", *Applications of remote sensing in agriculture. Proceedings of the 48th Easter School in Agricultural Science, University of Nottingham*, 3-7 April 1989., pp. 287–305, 1990.
- [64] D. Slaughter, D. Giles, and D. Downey, "Autonomous robotic weed control systems: A review," *Computers and electronics in agriculture*, vol. 61, no. 1, pp. 63–78, 2008.
- [65] J. M. Guerrero Hernández, *Sistema de visión para agricultura de precisión: identificación en tiempo real de líneas de cultivo y malas hierbas en campos de maíz*. PhD thesis, Universidad Complutense de Madrid, 2015.
- [66] T. Mueller-Sim, M. Jenkins, J. Abel, and G. Kantor, "The robotanist: a ground-based agricultural robot for high-throughput crop phenotyping," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3634–3639, IEEE, 2017.
- [67] S. Serafino, L. B. Cicerchia, C. Russo, H. Ramón, and I. C. d. l. P. De, "Robotic data collection platform with artificial vision applied to precision agriculture," in *III International Conference on Agro BigData and Decision Support Systems in Agriculture*, p. 41.
- [68] C. N'Doume, P. Lachenaud, A. Hussard, H. Nguyen, and A. Flori, "Etude de faisabilité pour l'élaboration d'une cartographie statistique d'inventaire des vergers café et cacao en côte d'ivoire par télédétection satellitale," 1999.
- [69] G. Forkuor, C. Conrad, M. Thiel, T. Ullmann, and E. Zougrana, "Integration of optical and Synthetic Aperture Radar imagery for improving crop mapping in Northwestern Benin, West Africa," *Remote sensing*, vol. 6, no. 7, pp. 6472–6499, 2014.
- [70] J. Rouse Jr, R. Haas, J. Schell, and D. Deering, "Monitoring vegetation systems in the great plains with erts," 1974.
- [71] C. F. Jordan, "Derivation of leaf-area index from quality of light on the forest floor," *Ecology*, vol. 50, no. 4, pp. 663–666, 1969.

- [72] A. R. Huete, "A soil-adjusted vegetation index (SAVI)," *Remote sensing of environment*, vol. 25, no. 3, pp. 295–309, 1988.
- [73] K. Smith, M. Steven, and J. Colls, "Use of hyperspectral derivative ratios in the red-edge region to identify plant stress responses to gas leaks," *Remote sensing of environment*, vol. 92, no. 2, pp. 207–217, 2004.
- [74] A. Dobermann and J. Ping, "Geostatistical integration of yield monitor data and remote sensing improves yield maps," *Agronomy journal*, vol. 96, no. 1, pp. 285–297, 2004.
- [75] M. Koller and S. Upadhyaya, "Prediction of processing tomato yield using a crop growth model and remotely sensed aerial images," *Transactions of the ASAE*, vol. 48, no. 6, pp. 2335–2341, 2005.
- [76] W. Yang, D. Huang, B. Tan, J. C. Stroeve, N. V. Shabanov, Y. Knyazikhin, R. R. Nemani, and R. B. Myneni, "Analysis of leaf area index and fraction of PAR absorbed by vegetation products from the terra MODIS sensor: 2000–2005," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 7, pp. 1829–1842, 2006.
- [77] K. Beerwinkle, J. Witz, and P. Schleider, "An automated, vertical looking, X-band radar system for continuously monitoring aerial insect activity," *Transactions of the ASAE*, vol. 36, no. 3, pp. 965–970, 1993.
- [78] S. Hobbs and W. Wolf, "An airborne radar technique for studying insect migration," *Bulletin of Entomological Research*, vol. 79, no. 4, pp. 693–704, 1989.
- [79] F. Hahn, "Automatic detection of black pulp mango in a sorting system," in *2007 ASAE Annual Meeting*, p. 1, American Society of Agricultural and Biological Engineers, 2007.
- [80] P. Mondal and M. Basu, "Adoption of precision agriculture technologies in India and in some developing countries: Scope, present status and strategies," *Progress in Natural Science*, vol. 19, no. 6, pp. 659–666, 2009.
- [81] J. Palacín, T. Pallejà, M. Tresanchez, R. Sanz, J. Llorens, M. Ribes-Dasi, J. Masip, J. Arno, A. Escola, and J. R. Rosell, "Real-time tree-foliage surface estimation using a ground laser scanner," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 4, pp. 1377–1383, 2007.
- [82] J. R. Rosell, J. Llorens, R. Sanz, J. Arno, M. Ribes-Dasi, J. Masip, A. Escolà, F. Camp, F. Solanelles, F. Gràcia, et al., "Obtaining the three-dimensional structure of tree orchards from remote 2D terrestrial LIDAR scanning," *Agricultural and Forest Meteorology*, vol. 149, no. 9, pp. 1505–1515, 2009.
- [83] L. Tang, L. Tian, B. Steward, and J. Reid, "Texture-based weed classification using gabor wavelets and neural network for real-time selective herbicide applications," *Urbana*, vol. 51, p. 61801, 1999.
- [84] F. Hahn, "Actual pathogen detection: sensors and algorithms-a review," *Algorithms*, vol. 2, no. 1, pp. 301–338, 2009.
- [85] J. R. Jensen, *Remote sensing of the environment: An earth resource perspective 2/e*. Pearson Education India, 2009.
- [86] K. D. Ehsani, R., "Yield monitors for specialty crops. advanced engineering systems for specialty crops: a review of precision agriculture for water, chemical, nutrient application, and yield monitoring.," *Agric. For. Res.*, vol. 340, pp. 31–34, 2010.
- [87] T. J. Trout, L. F. Johnson, and J. Gartung, "Remote sensing of canopy cover in horticultural crops," *HortScience*, vol. 43, no. 2, pp. 333–337, 2008.
- [88] F. Díaz Terreno, "Los territorios periurbanos de Córdoba: entre lo genérico y lo específico," *Revista Iberoamericana de Urbanismo*, no. 5, pp. 65–84, 2011.
- [89] B. Giobellina and M. Quinteros, *Perspectivas de la agricultura urbana y periurbana en Córdoba. Aportes del programa Pro Huerta a la producción agroecológica de alimentos*. Ediciones INTA, O-AUPA (Observatorio de Agricultura Urbana, Periurbana y Agroecología), Córdoba
- [90] R. A. E. Real Academia Española, *Diccionario*. 2020.

- [91] Y. Céliz, “Una construcción compleja de la interfase territorial. Revisión conceptual para la generación de variables de análisis,” *Papeles de Geografía*, 2020.
- [92] R. Capitanelli, “Clima,” *Geografía física de la provincia de Córdoba*, pp. 45–138, 1979.
- [93] E. J. Tártara, J. Apezteguía, A. Roberi, M. Bocco, and O. Adib, “Caracterización de los sistemas frutihortícolas bajo riego del cinturón verde de la ciudad de Córdoba,” 1998.
- [94] J. N. Volante, A. Collado, E. B. Ferreyra, M. F. Navarro, A. Pezzola, and M. I. Puentes, *Monitoreo de la cobertura y el uso del suelo a partir de sensores remotos - Cobertura del suelo de la Republica Argentina. Año 2006-2007 (LCCS-FAO)*.
- [95] Instituto de Planificación Metropolitana IPLAM – Ministerio de Agua Ambiente y Servicios Públicos, *Zonificación del Área Metropolitana de Córdoba*. 2014.
- [96] M. Di Pace and H. C. Bartrons, *Ecología de la ciudad*. Universidad Nacional de General Sarmiento, 2004.
- [97] R Core Team, R: *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019.
- [98] E. J. C. Suárez, “Tutorial sobre máquinas de vectores soporte (SVM),” *Dpto. de Inteligencia Artificial, ETS de Ingeniería Informática, Universidad Nacional de Educación a Distancia (UNED) - Madrid (Spain)*, Última versión:16/11/2016.
- [99] D. R. Cutler, T. C. Edwards Jr, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler, “Random forests for classification in ecology,” *Ecology*, vol. 88, no. 11, pp. 2783–2792, 2007.
- [100] A. Liaw, M. Wiener, et al., “Classification and regression by randomForest,” *R news*, vol. 2, no. 3, pp. 18–22, 2002.
- [101] L. Breiman, “Random Forests,” *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [102] QGIS Development Team, *QGIS Geographic Information System*. Open Source Geospatial Foundation, 2018.
- [103] C. Sánchez and N. A. Barberis, “Caracterización del territorio centro de la provincia de Córdoba,” *Estación Experimental Agropecuaria Manfredi*. Ediciones INTA, 2013.
- [104] H. Jeffreys, “An invariant form for the prior probability in estimation problems,” *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 186, no. 1007, pp. 453–461, 1946.
- [105] J. A. Richards and J. Richards, *Remote sensing digital image analysis*, vol. 3. Springer, 1999.
- [106] Universidad de Jaen, “Clasificación no supervisada,” no. cluster 6, p. 3, 2014.
- [107] R. A. Monserud and R. Leemans, “Comparing global vegetation maps with the Kappa statistic,” *Ecological modelling*, vol. 62, no. 4, pp. 275–293, 1992.
- [108] D. M. Powers, “Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation,” 2011.
- [109] B. Zheng, S. W. Myint, P. S. Thenkabail, and R. M. Aggarwal, “A support vector machine to identify irrigated crop types using time-series Landsat NDVI data,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 34, pp. 103–112, 2015.
- [110] A. Asgarian, A. Soffianian, and S. Pourmanafi, “Crop type mapping in a highly fragmented and heterogeneous agricultural landscape: A case of central iran using multi-temporal landsat 8 imagery,” *Computers and Electronics in Agriculture*, vol. 127, pp. 531–540, 2016.
- [111] G. Pulighe and F. Lupia, “Multitemporal geospatial evaluation of urban agriculture and (non)- sustainable food self-provisioning in milan, italy,” *Sustainability*, vol. 11, no. 7, p. 1846, 2019.

- [112] D. Y. Daga, L. Zulaica, and P. Vazquez, "The peri-urban mar del plata (argentina): digital classification of land uses and analysis of transformations in the horticultural belt," *Revista Geográfica de América Central*, no. 65, pp. 175–206, 2020.
- [113] García, CL; Piumetto, M; Teich, I; Morales, H; Kindgard, A; Fuentes, ML; Bosio, MJ; Ravelo, A., *Mapas de cobertura del suelo de la provincia de Córdoba 2017/2018 - Niveles 1 a 3*. Infraestructura de Datos Espaciales de Córdoba (IDECOR), Diciembre 2018.
- [114] M. Usman, R. Liedl, M. Shahid, and A. Abbas, "Land use/land cover classification and its change detection using multi-temporal modis ndvi data," *Journal of Geographical sciences*, vol. 25, no. 12, pp. 1479–1506, 2015.
- [115] F. Vuolo, M. Neuwirth, M. Immitzer, C. Atzberger, and W.-T. Ng, "How much does multitemporal sentinel-2 data improve crop type classification?," *International journal of applied earth observation and geoinformation*, vol. 72, pp. 122–130, 2018.
- [116] Z. Li, G. Chen, and T. Zhang, "Temporal attention networks for multitemporal multisensor crop classification," *IEEE Access*, vol. 7, pp. 134677–134690, 2019.
- [117] A. Htitiou, A. Boudhar, Y. Lebrini, R. Hadria, H. Lionboui, L. Elmansouri, B. Tychon, and T. Benabdelouahab, "The performance of random forest classification based on phenological metrics derived from sentinel-2 and landsat 8 to map crop cover in an irrigated semi-arid region," *Remote Sensing in Earth Systems Sciences*, vol. 2, no. 4, pp. 208–224, 2019.
- [118] N.-T. Son, C.-F. Chen, C.-R. Chen, and H.-Y. Guo, "Classification of multitemporal sentinel-2 data for field-level monitoring of rice cropping practices in taiwan," *Advances in Space Research*, 2020.
- [119] M. Baatz, C. Hoffmann, and G. Willhauck, "Progressing from object-based to object-oriented image analysis," in *Object-Based Image Analysis*, pp. 29–42, Springer, 2008.
- [120] T. Blaschke, "Object based image analysis for remote sensing," *ISPRS journal of photogrammetry and remote sensing*, vol. 65, no. 1, pp. 2–16, 2010.
- [121] T. Blaschke, "Object based image analysis: A new paradigm in remote sensing," in *ASPRS Annual Conference*, March, pp. 24–28, 2013.
- [122] T. Su and S. Zhang, "Object-based crop classification in hetao plain using random forest," *Earth Science Informatics*, pp. 1–13, 2020.
- [123] N. A. Mari, "Diseño de un sistema de alerta y respuesta temprana a incendios de vegetación–SARTiv," Master's thesis, Universidad Nacional de Córdoba, Facultad de Matemática, Astronomía y Física y el Instituto de Altos Estudios Espaciales Mario Gulich, 2012.
- [124] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern recognition*, vol. 26, no. 9, pp. 1277–1294, 1993.
- [125] T. Grippa, M. Lennert, B. Beaumont, S. Vanhuyse, N. Stephenne, and E. Wolff, "An opensource semi-automated processing chain for urban object-based classification," *Remote Sensing*, vol. 9, no. 4, p. 358, 2017.
- [126] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Computer vision, graphics, and image processing*, vol. 29, no. 1, pp. 100–132, 1985.
- [127] Y. J. Zhang, "Evaluation and comparison of different segmentation algorithms," *Pattern recognition letters*, vol. 18, no. 10, pp. 963–974, 1997.
- [128] M. D. Hossain and D. Chen, "Segmentation for object-based image analysis (OBIA): A review of algorithms and challenges from remote sensing perspective," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 150, pp. 115–134, 2019.
- [129] T. Blaschke, S. Lang, E. Lorup, J. Strobl, and P. Zeil, "Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications," *Environmental information for planning, politics and the public*, vol. 2, pp. 555–570, 2000.

- [130] M. Belgiu and O. Csillik, "Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis," *Remote sensing of environment*, vol. 204, pp. 509–523, 2018.
- [131] T. scikit-image development team. Created using Bootstrap and Sphinx, "Comparison of segmentation and superpixel algorithms." https://scikit-image.org/docs/dev/auto_examples/segmentation/plot_segmentations. (Accessed on 09/09/2019).
- [132] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, and the scikit-image contributors, "Scikit-image: image processing in Python," *PeerJ*, vol. 2, p. e453, 6 2014.
- [133] D. Clewley, P. Bunting, J. Shepherd, S. Gillingham, N. Flood, J. Dymond, R. Lucas, J. Armston, and M. Moghaddam, "A python-based open source system for geographic object-based image analysis (GEOBIA) utilizing raster attribute tables," *Remote Sensing*, vol. 6, no. 7, pp. 6111–6135, 2014.
- [134] L. Najman and M. Schmitt, "Geodesic saliency of watershed contours and hierarchical segmentation," 1996.
- [135] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.
- [136] Wikipedia©, "Canny edge detector." https://en.wikipedia.org/wiki/Canny_edge_detector. (Accessed on 09/09/2019).
- [137] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 5, pp. 603–619, 2002.
- [138] R. Adams and L. Bischof, "Seeded region growing," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 16, no. 6, pp. 641–647, 1994.
- [139] web Procesamiento de Señales Visión por Computadora by unamdsp, "Segmentación de imágenes con el algoritmo de mean-shift." <https://fierdetregauche.wordpress.com/2009/06/19/segmentacion-de-imagenes-con-el-algoritmo-de-mean-shift/>. (Posted on junio 19, 2009).
- [140] E. Momsen and M. Metz (GSoC Mentor), "i.segment - identifies segments (objects) from imagery data." <https://grass.osgeo.org/grass76/manuals/i.segment.html>. (Accessed on 09/09/2019).
- [141] G. Espindola, G. Câmara, I. Reis, L. Bins, and A. Monteiro, "Parameter selection for regiongrowing image segmentation algorithms using spatial autocorrelation," *International Journal of Remote Sensing*, vol. 27, no. 14, pp. 3035–3040, 2006.
- [142] Y. J. Zhang, "A survey on evaluation methods for image segmentation," *Pattern recognition*, vol. 29, no. 8, pp. 1335–1346, 1996.
- [143] B. Johnson, M. Bragais, I. Endo, D. Magcale-Macandog, and P. Macandog, "Image segmentation parameter optimization considering within-and between-segment heterogeneity at multiple scale levels: Test case for mapping residential areas using Landsat imagery," *ISPRS International Journal of Geo-Information*, vol. 4, no. 4, pp. 2292–2305, 2015.
- [144] O. Bustos and A. Guerrero, "Breve introducción a la matemática de la estadística espacial," *Ensayos Matemáticos*, vol. 20, pp. 1–115, 2011.
- [145] G. Antonioli, C. Basco, M. Ceccarelli, M. Metz, and M. Lennert, "r.texture-Generate images with textural features from a raster map." <https://grass.osgeo.org/grass76/manuals/r.texture.html>. (Accessed on 09/09/2019).
- [146] GRASS Development Team, *Geographic Resources Analysis Support System (GRASS GIS) Software*. Open Source Geospatial Foundation, USA, 2018.
- [147] E. Momsen, "Grass gsoc 2012 image segmentation. repository: Grass 7, browse at: i.segment." https://grasswiki.osgeo.org/wiki/GRASS_GSoC_2012_Image_Segmentation#Region_Growing_Algorithm. (Posted on 22 December 2015).

- [148] M. Lennert, “i.segment.uspo - Unsupervised segmentation parameter optimization.” <https://grass.osgeo.org/grass78/manuals/addons/i.segment.uspo.html>. (Accessed on 09/2019).
- [149] Y. Xu and R. Goodacre, “On splitting training and validation set: A comparative study of crossvalidation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning,” *Journal of Analysis and Testing*, vol. 2, no. 3, pp. 249–262, 2018.
- [150] M. Lennert, “v.class.mlr-provides supervised support vector machine classification.” <https://grass.osgeo.org/grass78/manuals/addons/v.class.mlr.html>. (Accessed on 09/09/2019).
- [151] O. Vlachopoulos, B. Leblon, J. Wang, A. Haddadi, A. LaRocque, and G. Patterson, “Delineation of crop field areas and boundaries from uas imagery using pbia and geobia with random forest classification,” *Remote Sensing*, vol. 12, no. 16, p. 2640, 2020.
- [152] V. Lebourgeois, S. Dupuy, É. Vintrou, M. Ameline, S. Butler, and A. Bégué, “A combined random forest and obia classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated sentinel-2 time series, vhrs and dem),” *Remote Sensing*, vol. 9, no. 3, p. 259, 2017.
- [153] X. Tian, E. Chen, Z. Li, Z. B. Su, F. Ling, L. Bai, and F. Wang, “Comparison of crop classification capabilities of spaceborne multi-parameter sar data,” in 2010 *IEEE International Geoscience and Remote Sensing Symposium*, pp. 359–362, IEEE, 2010.
- [154] Y. Zhang, P. S. Thenkabail, and P. Wang, “A bibliometric profile of the remote sensing open access journal published by MDPI between 2009 and 2018,” *Remote Sensing*, vol. 11, no. 1, 2019.
- [155] A. Sargolzaei, A. Abbaspour, and C. D. Crane, “Control of cooperative unmanned aerial vehicles: Review of applications, challenges, and algorithms,” *arXiv preprint arXiv:1908.02789*, 2019.
- [156] S. Bhandari, A. Raheja, M. R. Chaichi, R. L. Green, D. Do, F. H. Pham, M. Ansari, J. G. Wolf, T. M. Sherman, and A. Espinas, “Lessons learned from UAV-based remote sensing for precision agriculture,” in 2018 *International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 458–467, IEEE, 2018.
- [157] P. Cinat, S. F. Di Gennaro, A. Berton, and A. Matese, “Comparison of unsupervised algorithms for vineyard canopy segmentation from UAV multispectral images,” *Remote Sensing*, vol. 11, no. 9, p. 1023, 2019.
- [158] R.-Y. Lee, K.-C. Chang, D.-Y. Ou, and C.-H. Hsu, “Evaluation of crop mapping on fragmented and complex slope farmlands through random forest and object-oriented analysis using unmanned aerial vehicles,” *Geocarto International*, pp. 1–18, 2019.
- [159] J. G. A. Barbedo, “A review on the use of unmanned aerial vehicles and imaging sensors for monitoring and assessing plant stresses,” *Drones*, vol. 3, no. 2, 2019.
- [160] D. C. Tsouros, S. Bibi, and P. G. Sarigiannidis, “A review on UAV-based applications for precision agriculture,” *Information*, vol. 10, no. 11, p. 349, 2019.
- [161] Pix4D, S.A., “Pix4Dmapper 4.1 user manual,” *Pix4D SA: Lausanne, Switzerland*, 2017.
- [162] H. H. Initiative, “KoBoToolbox - a free open source suite of tools for field data collection for use in challenging environments. .” <https://github.com/kobotoolbox>. (Accessed on 09/2019).
- [163] T. Su and S. Zhang, “Local and global evaluation for remote sensing image segmentation,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, pp. 256–276, 2017.
- [164] C. H. W. de Souza, R. A. C. Lamparelli, J. V. Rocha, and P. S. G. Magalhães, “Mapping skips in sugarcane fields using object-based analysis of unmanned aerial vehicle (UAV) images,” *Computers and Electronics in Agriculture*, vol. 143, pp. 49–56, 2017.
- [165] K. Srivastava, A. J. Bhutoria, J. K. Sharma, A. Sinha, and P. C. Pandey, “Uavs technology for the development of gui based application for precision agriculture and environmental research,” *Remote Sensing Applications: Society and Environment*, vol. 16, p. 100258, 2019.

- [166] Y. Pan, J. Zhang, and K. Shen, "Crop area estimation from UAV transect and MSR image data using spatial sampling method: a simulation experiment," *Procedia Environmental Sciences*, vol. 7, pp. 110–115, 2011.
- [167] I. H. Yano, J. R. Alves, W. E. Santiago, and B. J. Mederos, "Identification of weeds in sugarcane fields through images taken by UAV and random forest classifier," *IFAC-PapersOnLine*, vol. 49, no. 16, pp. 415–420, 2016.
- [168] L. M. Dang, I. H. Syed, and I. Suhyeon, "Drone agriculture imagery system for radish wilt," *Journal of Applied Remote Sensing*, vol. 11, no. 1, pp. 016006–016006, 2017.

ANEXO A

GEOBIA_Sentinel-2 Detalle de modelo de clasificación RF ii g elegida

MEJORES VALORES DE AJUSTE

svmRadial sigma C
4 5,94E-09 2

rf mtry
9 156

RESUMEN DE LOS RESULTADOS DEL REMUESTREO

Call:
summary.resamples(object = resamps.cv)

Models: svmRadial, rf

Number of resamples: 50

Accuracy	Min	1st Qu.	Median	Mean	3rd Qu.	Max.
svmRadial	0,3888889	0,5	0,5835913	0,5730123	0,6315789	0,7647059
rf	0,5	0,6875	0,7295322	0,7476835	0,8308824	0,9411765

Kappa	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
svmRadial	0,258427	0,3816517	0,4818376	0,4657396	0,5401808	0,6950673
rf	0,3793103	0,6086127	0,6660056	0,6839838	0,7898319	0,9264069

DETALLE DE MODELO DE CLASIFICACIÓN RF II G

MATRICES DE CONFUSIÓN REMUESTREADAS

svmRadial

Matriz de confusión de validación cruzada (25 veces, repetida 10 veces)

(las entradas son el promedio porcentual de los recuentos de células en las muestras)

<i>predicción</i>	<i>referencia</i>					
	Tierra Arada	Arboleda	Construcción	Extensivo	Horticultura	Uso Mixto
Tierra Arada	2,7	0	0	0,9	0,8	0
Arboleda	0	9,9	3,9	1,9	2,6	3,6
Construcción	0	1,2	6,6	0	4,2	1,6
Extensivo	1	0,1	0	22,9	1,1	0
Horticultura	6,2	0,1	6,4	0,6	14,7	3,1
Uso Mixto	0,1	0,9	0,9	0,4	1,1	0,7

Accuracy (average) : 0.5733

rf

Matriz de confusión de validación cruzada (25 veces, repetida 10 veces)

(las entradas son el promedio porcentual de los recuentos de células en las muestras)

<i>predicción</i>	<i>referencia</i>					
	Tierra Arada	Arboleda	Construcción	Extensivo	Horticultura	Uso Mixto
Tierra Arada	4,8	0	0	0	1,7	1,1
Arboleda	0	9,1	3,6	0	0,3	1,3
Construcción	0	1,3	11	0	2,3	1,7
Extensivo	1,1	0	0	25,6	0,2	0
Horticultura	3,3	0,9	3	0	19,9	0,6
Uso Mixto	0,8	0,9	0,2	1,1	0	4,2

Accuracy (average) : 0.7456

RESULTADOS DETALLADOS DE VALIDACIÓN CRUZADA

svmRadial

	sigma	C	Accuracy	Kappa	AccuracySD	KappaSD
1	5,94E-09	0,25	0,5221237	0,3808051	0,06838423	0,0874698
2	5,94E-09	0,5	0,5725585	0,4552854	0,08173289	0,1008544
3	5,94E-09	1	0,5677193	0,4529928	0,08344872	0,1016176
4	5,94E-09	2	0,5730123	0,4657396	0,08899301	0,1076732
5	5,94E-09	4	0,5422298	0,4334606	0,0948496	0,1133633
6	5,94E-09	8	0,530451	0,4212558	0,09672929	0,1164428
7	5,94E-09	16	0,5227933	0,4101468	0,09998027	0,1208227
8	5,94E-09	32	0,5179871	0,4046259	0,10008875	0,1221057
9	5,94E-09	64	0,516899	0,4039173	0,10918389	0,1321668
10	5,94E-09	128	0,5092186	0,3944213	0,11297703	0,1358779

rf

	mtry	Accuracy	Kappa	AccuracySD	KappaSD
1	2	0,7121987	0,6403756	0,08883462	0,1106472
2	21	0,7174958	0,6466286	0,0990227	0,1241936
3	40	0,7220125	0,6524857	0,09896797	0,1241856
4	60	0,7319279	0,6648298	0,10191444	0,1275436
5	79	0,729888	0,6613562	0,09963776	0,1258165
6	98	0,7320784	0,6645005	0,09782094	0,1232616
7	118	0,7265173	0,6574932	0,09397212	0,1185807
8	137	0,7358769	0,6695523	0,10885417	0,1363606
9	156	0,7476835	0,6839838	0,10088861	0,1265094
10	176	0,740219	0,6749347	0,10499386	0,1315357

ANEXO B

Detalles de Cadena de Procesamiento GEOBIA

La arquitectura o cadena de procesamiento GEOBIA propuesta para la identificación de lotes de uso hortícola y cultivos hortícolas, está escrita en código Python 2.7 para ser ejecutada en forma operativa¹. La misma incluye funciones de GRASS GIS 7.6 (56)² y R 3.4.4 (120)³. Ésta es una modificación en base a la presentada por Grippa, et al., 2017 (57). La arquitectura propuesta se muestra en el diagrama de flujo de la Figura B.1; las modificaciones radican en las nuevas entradas (*inputs*), incorporación de procesos intermedios y salidas (*outputs*), asimismo se optimiza el trabajo con vectores y se obtiene métricas y productos en los tipo de archivos deseados.

El esquema de trabajo planteado opera con software e imágenes satelitales de acceso libre (Sentinel-2, ESA) y con imágenes de alta resolución espacial (Parrot Sequoia+). El procesamiento requiere un nivel de cómputo medio de acuerdo al tamaño de escena que se pretende trabajar. Se operó en Linux en un computador con procesador intel core i7 8th Generación, con 8 GB de RAM.

1 Requiere librerías de Python *os*, *sys*, *tempfile*, *pandas*, *matplotlib*, *math*, *subprocess*, *numpy*, *grass.script*, *shutil*, *glob* y *csv*.

2 Los módulos y add-ons utilizados en GRASS son *r.neighborhoodmatrix*, *i.segment.hierarchical*, *i.segment.stats*, *i.segment.uspo*.

3 En R se utiliza la librería *caret*, *sf*, *rgdal*.

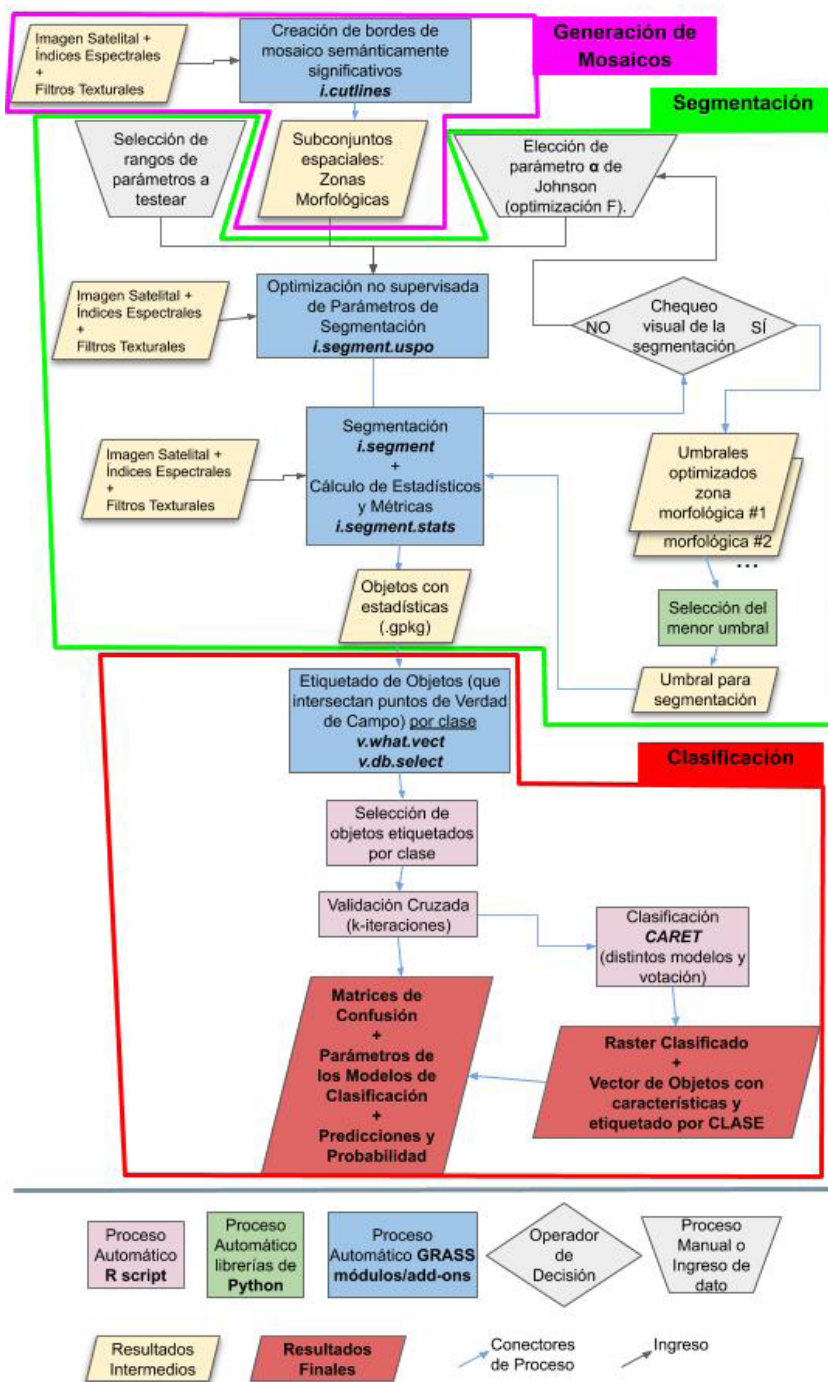


Figura B.1: Diagrama de cadena de procesamientos GEOBIA. Modificado de Grippa, et al. 2017 (57).

ANEXO C

GEOBIA_Parrot Sequoia+ Detalle de modelo de clasificación RF(e)(III) elegida

DETALLE DE MODELO DE CLASIFICACIÓN RF(E)(III)

MEJORES VALORES DE AJUSTE

svmRadial sigma C
4 0,01027441 2

Srf mtry
4 36

Sknn k
10 23

RESUMEN DE LOS RESULTADOS DEL REMUESTREO

Call:

summary.resamples(object = resamps.cv)

Models: svmRadial, rf, knn

Number of resamples: 50

Accuracy

	Min	1st Qu.	Median	Mean	3rd Qu.	Max.
svmRadial	0,5227273	0,6363636	0,673913	0,6646047	0,6971689	0,8
rf	0,5217391	0,6285389	0,6666667	0,6639564	0,7111111	0,8666667
knn	0,2826087	0,3478261	0,3913043	0,3849018	0,4222222	0,4772727

Kappa

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
svmRadial	0,4317343	0,5686724	0,6070414	0,6004039	0,6370044	0,7614841
rf	0,4368392	0,5574613	0,6022553	0,6026422	0,6621213	0,8417351
knn	0,1365188	0,1918926	0,2574179	0,2471752	0,2905503	0,3594937

MATRICES DE CONFUSIÓN REMUESTREADAS

SsvmRadial

Matriz de confusión de validación cruzada (25 veces, repetida 10 veces)
(las entradas son el promedio porcentual de los recuentos de células en las muestras)
referencia

predicción	Tierra Arada	Construcción	Arboleda	Caminos	Sombra	Brócoli	Poscosecha	Repollo	Acelga	Hinojo	Siembras	Invernadero
Tierra Arada	23,8	0,9	0,6	2,7	0,9	0	0,3	0	0	0,4	1,5	0,9
Construcción	0	0	0	0	0	0	0	0	0	0	0	0
Arboleda	0,3	0	6,4	1,2	0,9	0	0,6	0	0	0	0	0
Caminos	1,5	0,8	1,9	7,6	0	0	0,6	0	0	0	0,8	0
Sombra	0	0	0,4	0,4	4,8	0	0	0	0	0	0	0
Brócoli	0	0	0	0	0	4,7	0,7	2	0	0,8	0	0
Poscosecha	0	0	1,3	0,5	0	2,4	8,8	1,8	0,9	1	0,6	0
Repollo	0	0	0	0	0	0	0	0,1	0	0	0,2	0
Acelga	0	0	0	0	0	0	0	0	0	0	0	0
Hinojo	0	0	0	0	0	0	0	0	0	0,8	0	0
Siembras	1,1	0	0,5	0,8	0	0	1	0	0	0	9,4	0
Invernadero	0	0	0	0	0	0	0	0	0	0	0	0

Accuracy (average) : 0.6644

Srf

Matriz de confusión de validación cruzada (25 veces, repetida 10 veces)
(las entradas son el promedio porcentual de los recuentos de células en las muestras)

predicción	Tierra Arada	Construcción	Arboleda	Caminos	Sombra	Brócoli	Poscosecha	Repollo	Acelga	Hinojo	Siembras	Invernadero
Tierra Arada	23,6	0,4	0,6	2,3	0,3	0	0	0	0	0	1,2	0,7
Construcción	0	0,1	0	0,1	0	0	0	0	0	0	0	0
Arboleda	0,4	0,1	6,9	2,3	0,4	0	0,9	0	0,4	0,4	0	0
Caminos	2,2	1,2	2,1	6,8	0	0,1	0,8	0	0	0	0,9	0,1
Sombra	0	0	0	0,4	5,5	0	0	0	0	0	0,3	0
Brócoli	0	0	0	0	0	4,9	0,8	1,7	0	0,7	0,6	0
Poscosecha	0	0	1	0,7	0	1,8	8,3	1,7	0,4	0,3	0,4	0
Repollo	0	0	0	0	0	0	0,5	0,6	0	0,5	0,1	0
Acelga	0	0	0	0	0	0	0	0	0	0	0	0
Hinojo	0	0	0	0	0	0	0	0	0	0,8	0	0
Siembras	0,5	0	0,4	0,7	0,4	0,2	0,6	0	0	0,4	8,8	0,1
Invernadero	0	0	0	0	0	0	0	0	0	0	0	0

Accuracy (average) : 0.6636

Sknn

Matriz de confusión de validación cruzada (25 veces, repetida 10 veces)
 (las entradas son el promedio porcentual de los recuentos de células en las muestras)

predicción	Tierra Arada	Construcción	Arboleda	Caminos	Sombra	Brócoli	Poscosecha	Repollo	Acelga	Hinojo	Siembras	Invernadero
Tierra Arada	24,4	1,2	0,4	6,9	1,9	0	0	0	0	0,4	8,2	0,7
Construcción	0	0	0	0	0	0	0	0	0	0	0	0
Arboleda	0	0	4,7	2,3	1,2	3,8	4,4	1,4	0,1	1	1,6	0
Caminos	2	0,5	1,8	3,2	2,4	0,6	1,8	0,3	0	0	2,1	0,2
Sombra	0,3	0	0,4	0,4	0,9	0,3	0,4	0	0	0	0,3	0
Brócoli	0	0	0,3	0	0	0,1	0,3	0,2	0,1	0	0	0
Poscosecha	0	0	3,3	0,3	0,1	2,4	5	2	0,7	1,6	0,1	0
Repollo	0	0	0	0	0	0	0,1	0	0	0	0	0
Acelga	0	0	0	0	0	0	0	0	0	0	0	0
Hinojo	0	0	0	0	0	0	0	0	0	0	0	0
Siembras	0	0	0,1	0,3	0,3	0	0	0	0	0	0,1	0
Invernadero	0	0	0	0	0	0	0	0	0	0	0	0

Accuracy (average) : 0.3844

RESULTADOS DETALLADOS DE VALIDACIÓN CRUZADA

svmRadial

	sigma	C	Accuracy	Kappa	AccuracySD	KappaSD
1	0,01027441	0,25	0,5168329	0,4088477	0,04387365	0,05357096
2	0,01027441	0,5	0,578473	0,4908308	0,05012581	0,06149804
3	0,01027441	1	0,6389921	0,5677788	0,05904825	0,07059609
4	0,01027441	2	0,6646047	0,6004039	0,05235813	0,06231992
5	0,01027441	4	0,6581394	0,5934555	0,060847	0,07234401
6	0,01027441	8	0,6419107	0,5746834	0,06588503	0,07846643
7	0,01027441	16	0,6356445	0,5683144	0,06973495	0,08224671
8	0,01027441	32	0,6396877	0,5734732	0,07211077	0,08501963
9	0,01027441	64	0,6374923	0,5710308	0,07131562	0,08415508
10	0,01027441	128	0,6374923	0,5710308	0,07131562	0,08415508

Srf

	mtry	Accuracy	Kappa	AccuracySD	KappaSD
1	2	0,6533766	0,5875403	0,05901205	0,07086499
2	13	0,6599559	0,5975973	0,0593462	0,06978827
3	25	0,6604597	0,5984457	0,06443995	0,07556141
4	36	0,6639564	0,6026422	0,06667832	0,07859284
5	48	0,6629988	0,6017829	0,06377402	0,07458662
6	59	0,6576756	0,5954107	0,06591117	0,0777362
7	71	0,663022	0,6018777	0,06133327	0,07212128
8	82	0,6576545	0,5956157	0,06281565	0,07381018
9	94	0,6589396	0,5970219	0,06396763	0,07541877
10	106	0,6625888	0,6013107	0,06267161	0,07337084

Srknn

	k	Accuracy	Kappa	AccuracySD	KappaSD
1	5	0,3530078	0,2302187	0,05017917	0,05845133
2	7	0,3552502	0,2300135	0,05475378	0,06549494
3	9	0,3587197	0,2299329	0,04463883	0,053918
4	11	0,3645465	0,2347706	0,04834177	0,05788525
5	13	0,3645966	0,2326428	0,0449316	0,05363515
6	15	0,3714852	0,2403173	0,05836263	0,07001669
7	17	0,3816575	0,2500983	0,05599629	0,06893345
8	19	0,37844	0,2438807	0,04960556	0,05936706
9	21	0,3799743	0,2426636	0,04646723	0,05534747
10	23	0,3849018	0,2471752	0,04740212	0,05678864

DETALLE DE MODELO DE CLASIFICACIÓN RF(E)(III)

Acrónimos

CETT	Centro Espacial Teófilo Tabanera.
CONAE	Comisión Nacional de Actividades Espaciales.
CV	Cinturón Verde, área de producción de alimento de proximidad a la ciudad.
ESA	Agencia Espacial Europea, por sus siglas en inglés European Space Agency.
FAO	Organización de las Naciones Unidas para la Alimentación y la Agricultura, conocida como FAO por sus siglas en inglés.
GEOBIA	Análisis de imágenes basadas en objetos geográficos, por sus siglas en inglés Geographic Object Based Image Analysis.
GRS	Geociencia y Sensado Remoto, por sus siglas en inglés, geocience and remote sensing. Se utiliza como sinónimo de teledetección y sensado remoto.
IA	nteligencia Artificial.
IG	Instituto de Altos Estudios Espaciales "Mario Gulich".
INTA	Instituto Nacional de Tecnología Agropecuaria.
NDVI	Índice de vegetación de diferencia normalizada, conocido como NDVI por sus siglas en inglés.
NDWI	Índice de agua de diferencia normalizado, conocido como NDWI por sus siglas en inglés.
MAIE	Maestría en Aplicaciones de Información Espacial.
MS	Mean Shift, algoritmo de segmentación.
O-AUPA	Observatorio de Agricultura Urbana, Periurbana y Agroecología (INTA).
OMS	Organización Mundial de la Salud, conocido también como WHO por sus siglas en inglés.
RF	Random Forest, algoritmo de clasificación supervisada no-paramétrica.
RG	Region Growing, algoritmo de segmentación.
SIG	Sistemas de Información Geográfica. En inglés GIS geographic information system.
SNAP	Plataforma de aplicaciones Sentinel, SNAP por sus siglas en inglés Sentinel Application Platform.
SVM	Support Vector Machine, algoritmo de clasificación supervisada no-paramétrica.
VANT	Vehículo Aéreo No Tripulado, sinónimo “dron” o “drone”.

Glosario

Servicios Ecosistémicos

Los Servicios del ecosistema, son recursos o procesos de los ecosistemas naturales (bienes y servicios) que benefician a los seres humanos. Incluye productos como el aprovisionamiento de alimentos y procesos tales como la descomposición de desechos o la regulación biológica de plagas. Su definición fue formalizada por la Evaluación de los Ecosistemas del Milenio (EM) organizada por las Naciones Unidas en 2005. Un estudio de cuatro años que involucró a más de 1300 científicos del mundo entero. Agruparon los servicios ecosistémicos en cuatro categorías amplias: aprovisionamiento (como la producción de agua y de alimentos), regulación (control del clima y de las enfermedades), apoyo (para los ciclos de nutrientes y la polinización de cultivos) y cultural (beneficios espirituales y recreativos) [1]. Si bien vale mencionar que su definición aún está en permanente discusión en ámbitos científicos, en las esferas de la política pública y la gestión es esta la definición que se aplica.

GeoTIFF

Es un estándar de metadatos de dominio público que permite que información georreferenciada sea ajustado en un archivo de imagen de formato TIFF.

JPEG

Joint Photographic Experts Group (JPEG), es el nombre de un comité de expertos que creó un estándar de compresión y codificación de archivos e imágenes fijas. Es a menudo considerado como un formato de archivo, JPEG es el formato de imagen más común, utilizado por las cámaras fotográficas digitales y otros dispositivos de captura de imagen, junto con JPG/JFIF, que también es otro formato para el almacenamiento y la transmisión de imágenes fotográficas. Estas variaciones de formatos a menudo no se distinguen, y se llaman “JPEG”. Los archivos de este tipo se suelen nombrar con la extensión .jpg. Existen también tres variantes del estándar JPEG que comprimen la imagen sin pérdida de datos: JPEG 2000, JPEG-LS y Lossless JPEG [2].

ESRI Shapefile (SHP) ESRI Shapefile (SHP) es un formato de archivo informático propietario de datos espaciales desarrollado por la compañía ESRI, quien crea y comercializa software para Sistemas de Información Geográfica como Arc/Info o ArcGIS. Un shapefile es un formato vectorial de almacenamiento digital donde se guarda la localización de los elementos geográficos y los atributos asociados a ellos. No obstante carece de capacidad para almacenar información topológica. Es un formato multiarchivo, es decir está generado por varios ficheros informáticos.

GeoPackage Es un formato de datos abierto, no propietario, independiente de la plataforma y basado en estándares para sistemas de información geográfica implementado como un contenedor de base de datos SQLite 3 extendido (*.gpkg). Definido por Open Geospatial Consortium (OGC)¹. Es un formato de datos espaciales, de archivo único, que permite el almacenado de datos vectoriales, matrices de datos raster, atributos alfanuméricos y posibles extensiones. Podemos almacenar multitud de tipos de geometrías en un mismo archivo .gpkg: Point, Line, Polygon, MultiPoint, MultiLine, MultiPolygon, CompoundCurve, CurvedPolygon, MultiCurve y MultiSurface. GeoPackage, a diferencia de un shapefile, se ha diseñado para almacenar datos complejos y voluminosos (hasta 140 TB). Además los atributos de las geometrías pueden contener nombres muy largos, almacenando la topología. Los GeoPackages son particularmente útiles en dispositivos móviles y en entornos de comunicación en los que hay conectividad y ancho de banda limitados.²

1 <http://www.opengeospatial.org/standards/geopackage>

2 <http://www.geopackage.org/implementations.html>

Índice de tablas

Tabla 1.1.	20
Tabla 1.2.	29
Tabla 2.1.	40
Tabla 2.2.	48
Tabla 2.3.	50
Tabla 2.4.	50
Tabla 2.5.	51
Tabla 2.6.	54
Tabla 2.7.	55
Tabla 2.8.	56
Tabla 2.9.	57
Tabla 2.10.	57
Tabla 3.1.	70
Tabla 3.2.	83
Tabla 3.3.	85
Tabla 3.4.	90
Tabla 3.5.	92
Tabla 3.6.	92
Tabla 3.7.	93
Tabla 3.8.	93
Tabla 3.9.	96
Tabla 3.10.	96
Tabla 4.1.	105
Tabla 4.2.	119
Tabla 4.3.	120
Tabla 4.4.	120
Tabla 4.5.	121

Índice de Figuras

Figura 1.1: Mapas de cobertura y usos de suelo de la ciudad de Córdoba y zonas aledañas	21
Figura 1.2: Firma espectral de especies hortícolas	26
Figura 1.3: La escala y diversidad de cultivos, de los campos en el área de interfaz rural-urbano en el periurbano cordobés	30
Figura 1.4: Fotografía de campos de cultivos hortícolas diversos	32
Figura 1.5: Imagen de VANT RGB en un huerto de brócoli	33
Figura 2.1: Área de estudio en el Cinturón Verde de Córdoba	37
Figura 2.2: Características auxiliares	39
Figura 2.3: Mapa con geoposición de puntos de verdad de campo, canales funcionales del sistema de riego	44
Figura 2.4: Diagrama de flujo de trabajo para generar máscara de cobertura urbana y clasificaciones supervisadas	45
Figura 2.5: Diagrama de flujo de trabajo para generar máscara de interfaz rural-urbana y clasificaciones supervisadas	47
Figura 2.6: Diagrama de flujo de trabajo y método para la obtención del producto de clasificación jerárquica textural-espectral	49
Figura 2.7: Registro fotográfico del recorrido de Verdad de Campo	52
Figura 2.8: Zoom sobre el resultado de la clasificación Random Forest en contexto urbano	53
Figura 2.9: Clasificación Random Forest jerárquica de la máscara de interfaz rural-urbano	61
Figura 3.1: Esquema de algoritmo Mean Shift, un ciclo o iteración	71
Figura 3.2: Arquitectura modular propuesta para GEOBIA	78
Figura 3.3: Zonas Morfológicas: divide la optimización de parámetros en el proceso de segmentación	78
Figura 3.4: Medidas texturales disponibles en el módulo r.texture de GRASS sobre la zona de estudio	81
Figura 3.5: Capas de información espectral y textural generadas para segmentar y/o clasificar	82
Figura 3.6: Segmentación Manual sobre imagen Sentinel-2A	88
Figura 3.7: Resultado de pruebas de segmentación, sobre escena Sentinel-2 en periurbano de Córdoba	89
Figura 3.8: Clasificaciones votadas según cada segmentación elegida y distintas fuentes de datos	95
Figura 3.9: Clasificación GEOBIA versus clasificación Espectral-Textural de Pixel	98
Figura 4.1: Plataforma aérea de ala fija usada, eBee SQ	104
Figura 4.2: Plataforma eBee SQ en detalle	105
Figura 4.3: Gráfica representativa de reflectancia de vegetación e intersección con resolución radiométrica Parrot Sequoia+	106
Figura 4.4: eMotion®, software de gestión de vuelo de VANT	107
Figura 4.5: UAV Forecast, aplicación móvil para corroborar condiciones de vuelo con VANT	108
Figura 4.6: Equipo de trabajo, ejecución de vuelo y calibración radiométrica de los sensores multispectrales Parrot Sequoia+	109
Figura 4.7: Imagen multispectral cruda por banda, tomada por Parrot Sequoia+, sobre campo hortícola en Córdoba Capital	110
Figura 4.8: Observación de los metadatos de las imágenes del vuelo	110
Figura 4.9: Ortomosaico del vuelo por banda, tomada por Parrot Sequoia+, sobre campos hortícola en Córdoba Capital	111
Figura 4.10: Ortomosaicos procesados en Pix4D®, de captura con Parrot Sequoia+, sobre campo hortícola en Córdoba Capital	112
Figura 4.11: Puntos georeferenciados de capturas de la cámara Parrot Sequoia+	113
Figura 4.12: Área de vuelo sobre zona hortícola (Córdoba) y área de prueba (cuadrado rojo)	113
Figura 4.13: Imágenes sobre coberturas objetivos	115
Figura 4.14: Máscara de suelo desnudo sobre imagen multispectral Parrot Sequoia+	117
Figura 4.15: Segmentaciones de imagen obtenidas por Parrot Sequoia+	118
Figura 4.16: Mapas de cobertura y cultivos de lote hortícola en la ciudad de Córdoba	122
Figura B.1: Diagrama de cadena de procesamientos GEOBIA	142

El O-AUPA impulsa desde 2013 una visión integral, transdisciplinaria y multisectorial en un trabajo de Investigación-Acción Participativa que procura aportar a políticas públicas que desarrollen y protejan el área de regadío, principal abastecedora de alimentos frescos de proximidad del área Metropolitana de Córdoba. Para ello, trabaja en red con otras instituciones públicas, organizaciones, grupos de investigación del sistema público de Ciencia y Técnica, asociaciones de productores, entre otros.

Herramientas analíticas de valoración y cuantificación de la producción hortícola basada en sensores remotos es un desarrollo aplicado que estábamos necesitando para responder a la pregunta ¿Dónde y Qué se produce en el Cinturón Verde de Córdoba? Esta tesis fue defendida en la Maestría ofrecida por el Instituto Gulich-CONAE y la Facultad de Matemática, Astronomía y Física de la Universidad Nacional de Córdoba. Aborda una forma de aproximación al mapeo mediante un nuevo algoritmo del periurbano de Córdoba, donde se desarrollan múltiples actividades que representan diferentes usos de suelo. Nos brinda la primera imagen integrada de un área que en el debate académico tiene muchas formas de mencionarse: periurbano, área rururbana, interfase rural-urbana, borde o frontera urbano-rural, periferia... Un territorio en estudio que demanda nuevos marcos teóricos y herramientas para comprenderlo, monitorearlo y definir acciones sobre él en función de generar ciudades y territorios más resilientes y sustentables.



Instituto de Altos
Estudios Espaciales
"MARIO GULICH"



FAMAF
Facultad de Matemática,
Astronomía y Física



Instituto Nacional de
Tecnología Agropecuaria

Secretaría de Agricultura,
Ganadería y Pesca



Ministerio de Economía
Argentina